# Table of Contents

1. Main Page ........................................................................................................... 1
   Getting Started ........................................................................................................ 1
   XQuery Portal ........................................................................................................... 1
   Advanced User’s Guide ............................................................................................ 2

I. Getting Started ........................................................................................................ 3
   2. Getting Started ..................................................................................................... 4
      Overview ................................................................................................................ 4
   3. Startup .................................................................................................................. 6
      Getting Started ...................................................................................................... 6
      Requirements ......................................................................................................... 6
      Concurrent Operations .......................................................................................... 6
      Standalone ............................................................................................................. 6
      Graphical User Interface ....................................................................................... 7
      Client/Server ......................................................................................................... 7
      Server .................................................................................................................... 7
      Client ..................................................................................................................... 7
      HTTP Server ........................................................................................................ 8
      Changelog ............................................................................................................. 8

4. Command-Line Options ........................................................................................... 9
   Standalone ................................................................................................................ 9
   Server ....................................................................................................................... 11
   Client ....................................................................................................................... 12
   HTTP Server .......................................................................................................... 14
   GUI .......................................................................................................................... 15
   Changelog ............................................................................................................... 15

5. Start Scripts ............................................................................................................... 17
   Main Package .......................................................................................................... 17
      Windows: basex.bat .............................................................................................. 17
      Linux/Mac: basex .................................................................................................. 17
   HTTP Server ........................................................................................................... 18
      Windows: basexhttp.bat ....................................................................................... 18
      Linux/Mac: basexhttp .......................................................................................... 18
   Changelog ............................................................................................................... 18

II. User Interfaces ......................................................................................................... 20
6. Graphical User Interface ......................................................................................... 21
   Startup ...................................................................................................................... 21
   Create Database ..................................................................................................... 21
   Input Bar ................................................................................................................. 21
   Find ......................................................................................................................... 21
   XQuery ..................................................................................................................... 21
   Command ................................................................................................................. 22
   Text Editor ............................................................................................................... 22
   Project View ............................................................................................................ 22
   Realtime Options .................................................................................................... 22
   Visualizations ......................................................................................................... 22
   Look and Feels ........................................................................................................ 24
   Changelog ............................................................................................................... 24

7. Shortcuts .................................................................................................................. 25
   Editor ....................................................................................................................... 25
   Code Completions .................................................................................................. 25
   Editor Shortcuts ...................................................................................................... 25
   Changelog ............................................................................................................... 28

8. Database Server ...................................................................................................... 29
   Startup ...................................................................................................................... 29
   Working on Command-Line .................................................................................. 29
<table>
<thead>
<tr>
<th>Section</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>9. Standalone Mode</td>
<td>30</td>
</tr>
<tr>
<td>Startup</td>
<td>30</td>
</tr>
<tr>
<td>Create a Database</td>
<td>30</td>
</tr>
<tr>
<td>Execute a Query</td>
<td>30</td>
</tr>
<tr>
<td>Database Commands</td>
<td>30</td>
</tr>
<tr>
<td>Multiple Resources</td>
<td>31</td>
</tr>
<tr>
<td>Backup and Restore</td>
<td>31</td>
</tr>
<tr>
<td>10. Web Application</td>
<td>32</td>
</tr>
<tr>
<td>Servlet Container</td>
<td>32</td>
</tr>
<tr>
<td>Configuration</td>
<td>32</td>
</tr>
<tr>
<td>Available Services</td>
<td>33</td>
</tr>
<tr>
<td>Maven</td>
<td>34</td>
</tr>
<tr>
<td>Configuration</td>
<td>34</td>
</tr>
<tr>
<td>User Management</td>
<td>34</td>
</tr>
<tr>
<td>Changelog</td>
<td>34</td>
</tr>
<tr>
<td>11. DBA</td>
<td>36</td>
</tr>
<tr>
<td>Startup</td>
<td>36</td>
</tr>
<tr>
<td>First Steps</td>
<td>36</td>
</tr>
<tr>
<td>Editor</td>
<td>36</td>
</tr>
<tr>
<td>Changelog</td>
<td>37</td>
</tr>
<tr>
<td>12. Databases</td>
<td>39</td>
</tr>
<tr>
<td>Create Databases</td>
<td>39</td>
</tr>
<tr>
<td>Access Resources</td>
<td>39</td>
</tr>
<tr>
<td>XML Documents</td>
<td>39</td>
</tr>
<tr>
<td>Raw Files</td>
<td>40</td>
</tr>
<tr>
<td>HTTP Services</td>
<td>40</td>
</tr>
<tr>
<td>Update Resources</td>
<td>40</td>
</tr>
<tr>
<td>Export Data</td>
<td>41</td>
</tr>
<tr>
<td>In Memory Database</td>
<td>41</td>
</tr>
<tr>
<td>Changelog</td>
<td>41</td>
</tr>
<tr>
<td>13. Binary Data</td>
<td>43</td>
</tr>
<tr>
<td>Storage</td>
<td>43</td>
</tr>
<tr>
<td>Usage</td>
<td>43</td>
</tr>
<tr>
<td>14. Parsers</td>
<td>44</td>
</tr>
<tr>
<td>XML Parsers</td>
<td>44</td>
</tr>
<tr>
<td>GUI</td>
<td>44</td>
</tr>
<tr>
<td>Command Line</td>
<td>44</td>
</tr>
<tr>
<td>XQuery</td>
<td>44</td>
</tr>
<tr>
<td>HTML Parser</td>
<td>44</td>
</tr>
<tr>
<td>Installation</td>
<td>44</td>
</tr>
<tr>
<td>TagSoup Options</td>
<td>45</td>
</tr>
<tr>
<td>JSON Parser</td>
<td>46</td>
</tr>
<tr>
<td>GUI</td>
<td>46</td>
</tr>
<tr>
<td>Command Line</td>
<td>46</td>
</tr>
<tr>
<td>XQuery</td>
<td>46</td>
</tr>
<tr>
<td>CSV Parser</td>
<td>46</td>
</tr>
<tr>
<td>GUI</td>
<td>46</td>
</tr>
<tr>
<td>Command Line</td>
<td>46</td>
</tr>
<tr>
<td>XQuery</td>
<td>46</td>
</tr>
<tr>
<td>Text Parser</td>
<td>47</td>
</tr>
<tr>
<td>GUI</td>
<td>47</td>
</tr>
<tr>
<td>Command Line</td>
<td>47</td>
</tr>
<tr>
<td>XQuery</td>
<td>47</td>
</tr>
<tr>
<td>Changelog</td>
<td>47</td>
</tr>
<tr>
<td>15. Commands</td>
<td>48</td>
</tr>
<tr>
<td>Basics</td>
<td>48</td>
</tr>
<tr>
<td>Command Scripts</td>
<td>48</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------</td>
</tr>
<tr>
<td>General Commands</td>
<td>58</td>
</tr>
<tr>
<td>Updates</td>
<td>55</td>
</tr>
<tr>
<td>Querying</td>
<td>53</td>
</tr>
<tr>
<td>Database Operations</td>
<td>49</td>
</tr>
<tr>
<td>Administration</td>
<td>51</td>
</tr>
<tr>
<td>LCOPY</td>
<td>52</td>
</tr>
<tr>
<td>INFO DB</td>
<td>52</td>
</tr>
<tr>
<td>INFO INDEX</td>
<td>52</td>
</tr>
<tr>
<td>INFO STORAGE</td>
<td>53</td>
</tr>
<tr>
<td>Querying</td>
<td>53</td>
</tr>
<tr>
<td>LIST</td>
<td>53</td>
</tr>
<tr>
<td>XQUERY</td>
<td>53</td>
</tr>
<tr>
<td>RETRIEVE</td>
<td>53</td>
</tr>
<tr>
<td>FIND</td>
<td>54</td>
</tr>
<tr>
<td>TEST</td>
<td>54</td>
</tr>
<tr>
<td>REPO INSTALL</td>
<td>54</td>
</tr>
<tr>
<td>REPO LIST</td>
<td>54</td>
</tr>
<tr>
<td>REPO DELETE</td>
<td>54</td>
</tr>
<tr>
<td>Updates</td>
<td>55</td>
</tr>
<tr>
<td>ADD</td>
<td>55</td>
</tr>
<tr>
<td>DELETE</td>
<td>55</td>
</tr>
<tr>
<td>RENAME</td>
<td>55</td>
</tr>
<tr>
<td>REPLACE</td>
<td>56</td>
</tr>
<tr>
<td>STORE</td>
<td>56</td>
</tr>
<tr>
<td>OPTIMIZE</td>
<td>56</td>
</tr>
<tr>
<td>FLUSH</td>
<td>56</td>
</tr>
<tr>
<td>Server Administration</td>
<td>57</td>
</tr>
<tr>
<td>SHOW SESSIONS</td>
<td>57</td>
</tr>
<tr>
<td>SHOW USERS</td>
<td>57</td>
</tr>
<tr>
<td>KILL</td>
<td>57</td>
</tr>
<tr>
<td>User Management</td>
<td>57</td>
</tr>
<tr>
<td>CREATE USER</td>
<td>57</td>
</tr>
<tr>
<td>ALTER USER</td>
<td>57</td>
</tr>
<tr>
<td>ALTER PASSWORD</td>
<td>58</td>
</tr>
<tr>
<td>DROP USER</td>
<td>58</td>
</tr>
<tr>
<td>GRANT</td>
<td>58</td>
</tr>
<tr>
<td>PASSWORD</td>
<td>58</td>
</tr>
<tr>
<td>General Commands</td>
<td>58</td>
</tr>
<tr>
<td>RUN</td>
<td>58</td>
</tr>
<tr>
<td>EXECUTE</td>
<td>59</td>
</tr>
<tr>
<td>GET</td>
<td>59</td>
</tr>
<tr>
<td>SET</td>
<td>59</td>
</tr>
<tr>
<td>INFO</td>
<td>59</td>
</tr>
<tr>
<td>HELP</td>
<td>60</td>
</tr>
</tbody>
</table>
V. Query Features ................................................................. 86
19. XQuery ........................................................................ 87
20. XQuery 3.0 ................................................................. 88
   Enhanced FLWOR Expressions ......................................... 88
      group by ..................................................................... 88
      count ....................................................................... 89
      allowing empty .......................................................... 89
      window ..................................................................... 89
   Function Items ............................................................... 90
   Simple Map Operator ...................................................... 90
   Try/Catch ..................................................................... 91
   Switch ...................................................................... 91
   Expanded QNames ........................................................ 92
   Namespace Constructors ............................................... 92
   String Concatenations ................................................... 92
   External Variables ........................................................ 92
   Serialization .................................................................. 92
   Context Item ............................................................... 92
   Annotations .................................................................. 93
   Functions .................................................................... 93
   Changelog .................................................................... 93
21. Higher-Order Functions ................................................. 95
   Function Items ................................................................ 95
      Function Types .......................................................... 95
   Higher-Order Functions ................................................. 96
      Higher-Order Functions on Sequences ............................... 96
      Folds ..................................................................... 98
22. XQuery 3.1 ................................................................. 101
   Maps ........................................................................ 101
   Arrays ....................................................................... 101
      Atomization ............................................................... 102
   Lookup Operator ........................................................... 102
   Arrow Operator ............................................................. 103
   String Constructor ........................................................ 104
   Serialization ................................................................. 104
      Adaptive Serialization ................................................ 104
      JSON Serialization ..................................................... 104
   Functions .................................................................... 105
      Map Functions .......................................................... 105
      Array Functions ........................................................ 105
      JSON Functions ........................................................ 105
      fn:sort .................................................................... 106
      fn:contains-token .................................................... 106
      fn:parse-ietf-date ...................................................... 107
      fn:apply .................................................................. 107
      fn:random-number-generator .................................... 107
      fn:format-number ..................................................... 108
      fn:tokenize .............................................................. 108
      fn:trace .................................................................. 108
      fn:string-join ........................................................... 108
      fn:default-language ................................................ 108
   Binary Data .................................................................. 108
   Collations ................................................................... 108
   Enclosed Expressions .................................................... 109
   Pending Features .......................................................... 109
   Changelog ................................................................. 109
23. XQuery Extensions ...................................................... 110
   Option Declarations ..................................................... 110
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>35. Array Module</td>
<td>165</td>
</tr>
<tr>
<td>Conventions</td>
<td>165</td>
</tr>
<tr>
<td>Functions</td>
<td>165</td>
</tr>
<tr>
<td>array:size</td>
<td>165</td>
</tr>
<tr>
<td>array:get</td>
<td>165</td>
</tr>
<tr>
<td>array:append</td>
<td>165</td>
</tr>
<tr>
<td>array:subarray</td>
<td>165</td>
</tr>
<tr>
<td>array:remove</td>
<td>166</td>
</tr>
<tr>
<td>array:insert-before</td>
<td>166</td>
</tr>
<tr>
<td>array:head</td>
<td>166</td>
</tr>
<tr>
<td>array:tail</td>
<td>166</td>
</tr>
<tr>
<td>array:reverse</td>
<td>166</td>
</tr>
<tr>
<td>array:join</td>
<td>166</td>
</tr>
<tr>
<td>array:flatten</td>
<td>167</td>
</tr>
<tr>
<td>array:for-each</td>
<td>167</td>
</tr>
<tr>
<td>array:filter</td>
<td>167</td>
</tr>
<tr>
<td>array:fold-left</td>
<td>167</td>
</tr>
<tr>
<td>array:fold-right</td>
<td>167</td>
</tr>
<tr>
<td>array:for-each-pair</td>
<td>168</td>
</tr>
<tr>
<td>array:sort</td>
<td>168</td>
</tr>
<tr>
<td>Errors</td>
<td>168</td>
</tr>
<tr>
<td>Changelog</td>
<td>168</td>
</tr>
<tr>
<td>36. Binary Module</td>
<td>169</td>
</tr>
<tr>
<td>Conventions</td>
<td>169</td>
</tr>
<tr>
<td>Constants and Conversions</td>
<td>169</td>
</tr>
<tr>
<td>bin:hex</td>
<td>169</td>
</tr>
<tr>
<td>bin:bin</td>
<td>169</td>
</tr>
<tr>
<td>bin:octal</td>
<td>169</td>
</tr>
<tr>
<td>bin:to-octets</td>
<td>170</td>
</tr>
<tr>
<td>bin:from-octets</td>
<td>170</td>
</tr>
<tr>
<td>Basic Operations</td>
<td>170</td>
</tr>
<tr>
<td>bin:length</td>
<td>170</td>
</tr>
<tr>
<td>bin:part</td>
<td>170</td>
</tr>
<tr>
<td>bin:join</td>
<td>170</td>
</tr>
<tr>
<td>bin:insert-before</td>
<td>170</td>
</tr>
<tr>
<td>bin:pad-left</td>
<td>171</td>
</tr>
<tr>
<td>bin:pad-right</td>
<td>171</td>
</tr>
<tr>
<td>bin:find</td>
<td>171</td>
</tr>
<tr>
<td>Text Decoding and Encoding</td>
<td>171</td>
</tr>
<tr>
<td>bin:decode-string</td>
<td>171</td>
</tr>
<tr>
<td>bin:encode-string</td>
<td>171</td>
</tr>
<tr>
<td>Packing and Unpacking of Numeric Values</td>
<td>172</td>
</tr>
<tr>
<td>bin:pack-double</td>
<td>172</td>
</tr>
<tr>
<td>bin:pack-float</td>
<td>172</td>
</tr>
<tr>
<td>bin:pack-integer</td>
<td>172</td>
</tr>
<tr>
<td>bin:unpack-double</td>
<td>172</td>
</tr>
<tr>
<td>bin:unpack-float</td>
<td>173</td>
</tr>
<tr>
<td>bin:unpack-integer</td>
<td>173</td>
</tr>
<tr>
<td>bin:unpack-unsigned-integer</td>
<td>173</td>
</tr>
<tr>
<td>Bitwise Operations</td>
<td>173</td>
</tr>
<tr>
<td>bin:or</td>
<td>173</td>
</tr>
<tr>
<td>bin:xor</td>
<td>173</td>
</tr>
<tr>
<td>bin:and</td>
<td>174</td>
</tr>
<tr>
<td>bin:not</td>
<td>174</td>
</tr>
<tr>
<td>bin:shift</td>
<td>174</td>
</tr>
<tr>
<td>Errors</td>
<td>174</td>
</tr>
<tr>
<td>Changelog</td>
<td>174</td>
</tr>
<tr>
<td>37. Client Module</td>
<td>175</td>
</tr>
</tbody>
</table>
Fetch Module .................................................................................................................. 203
Conventions .................................................................................................................... 205
Errors ............................................................................................................................... 204
Functions ........................................................................................................................ 203
Changelog ....................................................................................................................... 201

Value Indexes .................................................................................................................. 194
Updates ............................................................................................................................. 195
Helper Functions .............................................................................................................. 200
Errors ............................................................................................................................... 201
Changelog ....................................................................................................................... 204
42. Fetch Module .................................................................................................................. 203
Conventions .................................................................................................................... 203
Functions ........................................................................................................................ 203
Errors ............................................................................................................................... 204
Changelog ....................................................................................................................... 204
43. File Module .................................................................................................................. 205
Conventions .................................................................................................................... 205
Read Operations .............................................................................................................. 205
file:list ............................................................................................................................... 205
file:children .................................................................................................................... 205
file:read-binary .............................................................................................................. 205
file:read-text .................................................................................................................. 206
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full-Text Module</td>
<td>213</td>
</tr>
<tr>
<td>General Functions</td>
<td>218</td>
</tr>
<tr>
<td>Changelog</td>
<td>217</td>
</tr>
<tr>
<td>Errors</td>
<td>217</td>
</tr>
<tr>
<td>Path Functions</td>
<td>210</td>
</tr>
<tr>
<td>File Properties</td>
<td>209</td>
</tr>
<tr>
<td>Write Operations</td>
<td>206</td>
</tr>
<tr>
<td>System Properties</td>
<td>211</td>
</tr>
<tr>
<td>Errors</td>
<td>211</td>
</tr>
<tr>
<td>Changelog</td>
<td>212</td>
</tr>
<tr>
<td>44. Full-Text Module</td>
<td>213</td>
</tr>
<tr>
<td>Conventions</td>
<td>213</td>
</tr>
<tr>
<td>Functions</td>
<td>213</td>
</tr>
<tr>
<td>Errors</td>
<td>217</td>
</tr>
<tr>
<td>Changelog</td>
<td>217</td>
</tr>
<tr>
<td>45. Geo Module</td>
<td>218</td>
</tr>
<tr>
<td>Conventions</td>
<td>218</td>
</tr>
<tr>
<td>General Functions</td>
<td>218</td>
</tr>
<tr>
<td>Errors</td>
<td>217</td>
</tr>
<tr>
<td>Changelog</td>
<td>217</td>
</tr>
<tr>
<td>geo:srid</td>
<td>219</td>
</tr>
<tr>
<td>geo:dimension</td>
<td>218</td>
</tr>
<tr>
<td>geo:geometry-type</td>
<td>218</td>
</tr>
</tbody>
</table>
geo:envelope .......................................................... 219
geo:as-text .......................................................... 219
geo:as-binary .......................................................... 219
geo:is-simple .......................................................... 220
geo:boundary .......................................................... 220
geo:contains .......................................................... 220
geo:point-on-surface ................................................. 222
Spatial Predicate Functions ........................................ 223
go:equals ............................................................... 223
go:disjoint ............................................................... 223
go:intersects ............................................................ 224
go:touches ............................................................... 224
go:crosses ............................................................... 224
go:within ............................................................... 225
go:contains ............................................................. 225
go:overlaps ............................................................. 226
go:relate ................................................................. 226
Analysis Functions ................................................ 227
go:distance ............................................................... 227
go:buffer ............................................................... 227
go:convex-hull .......................................................... 228
go:intersection .......................................................... 228
go:union ................................................................. 228
go:difference ............................................................ 229
go:sym-difference ....................................................... 229
Functions Specific to Geometry Type ................................ 229
go:x ................................................................. 229
go:y ................................................................. 230
go:z ................................................................. 230
go:start-point ........................................................... 230
go:end-point ............................................................ 230
go:is-closed ............................................................. 231
go:is-ring ................................................................ 231
go:point-n ............................................................... 231
go:exterior-ring ......................................................... 232
go:exterior-ring ......................................................... 232
go:interior-ring ........................................................ 232
go:interior-ring ........................................................ 232
Errors ................................................................. 233
Changelog ............................................................... 233
46. Hashing Module .................................................. 234
Conventions ............................................................. 234
Functions ................................................................. 234
  hash:md5 .............................................................. 234
  hash:sha1 ............................................................. 234
  hash:sha256 .......................................................... 234
  hash:hash ............................................................ 234
Errors ................................................................. 235
Changelog ............................................................... 235
47. Higher-Order Functions Module ................................ 236
Conventions ............................................................. 236
Functions ................................................................. 236
  hof:id .............................................................. 236
  hof:const ............................................................ 236
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>JsonML Format</td>
<td>256</td>
</tr>
<tr>
<td>Errors</td>
<td>257</td>
</tr>
<tr>
<td>Changelog</td>
<td>257</td>
</tr>
<tr>
<td>53. Map Module</td>
<td>259</td>
</tr>
<tr>
<td>Conventions</td>
<td>259</td>
</tr>
<tr>
<td>Functions</td>
<td>259</td>
</tr>
<tr>
<td>map:contains</td>
<td>259</td>
</tr>
<tr>
<td>map:entry</td>
<td>259</td>
</tr>
<tr>
<td>map:for-each</td>
<td>260</td>
</tr>
<tr>
<td>map:get</td>
<td>260</td>
</tr>
<tr>
<td>map:keys</td>
<td>260</td>
</tr>
<tr>
<td>map:merge</td>
<td>261</td>
</tr>
<tr>
<td>map:put</td>
<td>261</td>
</tr>
<tr>
<td>map:remove</td>
<td>261</td>
</tr>
<tr>
<td>map:size</td>
<td>261</td>
</tr>
<tr>
<td>Changelog</td>
<td>262</td>
</tr>
<tr>
<td>54. Math Module</td>
<td>263</td>
</tr>
<tr>
<td>Conventions</td>
<td>263</td>
</tr>
<tr>
<td>W3 Functions</td>
<td>263</td>
</tr>
<tr>
<td>math:pi</td>
<td>263</td>
</tr>
<tr>
<td>math:sqrt</td>
<td>263</td>
</tr>
<tr>
<td>math:asin</td>
<td>263</td>
</tr>
<tr>
<td>math:acos</td>
<td>263</td>
</tr>
<tr>
<td>math:atan</td>
<td>264</td>
</tr>
<tr>
<td>math:atan2</td>
<td>264</td>
</tr>
<tr>
<td>math:pow</td>
<td>264</td>
</tr>
<tr>
<td>math:exp</td>
<td>264</td>
</tr>
<tr>
<td>math:log</td>
<td>264</td>
</tr>
<tr>
<td>math:log10</td>
<td>264</td>
</tr>
<tr>
<td>Additional Functions</td>
<td>265</td>
</tr>
<tr>
<td>math:e</td>
<td>265</td>
</tr>
<tr>
<td>math:sinh</td>
<td>265</td>
</tr>
<tr>
<td>math:cosh</td>
<td>265</td>
</tr>
<tr>
<td>math:tanh</td>
<td>265</td>
</tr>
<tr>
<td>math:cr32</td>
<td>265</td>
</tr>
<tr>
<td>Changelog</td>
<td>265</td>
</tr>
<tr>
<td>55. Output Module</td>
<td>266</td>
</tr>
<tr>
<td>Conventions</td>
<td>266</td>
</tr>
<tr>
<td>Functions</td>
<td>266</td>
</tr>
<tr>
<td>out:nl</td>
<td>266</td>
</tr>
<tr>
<td>out:tab</td>
<td>266</td>
</tr>
<tr>
<td>out:format</td>
<td>266</td>
</tr>
<tr>
<td>Errors</td>
<td>266</td>
</tr>
<tr>
<td>Changelog</td>
<td>266</td>
</tr>
<tr>
<td>56. Process Module</td>
<td>267</td>
</tr>
<tr>
<td>Conventions</td>
<td>267</td>
</tr>
<tr>
<td>Functions</td>
<td>267</td>
</tr>
<tr>
<td>proc:system</td>
<td>267</td>
</tr>
<tr>
<td>proc:execute</td>
<td>267</td>
</tr>
<tr>
<td>proc:property</td>
<td>268</td>
</tr>
<tr>
<td>proc:property-names</td>
<td>268</td>
</tr>
<tr>
<td>Errors</td>
<td>268</td>
</tr>
<tr>
<td>Changelog</td>
<td>268</td>
</tr>
<tr>
<td>57. Profiling Module</td>
<td>269</td>
</tr>
<tr>
<td>Conventions</td>
<td>269</td>
</tr>
</tbody>
</table>
Functions ................................................................. 269
  prof:time .......................................................... 269
  prof:mem ........................................................... 269
  prof:sleep ......................................................... 269
  prof:human ....................................................... 269
  prof:dump ........................................................ 270
  prof:variables .................................................. 270
  prof:current-ms ............................................... 270
  prof:current-ns ............................................... 270
  prof:void ....................................................... 270
Changelog ............................................................ 270
58. Random Module .................................................... 272
Conventions .......................................................... 272
Functions ............................................................. 272
  random:double ................................................ 272
  random:integer ............................................... 272
  random:seeded-double ..................................... 272
  random:seeded-integer .................................. 272
  random:gaussian ............................................ 272
  random:uuid .................................................. 273
Errors ................................................................. 273
Changelog ............................................................ 273
59. Repository Module ................................................ 274
Conventions .......................................................... 274
Functions ............................................................. 274
  repo:install .................................................. 274
  repo:delete .................................................. 274
  repo:list ...................................................... 274
Errors ................................................................. 274
Changelog ............................................................ 275
60. Request Module .................................................... 276
Conventions .......................................................... 276
  General Functions ............................................. 276
    request:method ............................................ 276
    request:attribute ....................................... 276
  URI Functions .................................................. 276
    request:scheme ............................................ 276
    request:hostname ....................................... 277
    request:port .............................................. 277
    request:path ............................................. 277
    request:query ............................................ 277
    request:uri ............................................... 277
    request:context-path .................................. 277
  Connection Functions ......................................... 277
    request:address .......................................... 277
    request:remote-hostname ................................ 278
    request:remote-address ................................ 278
    request:remote-port .................................... 278
  Parameter Functions ........................................... 278
    request:parameter-names ................................ 278
    request:parameter ....................................... 278
  Header Functions ................................................ 278
    request:header-names ................................... 278
    request:header .......................................... 278
  Cookie Functions ................................................ 279
    request:cookie-names ................................... 279
    request:cookie ........................................... 279
Changelog ............................................................ 279

xvii
91. Query Mode ................................................................. 376
   Usage ........................................................................ 376
   PHP Example ............................................................ 376
   Changelog .................................................................. 377
92. Server Protocol ............................................................. 378
   Workflow ...................................................................... 378
   Transfer Protocol ......................................................... 378
   Example ...................................................................... 380
   Constructors and Functions ........................................... 381
93. Server Protocol: Types .................................................... 383
   XDM Meta Data ............................................................ 383
   Type IDs ..................................................................... 383
94. Java Examples ................................................................ 385
   Local Examples ........................................................... 385
   Server Examples .......................................................... 385
   XQuery Module Examples ............................................... 385
   Client API .................................................................... 386
   REST API ..................................................................... 386
   XML:DB API (deprecated) ............................................... 386
95. Docker ........................................................................... 387
   Running a BaseX Container ............................................. 387
   Administration with DBA ............................................... 387
   Security Considerations ............................................... 387
   Running your own Application in a Docker Image .............. 388
   Example: DBA ............................................................. 388
   Advanced Usage .......................................................... 388
X. Advanced User's Guide ...................................................... 390
96. Advanced User's Guide ..................................................... 391
97. Configuration ................................................................. 392
   Configuration Files ......................................................... 392
   Home Directory ............................................................ 392
   Database Directory ......................................................... 392
   Log Files ...................................................................... 392
   Changelog .................................................................... 393
98. Backups ......................................................................... 394
   GUI Example ............................................................... 394
   Console Example .......................................................... 394
99. Catalog Resolver ............................................................. 395
   Overview ...................................................................... 395
   XML Entity and URI Resolvers ....................................... 395
   Using other Resolvers .................................................. 396
   More Information ........................................................ 396
100. Storage Layout .............................................................. 397
    Data Types .................................................................... 397
    Database Files ............................................................. 397
    Meta Data, Name/Path/Doc Indexes: inf ......................... 397
    Node Table: tbl, tbl1 ................................................... 398
    Texts: txt, atv ........................................................... 398
    Value Indexes: txt1, txt, atvl, atvr ................................. 398
    Full-Text Fuzzy Index: ftxx, ftxy, ftxz ............................ 398
101. Node Storage ................................................................. 399
    Node Table .................................................................... 399
    PRE Value .................................................................... 399

xxii
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>403</td>
</tr>
<tr>
<td>Commands</td>
<td>402</td>
</tr>
<tr>
<td>Changelog</td>
<td>402</td>
</tr>
<tr>
<td>XQuery Update</td>
<td>403</td>
</tr>
<tr>
<td>External Side Effects</td>
<td>403</td>
</tr>
<tr>
<td>Process Locking</td>
<td>404</td>
</tr>
<tr>
<td>File-System Locks</td>
<td>405</td>
</tr>
<tr>
<td>Update Operations</td>
<td>405</td>
</tr>
<tr>
<td>Database Locks</td>
<td>405</td>
</tr>
<tr>
<td>Changelog</td>
<td>405</td>
</tr>
<tr>
<td>102. User Management</td>
<td>401</td>
</tr>
<tr>
<td>Rules</td>
<td>401</td>
</tr>
<tr>
<td>103. Transaction Management</td>
<td>403</td>
</tr>
<tr>
<td>Introduction</td>
<td>403</td>
</tr>
<tr>
<td>XQuery Update</td>
<td>403</td>
</tr>
<tr>
<td>Concurrency Control</td>
<td>403</td>
</tr>
<tr>
<td>External Side Effects</td>
<td>403</td>
</tr>
<tr>
<td>Limitations</td>
<td>404</td>
</tr>
<tr>
<td>Process Locking</td>
<td>404</td>
</tr>
<tr>
<td>File-System Locks</td>
<td>405</td>
</tr>
<tr>
<td>Update Operations</td>
<td>405</td>
</tr>
<tr>
<td>Database Locks</td>
<td>405</td>
</tr>
<tr>
<td>Changelog</td>
<td>405</td>
</tr>
<tr>
<td>104. Logging</td>
<td>406</td>
</tr>
<tr>
<td>Format</td>
<td>406</td>
</tr>
<tr>
<td>105. Statistics</td>
<td>408</td>
</tr>
<tr>
<td>Databases</td>
<td>408</td>
</tr>
<tr>
<td>Sources</td>
<td>410</td>
</tr>
<tr>
<td>106. Twitter</td>
<td>412</td>
</tr>
<tr>
<td>BaseX as Twitter Storage</td>
<td>412</td>
</tr>
<tr>
<td>Twitter’s Streaming Data</td>
<td>412</td>
</tr>
<tr>
<td>Statistics</td>
<td>412</td>
</tr>
<tr>
<td>Example Tweet (JSON)</td>
<td>413</td>
</tr>
<tr>
<td>Example Tweet (XML)</td>
<td>414</td>
</tr>
<tr>
<td>BaseX Performance</td>
<td>415</td>
</tr>
<tr>
<td>Insert with XQuery Update</td>
<td>415</td>
</tr>
<tr>
<td>XI. Use Cases</td>
<td>407</td>
</tr>
<tr>
<td>107. Block Storage</td>
<td>399</td>
</tr>
<tr>
<td>ID Value</td>
<td>399</td>
</tr>
<tr>
<td>Block Storage</td>
<td>399</td>
</tr>
<tr>
<td>108. Sources</td>
<td>410</td>
</tr>
</tbody>
</table>
Chapter 1. Main Page

Read this entry online in the BaseX Wiki.

BaseX GUI

Welcome to the documentation of BaseX!

BaseX is both a light-weight, high-performance and scalable XML Database and an XQuery 3.1 Processor with full support for the W3C Update and Full Text extensions. It focuses on storing, querying, and visualizing large XML and JSON documents and collections. A visual frontend allows users to interactively explore data and evaluate XQuery expressions in realtime. BaseX is platform-independent and distributed under the free BSD License (find more in Wikipedia).

This documentation is based on BaseX 8.4. New and updated features are highlighted and can also be searched for.

We invite you to contribute to our Wiki! If you have questions and if you are looking for a direct contact to the developer team and users, write to our basex-talk mailing list.

Getting Started

The getting started section gives you a quick introduction to BaseX. We suggest you to start with the Graphical User Interface as this is the easiest way to access your XML data, and to get an idea of how XQuery and BaseX works.

XQuery Portal

More information on using the wide range of XQuery functions and performing XPath and XQuery requests with BaseX can be found in our XQuery Portal.

Developer Section

The developer section provides useful information for developers. Here you can find information on our supported client APIs and HTTP services, and we present different ways how you can integrate BaseX into your own project.
Advanced User's Guide

Information for advanced users can be found in our advanced user's guide, which contains details on the BaseX storage, the Client/Server architecture, and some querying features.
Part I. Getting Started
Chapter 2. Getting Started

Read this entry online in the BaseX Wiki.

This page is one of the Main Sections of the documentation. It gives a quick introduction on how to start, run, and use BaseX.

Overview

First Steps

• Startup: How to get BaseX running

• Command-Line Options

User Interfaces

• Graphical User Interface (see available Shortcuts)

• Database Server: The client/server architecture

• Standalone Mode: The command-line interface

• Web Application: The HTTP server

• DBA: Browser-based database administration

General Info

• Databases: How databases are created, populated and deleted

• Parsers: How different input formats can be converted to XML

• Commands: Full overview of all database commands

• Options: Listing of all database options

Integration

• Integrating oXygen

• Integrating Eclipse

Tutorials and Slides

• BaseX for Dummies (I), written by Paul Swennenhuizen

• BaseX for Dummies (I), ZIP Package

• BaseX for Dummies (II)

• BaseX Adventures. Written by Neven Jovanović

• Tutorial. Written by Imed Bouchrika

• XQuery Tutorial; W3 Schools
• XQuery: A Guided Tour; from the book "XQuery from the Experts"

• BaseX Users Meetups; slides and examples

• Stack Overflow; questions tagged with baseX

• XQuery Summer Institute; exercises
Chapter 3. Startup

Read this entry online in the BaseX Wiki.

This article is part of the Getting Started Guide. It tells you how to get BaseX running.

Getting Started

First of all, download BaseX from our homepage. The following distributions are available:

- the **Core Package** is a JAR file, which contains the database code, the query processor and the GUI frontend. It runs completely without additional libraries.
- the **ZIP Archive** and the **Windows Installer** contain libraries for web applications and advanced features, Start Scripts, and some additional optional files.
- the **WAR Archive** can be embedded in existing Java web servers.

Some additional distributions are available from the download page, most of which contain only the core package and, optionally, scripts for starting BaseX.

BaseX is very light-weight. It can be run and used in various ways:

- as standalone application, using the **Standalone** mode or the **Graphical User Interface**, or
- as **Client/Server** application, or
- as **Web Application**, called from a web server.

It can also be embedded as a library in your own application.

Requirements

BaseX is platform-independent and runs on any system that provides an implementation of Java (JRE). Since **Version 8.0** of BaseX, Java 7 is mandatory, because it provides better file handling support, and because Oracle stopped public support for older versions.

BaseX has been tested on several platforms, including Windows (2000, XP, Vista, 7), Max OS X (10.x), Linux (SuSE xxx, Debian, Ubuntu) and OpenBSD (4.x).

Concurrent Operations

If you plan to perform concurrent read and write operations on a single database, you should use the client/server architecture or deploy it as web application. You may safely open the same database in different JVMs (Java virtual machines) for read-only access, and you won’t encounter any problems when reading from or writing to different databases, but your update operations will be rejected if the database to be written to is currently opened by another virtual machine.

Standalone

The **Standalone Mode** can be used to execute XQuery expressions or run database commands on command line. It can also be used both for scripting and batch processing your XML data. It can be started as follows (get more information on all **Startup Options**):

- Run one of the **baseX** or **baseX.bat** scripts.
- Execute the following command: java -cp BaseX.jar org.baseX.BaseX
- On Windows: Double click on the **BaseX** icon.
It is important to remember that the standalone mode does not interact with the Client/Server architecture.

**Graphical User Interface**

The GUI is the visual interface to the features of BaseX. It can be used to create new databases, perform queries or interactively explore your XML data.

It can be started as follows (get more information on all Startup Options):

- Double click on the BaseX.jar file.
- Run one of the basexgui or basexgui.bat scripts.
- Execute the following command: `java -cp BaseX.jar org.basex.BaseXGUI`
- On Windows: Double click on the BaseX GUI icon.
- For Maven users: type in `mvn exec:java` in the main directory of the basex project.

Note that the GUI does not interact with the client/server architecture.

**Client/Server**

**Server**

The Database Server comes into play if BaseX is to be used by more than one user (client). It handles concurrent read and write transactions, provides user management and logs all user interactions.

By default, the server listens to the port 1984. There are several ways of starting and stopping the server (get more information on all Startup Options):

- Run one of the basexserver or basexserver.bat scripts. Add the stop keyword to gracefully shut down the server.
- Execute the following command: `java -cp BaseX.jar org.basex.BaseXServer`. Again, the stop keyword will ensure a graceful shutdown.
- On Windows: Double click on the BaseX Server icon, which will also start the HTTP Server, or the BaseX Server (stop) icon.

Pressing Ctrl+c will close all connections and databases and shut down the server process.

**Client**

The BaseX Client interface can be used to send commands and queries to the server instance on command line. It can be started as follows (get more information on all Startup Options):

- Run one of the basexclient or basexclient.bat scripts.
- Execute the following command: `java -cp BaseX.jar org.basex.BaseXClient`
- On Windows: Double click on the BaseX Client icon.

The default admin user can be used to connect to the server:

- **Username:** admin
- **Password:** admin

The password should be changed with the PASSWORD command after the first login.
We provide additional clients in various programming languages.

HTTP Server

With the HTTP Server, BaseX can be used as Web Application. It provides access to the REST, RESTXQ and WebDAV services. An instance of the Jetty Web Server will be created, which by default listens to the port 8984. Additionally, the BaseX Server will be started, which is accessible on port 1984.

It can be started as follows (get more information on all Startup Options):

- Run one of the basexhttp or basexhttp.bat scripts. Call the script with the stop keyword to gracefully shut down the server.
- On Windows: Double click on the BaseX Server or BaseX Server (stop) icon.
- BaseX can also be deployed as web servlet.

In the first two cases, the command-line output will look similar to the following lines (the JSP warning message can be ignored):

```java
[main] INFO org.eclipse.jetty.server.AbstractConnector - Started
SelectChannelConnector@0.0.0.0:8984
HTTP Server was started (port: 8984).
```

After that, you can e.g. open your browser and navigate to the RESTXQ start page http://localhost:8984.

Changelog

Version 8.0

- Update: Switched to Java 7

Version 7.0

- Updated: BaseXJAXRX has been replaced with BaseXHTTP
Chapter 4. Command-Line Options

Read this entry online in the BaseX Wiki.

This article is part of the Getting Started Guide. It gives more details on the command-line options of all BaseX Startup modes.

Options can be specified multiple times. All options are evaluated in the given order (in earlier versions, the sequential evaluation was limited to the specified inputs, query files, queries and commands, while all other options were initially set). The standard input can be parsed by specifying a single dash (−) as argument.

Standalone

Launch the console mode

$ basex
BaseX [Standalone]
Try "help" to get more information.
> _

Available command-line flags can be listed with −h:

$ basex −h
BaseX [Standalone]
Usage: basex [−bcdiLoqrRsuvVwxXz] [input]
   [input]     Execute input file or expression
   −b< pars >  Bind external query variables
   −c< input > Execute commands from file or string
   −d          Activate debugging mode
   −i< input > Open XML file or database
   −I< input > Assign input string to context
   −o< output > Write output to file
   −q<expr >   Execute XQuery expression
   −r<num >    Set number of query executions
   −R          Turn query execution on/off
   −s< pars >  Set serialization parameter(s)
   −t[path]    Run tests in file or directory
   −u          Write updates back to original files
   −v/V        Show (all) process info
   −w          Preserve whitespaces from input files
   −x          Show query execution plan
   −X          Show query plan before/after compilation
   −z          Skip output of results

The meaning of all flags is listed in the following table. If an equivalent database option exists (which can be specified via the SET command), it is listed as well. For the examples to work escaping some characters might be necessary, depending on your Operating System.

<p>| Flag  | Description                                                                 | Option | Default | Examples               |
|-------|-----------------------------------------------------------------------------|--------|---------|                       |
| [input] | Evaluates the specified input:                                             |        |         |                       |
|       | • The input string may point to an existing file. If the file suffix is .bxs, the file content will be evaluated as Command Script; otherwise, it will be evaluated as XQuery expression. |        |         | • &quot;doc('X')//head&quot;• query.xq• commands.bxs |
|       | • Otherwise, the input string itself is evaluated as XQuery expression.     |        |         |                       |</p>
<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>-b &lt;pars&gt;</code></td>
<td>Binds external variables to XQuery expressions. This flag may be specified multiple times. Variables names and their values are delimited by equality signs (<code>=</code>). The names may be optionally prefixed with dollar signs. If a variable uses a namespace different to the default namespace, it can be specified with the Clark Notation.</td>
</tr>
<tr>
<td>-binding <code>bv=example</code></td>
<td><code>declare variable $v external; $v</code></td>
</tr>
<tr>
<td><code>b(URL)ln=value</code></td>
<td><code>declare namespace ns='URL'; declare variable $ns:ln external; $ns:ln</code></td>
</tr>
<tr>
<td><code>-c &lt;input&gt;</code></td>
<td>Executes Commands:</td>
</tr>
<tr>
<td></td>
<td>• Several commands in the input can be separated by semicolons (<code>;</code>).</td>
</tr>
<tr>
<td></td>
<td>• If the specified input is a valid file reference or URL, its content will be executed instead. Empty lines and lines starting with the number sign <code>#</code> will be ignored.</td>
</tr>
<tr>
<td><code>-d</code></td>
<td>Toggles the debugging mode. Debugging information is output to standard error.</td>
</tr>
<tr>
<td><code>-i &lt;input&gt;</code></td>
<td>Opens an XML file or database specified by the argument. The opened input can then be processed by a command or XQuery expression.</td>
</tr>
<tr>
<td><code>-I &lt;input&gt;</code></td>
<td>Assigns an input string as item of type <code>xs:untypedAtomic</code> to the query context.</td>
</tr>
<tr>
<td><code>-o &lt;file&gt;</code></td>
<td>All command and query output is written to the specified file.</td>
</tr>
<tr>
<td><code>-q &lt;expr&gt;</code></td>
<td>Executes the specified string as XQuery expression.</td>
</tr>
<tr>
<td><code>-r &lt;num&gt;</code></td>
<td>Specifies how often a specified query will be evaluated.</td>
</tr>
<tr>
<td><code>-R</code></td>
<td>Specifies if a query will be executed or parsed only.</td>
</tr>
<tr>
<td><code>-s &lt;pars&gt;</code></td>
<td>Specifies parameters for serializing XQuery results; see Serialization for more details. This flag may be specified multiple times. Key and values are separated by the equality sign (<code>=</code>).</td>
</tr>
<tr>
<td><code>-t</code></td>
<td>Runs all Unit tests in the specified file or directory.</td>
</tr>
<tr>
<td><code>-u</code></td>
<td>Propagates updates on input files back to disk.</td>
</tr>
<tr>
<td><code>-v</code></td>
<td>Prints process and timing information to the standard output.</td>
</tr>
</tbody>
</table>
Command-Line Options

- **V**  
  Prints detailed query information to the **standard output**, including details on the compilation and profiling steps.  
  **QUERYINFO**  
  **false**

- **w**  
  Specifies if whitespaces in XML text nodes should be chopped (which is the default) or preserved.  
  **CHOP**  
  **true**

- **x**  
  This flag turns on the output of the query execution plan, formatted in **XML**.  
  **XMLPLAN**  
  **false**

- **X**  
  Generates the query plan before or after query compilation. **–x** needs to be activated to make the plan visible.  
  **COMPPPLAN**  
  **true**

- **z**  
  Turns the serialization of XQuery results on/off. This flag is useful if the query is profiled or analyzed.  
  **SERIALIZE**  
  **true**

Server

Launch the server

```
$ basexserver
BaseX [Server]
Server was started (port: 1984)
```

Available command-line flags can be listed with **–h**:

```
$ basexserver -h
BaseX [Server]
Usage: basexserver [--cdinpSz] [stop]
    stop    Stop running server
    -c<cmd>  Execute initial database commands
    -d      Activate debugging mode
    -n<name> Set host the server is bound to
    -p<num>  Set server port
    -S      Start as service
    -z      Suppress logging
```

The flags have the following meaning (equivalent database options are shown in the table as well). For the examples to work escaping some characters might be necessary, depending on your Operating System.

<table>
<thead>
<tr>
<th>Flag</th>
<th>Description</th>
<th>Option</th>
<th>Default</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>stop</td>
<td>Stops a local database server instance and quits.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-c&lt;cmd&gt;</td>
<td>Launches database commands before the server itself is started. Several commands can be separated by semicolons (;).</td>
<td></td>
<td></td>
<td>–c&quot;open database;info&quot;</td>
</tr>
<tr>
<td>-d</td>
<td>Turns on the debugging mode. Debugging information is output to <strong>standard error</strong>.</td>
<td>DEBUG</td>
<td>false</td>
<td></td>
</tr>
<tr>
<td>-n&lt;name&gt;</td>
<td>Specifies the host the server will be bound to.</td>
<td>SERVERHOST</td>
<td></td>
<td>–p127.0.0.1</td>
</tr>
<tr>
<td>-p&lt;num&gt;</td>
<td>Specifies the port on which the server will be addressable.</td>
<td>SERVERPORT</td>
<td>1984</td>
<td>–p9999</td>
</tr>
<tr>
<td>-S</td>
<td>Starts the server as service (i.e., in the background).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-z</td>
<td>Does not generate any log files.</td>
<td>LOG</td>
<td>true</td>
<td></td>
</tr>
</tbody>
</table>
Multiple -c and -i flags can be specified. All other options will be set before any other operation takes place. The specified inputs, query files, queries and commands will be subsequently evaluated after that in the given order. The standard input can be parsed by specifying a single dash (-) as argument.

**Client**

Launch the console mode communicating with the server

The user name and password will be requested. The default user/password combination is **admin/admin**:

```
$ basexclient
Username: admin
Password: *****
BaseX [Client]
Try "help" to get more information.
> _
```

Available command-line flags can be listed with -h:

```
$ basexclient -h
BaseX [Client]
Usage: basexclient [-bcdilnopqrRsUvVwxXz] [input]
  [input] Execute input file or expression
  -b<pars> Bind external query variables
  -c<input> Execute commands from file or string
  -d          Activate debugging mode
  -i<input>   Open XML file or database
  -I<input>   Assign input string to context
  -n<name>    Set server (host) name
  -o<output>  Write output to file
  -p<port>    Set server port
  -P<pass>    Specify user password
  -q<expr>    Execute XQuery expression
  -r<num>     Set number of query executions
  -R          Turn query execution on/off
  -s<pars>    Set serialization parameter(s)
  -U<name>    Specify user name
  -v/V        Show (all) process info
  -w          Preserve whitespaces from input files
  -x          Show query execution plan
  -X          Show query plan before/after compilation
  -z          Skip output of results
```

The flags have the following meaning (equivalent database options are shown in the table as well). For the examples to work escaping some characters might be necessary, depending on your Operating System.

<table>
<thead>
<tr>
<th>Flag</th>
<th>Description</th>
<th>Option</th>
<th>Default</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>[input]</td>
<td>Evaluates the specified input:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The input string may point to an existing file. If the file suffix is .bxs,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>the file content will be evaluated as <strong>Command Script</strong>; otherwise, it will</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>be evaluated as XQuery expression.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Otherwise, the input string itself is evaluated as XQuery expression.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-b&lt;pars&gt;</td>
<td>Binds external variables to XQuery expressions. This flag may be specified</td>
<td></td>
<td>BINDINGS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>multiple times. Variables names and their values can be specified in the</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>format: &quot;variable=expression&quot; or &quot;variable=expression&quot;.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-b $v=example $d=example</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
values are delimited by equality signs (=). The names may be optionally prefixed with dollar signs. If a variable uses a namespace different to the default namespace, it can be specified with the Clark Notation or Expanded QName Notation.

<table>
<thead>
<tr>
<th>Command-Line Options</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>-c &lt;input&gt;</code></td>
<td>Executes Commands:</td>
</tr>
<tr>
<td>• Several commands in the input can be separated by semicolons (;).</td>
<td></td>
</tr>
<tr>
<td>• If the specified input is a valid file reference or URL, its content will be executed instead. Empty lines and lines starting with the number sign # will be ignored.</td>
<td></td>
</tr>
<tr>
<td><code>-d</code></td>
<td>Toggles the debugging mode. Debugging information is output to standard error.</td>
</tr>
<tr>
<td><code>-i &lt;input&gt;</code></td>
<td>Opens an XML file or database specified by the argument. The opened input can then be processed by a command or XQuery expression.</td>
</tr>
<tr>
<td><code>-I &lt;input&gt;</code></td>
<td>Assigns an input string as item of type xs:untypedAtomic to the query context.</td>
</tr>
<tr>
<td><code>-n &lt;name&gt;</code></td>
<td>Specifies the host name on which the server is running.</td>
</tr>
<tr>
<td><code>-o &lt;file&gt;</code></td>
<td>All command and query output is written to the specified file.</td>
</tr>
<tr>
<td><code>-p &lt;num&gt;</code></td>
<td>Specifies the port on which the server is running.</td>
</tr>
<tr>
<td><code>-P &lt;pass&gt;</code></td>
<td>Specifies the user password. If this flag is omitted, the password will be requested on command line. Warning: when the password is specified with this flag, it may get visible to others.</td>
</tr>
<tr>
<td><code>-q &lt;expr&gt;</code></td>
<td>Executes the specified string as XQuery expression.</td>
</tr>
<tr>
<td><code>-r &lt;num&gt;</code></td>
<td>Specifies how often a specified query will be evaluated.</td>
</tr>
<tr>
<td><code>-R</code></td>
<td>Specifies if a query will be executed or parsed only.</td>
</tr>
<tr>
<td><code>-s &lt;pars&gt;</code></td>
<td>Specifies parameters for serializing XQuery results; see Serialization for more details. This flag may be specified multiple times. Key and values are separated by the equality sign (=).</td>
</tr>
</tbody>
</table>
### Command-Line Options

<table>
<thead>
<tr>
<th>Flag</th>
<th>Description</th>
<th>Option</th>
<th>Default</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>-U&lt;name&gt;</code></td>
<td>Specifies the user name. If this flag is omitted, the user name will be requested on command line.</td>
<td>USER</td>
<td>-Uadmin</td>
<td></td>
</tr>
<tr>
<td><code>-v</code></td>
<td>Prints process and timing information to the standard output.</td>
<td>false</td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>-V</code></td>
<td>Prints detailed query information to the standard output, including details on the compilation and profiling steps.</td>
<td>QUERYINFO</td>
<td>false</td>
<td></td>
</tr>
<tr>
<td><code>-w</code></td>
<td>Specifies if whitespaces in XML text nodes should be chopped (which is the default) or preserved.</td>
<td>CHOP</td>
<td>chop</td>
<td></td>
</tr>
<tr>
<td><code>-x</code></td>
<td>This flags turn on the output of the query execution plan, formatted in XML.</td>
<td>XMLPLAN</td>
<td>false</td>
<td></td>
</tr>
<tr>
<td><code>-X</code></td>
<td>Generates the query plan before or after query compilation. <code>-x</code> needs to be activated to make the plan visible.</td>
<td>COMPPLAN</td>
<td>after</td>
<td></td>
</tr>
<tr>
<td><code>-z</code></td>
<td>Turns the serialization of XQuery results on/off. This flag is useful if the query is profiled or analyzed.</td>
<td>SERIALIZE</td>
<td>true</td>
<td></td>
</tr>
</tbody>
</table>

### HTTP Server

Launch the HTTP server

```
$ basexhttp
BaseX [Server]
Server was started (port: 1984)
HTTP Server was started (port: 8984)
```

Available command-line flags can be listed with `-h`:

```
$ basexhttp -h
BaseX [HTTP]
Usage: basexhttp [-dhlnpPRUWz] [stop]
    stop      Stop running server
    -d        Activate debugging mode
    -h<port>  Set port of HTTP server
    -l        Start in local mode
    -n<name>  Set host name of database server
    -p<port>  Set port of database server
    -P<pass>  Specify user password
    -s<port>  Specify port to stop HTTP server
    -S        Start as service
    -U<name>  Specify user name
    -z        Suppress logging
```

The flags have the following meaning (equivalent database options are shown in the table as well). For the examples to work escaping some characters might be necessary, depending on your Operating System.

<table>
<thead>
<tr>
<th>Flag</th>
<th>Description</th>
<th>Option</th>
<th>Default</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>stop</td>
<td>Stops a local HTTP server and quits. The database server will be stopped as well, unless <code>-l</code> is specified.</td>
<td>pom.xml</td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>-d</code></td>
<td>Turns on the debugging mode. Debugging information is output to standard error.</td>
<td>DEBUG</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Command-Line Options

- **-h<num>** Specifies the port on which the HTTP server will be addressable.
  - jetty.xml 8984 -h9999
- **-l** Starts the server in *local mode*, and executes all commands in the embedded database context.
- **-n[name]** Specifies the host name on which the server is running.
  - LOCAL http://nservers.basex.org
- **-p<num>** Specifies the port on which the database server will be addressable.
  - SERVERPORT 1984 -p9998
- **-P<pass>** Specifies a user password, which will be used by the HTTP services to open a new session. If this flag is omitted, and if `-U` was specified, the password will be requested on command line. *Warning*: when the password is specified with this flag, it may get visible to others.
  - PASSWORD -Uadmin -Padmin
- **-s<num>** Specifies the port that will be used to stop the HTTP server.
  - STOPPORT or pom.xml 8985
- **-S** Starts the server as service (i.e., in the background).
- **-U<name>** Specifies a user name, which will be used by the HTTP services for opening a new session.
  - USER -Uadmin
- **-z** Does not generate any log files.
  - LOG

### GUI

Launch the GUI

```
$ basexgui [file]
```

One or more XML and XQuery files can be passed on as parameters. If an XML file is specified, a database instance is created from this file, or an existing database is opened. XQuery files are opened in the XQuery editor.

### Changelog

Version 8.2
- Removed: Event ports, `-e`.

Version 8.1
- Added: Bind input strings to the query context with `-I`.

Version 8.0
- Removed: Command-line option `-L` (results will now be automatically separated by newlines).

Version 7.9
- Added: Runs tests in file or directory with `-t`.
  - Removed: interactive server mode.
• Added: Specify if a query will be executed or parsed only with \(-R\).

Version 7.7
• Added: Bind host to the BaseX Server with \(-n\).

Version 7.5
• Added: detection of Command Scripts.
• Removed: HTTP server flags \(-R\), \(-W\), and \(-X\).

Version 7.3
• Updated: all options are now evaluated in the given order.
• Updated: Create main-memory representations for specified sources with \(-i\).
• Updated: Options \(-C/-c\) and \(-q/[input]\) merged.
• Updated: Option \(-L\) also separates serialized items with newlines (instead of spaces).

Version 7.2
• Added: RESTXQ Service

Version 7.1.1
• Added: Options \(-C\) and \(-L\) in standalone and client mode.

Version 7.1
• Updated: Multiple query files and \(-c/-i/-q\) flags can be specified.
Chapter 5. Start Scripts

Read this entry online in the BaseX Wiki.

The following scripts, which are referenced in the Startup and Command-Line Options articles, are also included in the Windows and ZIP distributions.

• We recommend you to manually add the bin directory of your BaseX directory to the PATH variable of your environment.

• The Windows installer automatically adds the project’s bin directory to your path environment.

• If you work with Maven, you can directly run the scripts in the basex-core/etc and basex-api/etc sub-directories of the project.

If BaseX terminates with an Out of Memory error, you can assign more RAM via the -Xmx flag (see below).

Main Package

The following scripts can be used to launch the standalone version of BaseX. Please replace the class name in org.basex.BaseX with either BaseXClient, BaseXServer, or BaseXGUI to run the client, server or GUI version.

Windows: basex.bat

@echo off
setLocal EnableDelayedExpansion

REM Path to this script
set PWD=%~dp0

REM Core and library classes
set CP=%PWD%/../BaseX.jar
set LIB=%PWD%/../lib
for /R "%LIB%" %%a in (*.jar) do set CP=!CP!;%%a

REM Options for virtual machine
set VM=-Xmx512m

REM Run code
java -cp "$CP%" $VM% org.basex.BaseX %*

Linux/Mac: basex

#!/bin/bash

# Path to this script
FILE="$(basename "$BASH_SOURCE")"
while [ -h "$FILE" ] ; do
  SRC="$(readlink "$FILE")"
  FILE="$( cd -P "$SRC" )"
  SRC="$( readlink "$FILE" )"
  FILE="" # Repeat until dereference.
done
BX="$( cd -P "$BX" )"

# Core and library classes
CP="$BX/BaseX.jar"
CP="$CP$(for JAR in "$BX"/lib/*.jar; do echo -n ":$JAR" done)"

# Options for virtual machine
VM=-Xmx512m

# Run code
java -cp "$CP" $VM org.basex.BaseX "$@

HTTP Server

The scripts for starting the HTTP server, which gives access to the REST, RESTXQ and WebDAV services, can be found below.

Windows: basexhttp.bat

@echo off
setLocal EnableDelayedExpansion

REM Path to this script
set PWD=%~dp0

REM Core and library classes
set CP=%PWD%/../BaseX.jar
set LIB=%PWD%/../lib
for /R "%LIB%" %%a in (*.jar) do set CP=!CP!;%%a
for /R "%LIB%" %%a in (*.jar) do set CP=!CP!;%%a

REM Options for virtual machine
set VM=-Xmx512m

REM Run code
java -cp "%CP%;." %VM% org.basex.BaseXHTTP %*

Linux/Mac: basexhttp

#!/bin/bash
# Path to this script
FILE="$(readlink $(basename "$BASH_SOURCE[0]"))"
while [ -h "$FILE" ] ; do
    SRC="$(readlink "$FILE")"
    FILE="$( cd -P "$SRC" && pwd )/$(basename "$SRC")"
    done
BX="$FILE"
BXCORE="$BX/..

# API, core, and library classes
CP="$BX/BaseX.jar$(printf ":s""$BX/BaseX.jar""$BX/lib/"*.jar "$BXCORE/target/classes""$BXCORE/lib/"*.jar")"

# Options for virtual machine
VM=-Xmx512m

# Run code
java -cp "$CP" $VM org.basex.BaseXHTTP "$@

Changelog

Version 7.5

• Updated: Static dependencies removed from Windows batch scripts.

Version 7.2
• Updated: The BaseXHTTP start class moved from org.basex.api to org.basex.

Version 7.0

• Updated: The basexjaxrx scripts have been replaced with the basexhttp scripts.
Part II. User Interfaces
Chapter 6. Graphical User Interface

Read this entry online in the BaseX Wiki.

This page is part of the Getting Started Section. The BaseX homepage gives you a visual impression of the graphical user interface (GUI) of BaseX, and an introductory video is available, which presents some of its interactive features.

Startup

First of all, launch a GUI instance of BaseX. Depending on your operating system, double click on the BaseX GUI start icon or run the basexgui script. Beside that, some more startup options are available.

Create Database

Select Database → New and browse to an XML document of your choice. As an example, you can start with the factbook.xml document, which contains statistical information on the world's countries. It is included in our official releases and can also be downloaded (1.3 MB). If you type nothing in the input field, an empty database will be created. Next, choose the OK button, and BaseX will create a database that you can visually explore and query.

If no XML document is available, the Text Editor can also be used to create an initial XML document. After saving the entered XML document to harddisk, it can be supplied as database input in the above dialog.

Input Bar

On top of the main window, you can find an input bar. The attached dropdown menu provides three modes: Find, XQuery and Command:

The upcoming example queries can all be used with an instance of the factbook database:

Find

In the Find mode, the input bar can be used to find single elements and texts in the currently opened database. The following syntax is supported:

<table>
<thead>
<tr>
<th>Query</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>city</td>
<td>Find elements named city, and texts containing this token.</td>
</tr>
<tr>
<td>=India</td>
<td>Find texts matching the exact string India.</td>
</tr>
<tr>
<td>~Cing</td>
<td>Find texts equal or similar to the token Cingdom.</td>
</tr>
<tr>
<td>id</td>
<td>Find attributes named id and attribute values containing this token.</td>
</tr>
<tr>
<td>@=f0_119</td>
<td>Find attribute values matching the exact string f0_119.</td>
</tr>
<tr>
<td>&quot;European Chinese&quot;</td>
<td>Find texts containing the phrase &quot;European Chinese&quot;.</td>
</tr>
<tr>
<td>//city</td>
<td>Leading slash: Interpret the input as XPath expression (see below).</td>
</tr>
</tbody>
</table>

XQuery

In the XQuery mode, XPath and XQuery expressions can be entered in the input bar.
To evaluate the following example queries: enter them in the input bar, and press ENTER or click on the START button:

<table>
<thead>
<tr>
<th>Query</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>//country</td>
<td>Return all country elements.</td>
</tr>
<tr>
<td>//country[name = &quot;Switzerland&quot;]</td>
<td>Return the country element of &quot;Switzerland&quot;.</td>
</tr>
<tr>
<td>for $city in //city where $city/po plation &gt; 1000000 order by $city</td>
<td>Return the names of all cities with a population larger than one million and order the results by the name of the city.</td>
</tr>
</tbody>
</table>

**Command**

In the **Command** mode, **BaseX Commands** can be entered and executed. Just try the following examples:

- **INFO**: Returns system information.
- **CREATE DB TEST**: Creates an empty database named "TEST".
- **LIST**: Lists all databases.

**Text Editor**

The powerful text editor can be used to write and edit **XQuery** code (simple queries, modules), **XML documents**, **Command Scripts**, and any other textual files:

- XQuery and command scripts can be **executed** by clicking on the green triangle.
- XQuery, XML and JavaScript will be **syntax highlighted**.
- XQuery and XML will automatically **parsed** with each key click, and **errors** will be highlighted.

Numerous **keyboard shortcuts** are available to speed up editing and debugging.

**Project View**

The Project View is attached to the editor panel. It displays all files of the current project directory in a tree structure. Files can be renamed and deleted by right-clicks on the files. The project directory can be changed by clicking on the button with the three dots... in the upper left corner of the project view panel.

Erroneous XQuery modules, and all queries importing these modules, will be highlighted. Highlighting can be disabled in the GUI preferences dialog.

**Realtime Options**

Via the **Options** menu, you can change how queries are executed and visualized:

- **Realtime Execution**: If realtime execution is enabled, your searches and queries will be executed with each key click and the results will be instantly shown.

- **Realtime Filtering**: If enabled, all visualizations will be limited to the actual results in realtime. If this feature is disabled, the query results are highlighted in the visualizations and can be explicitly filtered with the 'Filter' button.

**Visualizations**

The BaseX GUI offers various visualizations, which help you to explore your XML data instances from different perspectives:
Text View **Result**

Displays query results and other textual output. Query results can be saved in a file.

Tree View **Tree**

Displays all XML nodes in a top down tree with edges and nodes. You can change some settings of the Tree in the Menu Options → Tree Options.

Map View **Map**

Displays all data in a TreeMap. All nodes of the XML document are represented as rectangles, filling the complete area. You can choose different layout algorithms in the Menu Options → Map Layout.

Folder View **Folder**

Displays all nodes in an Explorer-like folder view. Nodes can be expanded or closed by clicking on the arrows.

Scatterplot View **Plot**

Displays all nodes in a scatterplot, which is particularly helpful if you want to explore analyze your data. Three drop down menus allow custom axis assignments.

The Table View **Table**

Comes in handy if your data is highly regular. It displays all nodes in a table with rows and columns. Different assignments can be chosen by clicking on the arrow in the right upper corner.
Graphical User Interface

Explorer View **Explorer**

Can be used to explore the contents of your database via drop-down menus, search fields and double sliders.

Info View **Query Info**

Helpful for analyzing the query plans of your XQuery expressions. It also displays information on the compilation and evaluation of queries.

**Look and Feels**

By default, the Look and Feel of your operating system will be used in the GUI. In the Preferences dialog, you can choose among some more window themes.

The JTattoo library offers some more look and feels. If you download and copy the JTattoo jar file into the lib directory provided by the ZIP and EXE distribution of BaseX, 13 additional looks and feels will get available.

**Changelog**

Version 8.4

- Added: highlighting of erroneous XQuery modules in the project view.

Version 8.0

- Updated: support for dark look and feels; support for JTattoo library
Chapter 7. Shortcuts

Read this entry online in the BaseX Wiki.

This article is about the GUI of BaseX. It gives you an overview of the most important hotkeys available in the visual frontend.

Editor

Code Completions

The GUI editor provides various code completions, which simplify the authoring of complex XQuery applications. Opening elements, comments, quotes or brackets will automatically be closed, and new lines will automatically be indented.

If some characters have been entered, and if the shortcut for code completions is pressed (Ctrl Space), a popup menu will appear and provides some code templates. If only one completion is possible, it will automatically be inserted.

Editor Shortcuts

The text editor can be used to create, edit, save and execute XQuery expressions, XML documents and any other textual files.

Custom Editing

<table>
<thead>
<tr>
<th>Description</th>
<th>Win/Linux</th>
<th>Mac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performs Code Completions</td>
<td>Ctrl Space</td>
<td>Ctrl Space</td>
</tr>
<tr>
<td>Sort lines</td>
<td>Ctrl U</td>
<td># U</td>
</tr>
<tr>
<td>Format code (experimental)</td>
<td>Ctrl Shift F</td>
<td># Shift F</td>
</tr>
<tr>
<td>(Un)comment selection/line</td>
<td>Ctrl K</td>
<td># K</td>
</tr>
<tr>
<td>Delete complete line</td>
<td>Ctrl Shift D</td>
<td># Shift D</td>
</tr>
<tr>
<td>Lower case</td>
<td>Ctrl Shift L</td>
<td># Shift L</td>
</tr>
<tr>
<td>Upper case</td>
<td>Ctrl Shift U</td>
<td># Shift U</td>
</tr>
<tr>
<td>Title case</td>
<td>Ctrl Shift T</td>
<td># Shift T</td>
</tr>
</tbody>
</table>

Finding

<table>
<thead>
<tr>
<th>Description</th>
<th>Win/Linux</th>
<th>Mac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Find files (or currently highlighted text)</td>
<td>Ctrl H</td>
<td># Shift H</td>
</tr>
<tr>
<td>Jump to next error in project</td>
<td>Ctrl . (period)</td>
<td># . (period)</td>
</tr>
<tr>
<td>Jump to currently edited file</td>
<td>Ctrl J</td>
<td># J</td>
</tr>
<tr>
<td>Go to line</td>
<td>Ctrl L</td>
<td># L</td>
</tr>
<tr>
<td>Find and replace text</td>
<td>Ctrl F</td>
<td># F</td>
</tr>
<tr>
<td>Find next instance of text</td>
<td>F3Ctrl G</td>
<td># F3# G</td>
</tr>
<tr>
<td>Find previous instance of text</td>
<td>Shift F3Ctrl Shift G</td>
<td># Shift F3# Shift G</td>
</tr>
</tbody>
</table>

25
# Shortcuts

## Standard Editing

<table>
<thead>
<tr>
<th>Description</th>
<th>Win/Linux</th>
<th>Mac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undo recent changes</td>
<td>Ctrl Z</td>
<td># Z</td>
</tr>
<tr>
<td>Redo recent changes</td>
<td>Ctrl Y</td>
<td># Shift Z</td>
</tr>
<tr>
<td>Cut selection</td>
<td>Ctrl X Ctrl Delete</td>
<td># X</td>
</tr>
<tr>
<td>Copy selection to clipboard</td>
<td>Ctrl C Ctrl Insert</td>
<td># C</td>
</tr>
<tr>
<td>Paste from clipboard</td>
<td>Ctrl V Shift Insert</td>
<td># V</td>
</tr>
<tr>
<td>Select All</td>
<td>Ctrl A</td>
<td># A</td>
</tr>
<tr>
<td>Delete character left of cursor</td>
<td>Backspace</td>
<td>Backspace</td>
</tr>
<tr>
<td>Delete character right of cursor</td>
<td>Delete</td>
<td>Delete (fn Backspace)</td>
</tr>
<tr>
<td>Delete word left of cursor</td>
<td>Ctrl Backspace</td>
<td>Alt Backspace</td>
</tr>
<tr>
<td>Delete word right of cursor</td>
<td>Ctrl Delete</td>
<td>Alt Delete</td>
</tr>
<tr>
<td>Delete text left of cursor</td>
<td>Ctrl Shift Backspace</td>
<td># Backspace</td>
</tr>
<tr>
<td>Delete text right of cursor</td>
<td>Ctrl Shift Delete</td>
<td># Delete</td>
</tr>
</tbody>
</table>

## Navigation

<table>
<thead>
<tr>
<th>Description</th>
<th>Win/Linux</th>
<th>Mac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Move one character to the left/right</td>
<td>←/-→</td>
<td>←/-→</td>
</tr>
<tr>
<td>Move one word to the left/right</td>
<td>Ctrl ←/-→</td>
<td>Alt ←/-→</td>
</tr>
<tr>
<td>Move to beginning/end of line</td>
<td>Home/End</td>
<td># ←/-→</td>
</tr>
<tr>
<td>Move one line up/down</td>
<td>↑/↓</td>
<td>↑/↓</td>
</tr>
<tr>
<td>Move one screen-full up/down</td>
<td>Page ↑/↓</td>
<td>Page ↑/↓ (fn ↑/↓)</td>
</tr>
<tr>
<td>Move to top/bottom</td>
<td>Ctrl Home/End</td>
<td>#/# (# ↑/↓)</td>
</tr>
<tr>
<td>Scroll one line up/down</td>
<td>Ctrl ↑/↓</td>
<td>Alt ↑/↓</td>
</tr>
</tbody>
</table>

## GUI

### Global Shortcuts

The following shortcuts are available from most GUI components:

<table>
<thead>
<tr>
<th>Description</th>
<th>Win/Linux</th>
<th>Mac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jump to input bar</td>
<td>F6</td>
<td># F6</td>
</tr>
<tr>
<td>Jump to next/previous panel</td>
<td>Ctrl (Shift) Tab</td>
<td>Ctrl (Shift) Tab</td>
</tr>
<tr>
<td>Increase/Decrease font size</td>
<td>Ctrl +/-</td>
<td># +/-</td>
</tr>
<tr>
<td>Reset font size</td>
<td>Ctrl 0</td>
<td># 0</td>
</tr>
</tbody>
</table>

### Menu Shortcuts

The following commands and options are also linked from the main menu:
## Shortcuts

### Database

<table>
<thead>
<tr>
<th>Description</th>
<th>Win/Linux</th>
<th>Mac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create new database</td>
<td>Ctrl N</td>
<td># N</td>
</tr>
<tr>
<td>Open/manage existing databases</td>
<td>Ctrl M</td>
<td># M</td>
</tr>
<tr>
<td>View/edit database properties</td>
<td>Ctrl D</td>
<td># D</td>
</tr>
<tr>
<td>Close opened database</td>
<td>Ctrl Shift W</td>
<td># Shift W</td>
</tr>
<tr>
<td>Exit application</td>
<td>Ctrl Q</td>
<td># Q</td>
</tr>
</tbody>
</table>

### Editor

<table>
<thead>
<tr>
<th>Description</th>
<th>Win/Linux</th>
<th>Mac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create new tab</td>
<td>Ctrl T</td>
<td># T</td>
</tr>
<tr>
<td>Open existing file</td>
<td>Ctrl O</td>
<td># O</td>
</tr>
<tr>
<td>Save file</td>
<td>Ctrl S</td>
<td># S</td>
</tr>
<tr>
<td>Save copy of file</td>
<td>Ctrl Shift S</td>
<td># Shift S</td>
</tr>
<tr>
<td>Close tab</td>
<td>Ctrl W, Ctrl F4</td>
<td># W, # F4</td>
</tr>
</tbody>
</table>

### View

<table>
<thead>
<tr>
<th>Description</th>
<th>Win/Linux</th>
<th>Mac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toggle query/text editor</td>
<td>Ctrl E</td>
<td># E</td>
</tr>
<tr>
<td>Toggle project structure</td>
<td>Ctrl P</td>
<td># P</td>
</tr>
<tr>
<td>Toggle result view</td>
<td>Ctrl R</td>
<td># R</td>
</tr>
<tr>
<td>Toggle query info view</td>
<td>Ctrl I</td>
<td># I</td>
</tr>
</tbody>
</table>

### Options

<table>
<thead>
<tr>
<th>Description</th>
<th>Win/Linux</th>
<th>Mac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open preference dialog</td>
<td>Ctrl Shift P</td>
<td># , (comma)</td>
</tr>
</tbody>
</table>

### Visualization

<table>
<thead>
<tr>
<th>Description</th>
<th>Win/Linux</th>
<th>Mac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toggle map view</td>
<td>Ctrl 1</td>
<td># 1</td>
</tr>
<tr>
<td>Toggle tree view</td>
<td>Ctrl 2</td>
<td># 2</td>
</tr>
<tr>
<td>Toggle folder view</td>
<td>Ctrl 3</td>
<td># 3</td>
</tr>
<tr>
<td>Toggle plot view</td>
<td>Ctrl 4</td>
<td># 4</td>
</tr>
<tr>
<td>Toggle table view</td>
<td>Ctrl 5</td>
<td># 5</td>
</tr>
<tr>
<td>Toggle explorer view</td>
<td>Ctrl 6</td>
<td># 6</td>
</tr>
</tbody>
</table>

### Help

<table>
<thead>
<tr>
<th>Description</th>
<th>Win/Linux</th>
<th>Mac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Show Help</td>
<td>F1</td>
<td>F1</td>
</tr>
</tbody>
</table>

Additionally, the names of HTML entities will be converted to their Unicode representation (as an example, Åuml; will be translated to ä).
Changelog

Version 8.4

• Added: Lower case (Ctrl Shift L), Upper case (Ctrl Shift U), Title case (Ctrl Shift T)

Version 8.0

• Added: New code completions, popup menu

Version 7.8.2

• Added: Sort lines (Ctrl U)

Version 7.8

• Added: Code Completions, Project (Ctrl P), Find Files (Ctrl Shift F)

Version 7.5

• Added: go to line (Ctrl F)

Version 7.3

• Added: delete complete line (Ctrl Shift D), jump to highlighted error (Ctrl .)
Chapter 8. Database Server

Read this entry online in the BaseX Wiki.

This step by step tutorial is part of the Getting Started Guide. It shows you how to run BaseX in client-server mode from a terminal. You can copy and paste all commands to get them running on your machine. After you finished this tutorial, you will be familiar with the basic administration of BaseX. Visit the commands section for a complete list of database commands.

Startup

First, launch a Server and Client instance of BaseX: double click on the BaseX Server/Client icons, or run the basexserver and basexclient scripts. Follow this link for more information (or check out the additional command-line options).

Working on Command-Line

The BaseX command-line client provides similar features to the Standalone Mode. The major difference is that all commands will be executed by the BaseX server instance. As a consequence, paths/URIs to resources need to be resolvable by the server (file contents will not be transfered to the server).

Username and password can also be specified as command-line option. To evaluate commands without entering the console mode, you can use the -c option on the command line:

```
basexclient -V -U admin -P admin -c "CREATE DB input <example/>; XQUERY /"
```

Database 'input' created in 13.85 ms.
<example/>
Query:
/

Parsing: 0.18 ms
Compiling: 0.04 ms
Evaluating: 0.12 ms
Printing: 0.07 ms
Total Time: 0.41 ms

Hit(s): 1 Item
Updated: 0 Items
Printed: 10 Bytes
Read Locking: local [input]
Write Locking: none

Query "user" executed in 0.41 ms.
Chapter 9. Standalone Mode

Read this entry online in the BaseX Wiki.

This page is part of the Getting Started Section. BaseX offers a standalone console mode from which all database commands can be executed. The article on the Database Server provides numerous examples for running commands in the console mode (note that the GUI does not interact with the client/server architecture).

Startup

First of all, please launch a standalone version of BaseX: double click on the BaseX icon, or run the basex script. Follow this link for more information (or check out the additional command-line options).

Create a Database

• To create a database you need an XML document, e.g. factbook.xml.
• Save this document to your working directory.
• Type in the following command to create and open the database:

  > CREATE DB factbook factbook.xml

  factbook is the name of the database
  factbook.xml is the initial input of the database

Where is the database stored?

By default, databases are stored in the BaseXData directory, which is located in your home folder. Depending on your Configuration, the location of your home folder varies. For example, on a Mac it's /Users/John, if your name is John. If you are working with the ZIP or EXE distribution, the directory will be named data and reside in the application directory.

Execute a Query

The XQUERY command lets you run a query.

• For example, this query returns all country nodes in the currently opened database.

  > XQUERY //country

• You can also run queries in files:

  > RUN /Users/John/query.xq

Database Commands

• The following command lists all databases than can be opened by the currently logged in user:

  > LIST

• To open an existing database, execute the following:

  > OPEN factbook

• To get information on the currently opened database, type:

  > INFO
• You can also address a database within your query with the `db:open` function:

```
> XQUERY db:open("factbook")//country
```

• To close the current database, please type:

```
> CLOSE
```

• Use the `DROP DB` command to delete a database:

```
> DROP DB factbook
```

**Multiple Resources**

One database can contain not only a single, but millions of documents. All documents can have a different structure.

With the following commands, you can create an empty database and add two documents. It is also possible to address resources via URLs:

```
> CREATE DB store > ADD factbook.xml > ADD http://files.basex.org/xml/xmark.xml
```

• Deleting a document from a database is easy:

```
> DELETE factbook.xml
```

Make sure that the database, which contains the addressed document, is currently opened.

**Backup and Restore**

• To backup your database, type:

```
> CREATE BACKUP factbook
```

• To restore your database, type:

```
> RESTORE factbook
```

The backup file is stored in the database directory. The contains the name of the database and a timestamp: `[db-name]-[timestamp].zip`. If a database is restored, and if several backups exist, the backup with the newest timestamp is taken.
Chapter 10. Web Application

Read this entry online in the BaseX Wiki.

BaseX can be used as Web Application. The following three HTTP services are available:

- **RESTXQ** allows you to write web applications with XQuery.
- **REST** offers a RESTful API for accessing database resources via URLs, and
- **WebDAV** provides access to databases via the file system.

This article describes different ways of deploying and configuring these services. The services can be deployed as follows:

- as standalone application by running the BaseX HTTP Server,
- as web servlet in a Servlet Container, and
- as web servlet, using Maven.

**Servlet Container**

In order to deploy BaseX HTTP Services in a servlet container, you may download the WAR distribution of BaseX from the download site or compile it via `mvn compile war:war` in the basex-api package. The WAR file can then be deployed following the instructions of the corresponding servlet container (jetty, tomcat).

Configuring port, context path, etc. can be done by following the corresponding instructions of the used servlet container. This is needed if you want to replace the default URL path (e.g. `http://localhost:8080/rest`) with a custom one (e.g. `http://localhost:8080/BaseX711/rest`).

If run on a Jetty server you may use a `jetty.xml` file for detailed server configuration. You can e.g. enable SSL connections or Jetty logging. Place the `jetty.xml` right next to the `web.xml`. For detailed configuration refer to the [Jetty Documentation](http://www.eclipse.org/jetty/documentation). A sample `jetty.xml` is placed in the basex-api package.

To run on Apache Tomcat, start the tomcat server and add any *.war* distribution to deploy using the Tomcat web interface. By default, the interface is accessible via `http://localhost:8080/manager/html/`.

**Configuration**

All database options can be specified in the `web.xml` file. They need to be represented as context parameters and prefixed with `org.basex`. The most important options for the web application context are as follows:

<table>
<thead>
<tr>
<th>Option</th>
<th>Default</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>USER</td>
<td>admin</td>
<td>Applies to REST and WebDAV service: If no user is specified, the credentials must be passed on by the client. Please check by yourself if it is safe to store your credentials in plain text.</td>
</tr>
<tr>
<td>PASSWORD</td>
<td>admin</td>
<td>Applies to REST and WebDAV service: If no password is specified, it must be passed on by the client. Please check by yourself if it is safe to store your credentials in plain text.</td>
</tr>
<tr>
<td>HTTPLOCAL</td>
<td>false</td>
<td>Operation mode. By default, a database server instance will be started, as soon as the first HTTP service is called. The database server can be disabled by setting this flag to <code>true</code>.</td>
</tr>
<tr>
<td>RESTXQPATH</td>
<td>.</td>
<td>Relative or absolute directory referencing the RESTXQ modules. By default, the option points to the standard web application directory.</td>
</tr>
<tr>
<td>RESTPATH</td>
<td>.</td>
<td>Relative or absolute directory referencing queries and command-scripts that can be invoked via the run operation of REST. By default, the option points to the standard web application directory.</td>
</tr>
</tbody>
</table>
AUTHMETHOD | Basic |
The default authentication method proposed by the server. The available methods are Basic and Digest.

Path options may contain an absolute or relative path. If a relative path is specified, its root will be the servlet (webapp) path:

```xml
<context-param>
  <param-name>org.basex.dbpath</param-name>
  <!-- will be rewritten to ..../webapp/WEB-INF/data -->
  <param-value>WEB-INF/data</param-value>
</context-param>
<context-param>
  <param-name>org.basex.repopath</param-name>
  <!-- will be kept as is -->
  <param-value>f:/basex/repository</param-value>
</context-param>
```

Context parameters can be requested from XQuery via `proc:property-names` and `proc:property`. How to set these options is specific to the servlet container. For example, in Jetty it can be done by overriding the `web.xml` file. Another option is to directly edit the `WEB-INF/web.xml` file in the WAR archive (WAR files are simple ZIP files). Refer to the sample `web.xml` of the basex-api package.

Different credentials can be assigned to the REST and WebDAV service by specifying local init parameters. In the following example, specific credentials are set for the REST service:

```xml
<servlet>
  <servlet-name>REST</servlet-name>
  <servlet-class>org.basex.http.rest.RESTServlet</servlet-class>
  <init-param>
    <param-name>org.basex.user</param-name>
    <param-value>rest-user</param-value>
  </init-param>
  <init-param>
    <param-name>org.basex.password</param-name>
    <param-value>(:87!7X3$o3p</param-value>
  </init-param>
</servlet>
```

**Available Services**

To enable or disable one of the provided services, the corresponding servlet entry in the `web.xml` file needs to be removed/commented. The default URL paths are listed in the following table:

<table>
<thead>
<tr>
<th>Service</th>
<th>URL</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default web server</td>
<td>http://[host]:[port]/ [servlet_context_path]/ staticBefore: http://[host]:[port]/ [servlet_context_path]</td>
<td>Access your standard web files (e.g. HTML, JavaScript or CSS).</td>
</tr>
<tr>
<td>RESTXQ</td>
<td>http://[host]:[port]/ [servlet_context_path]</td>
<td>Create XQuery web services and applications.</td>
</tr>
<tr>
<td></td>
<td>http://[host]:[port]/ [servlet_context_path]/ restxq</td>
<td></td>
</tr>
<tr>
<td>REST</td>
<td>http://[host]:[port]/ [servlet_context_path]/ rest</td>
<td>Access XML database and its resources.</td>
</tr>
<tr>
<td>WebDAV</td>
<td>http://[host]:[port]/[servlet_context_path]/webdav orwebdav://[host]:[port]/[servlet_context_path]/webdav (depending on client)</td>
<td>Access databases via the filesystem.</td>
</tr>
</tbody>
</table>

**Maven**

Checkout the BaseX sources via Eclipse or Git. Execute `mvn install` in the `baseX-core` project folder and then `mvn install jetty:run` in the `baseX-api` project folder. This will start a Jetty instance in which the servlets will be deployed.

**Configuration**

The same options as in the case of deployment in a servlet container apply. In this case, however, there is no WAR archive. Instead, Jetty looks up all files in the directory `baseX-api/src/main/webapp`. Jetty and servlet options can be configured in the `jetty.xml` and `web.xml` files as described above in the Servlet Container Configuration. The Jetty stop port can be changed in the Maven Jetty Plugin session in the `pom.xml` file.

**User Management**

By default, the REST and WebDAV services require client-side authentication. Default credentials can be stored server-side in the `web.xml` file or specified via command-line arguments. If the HTTP server is started with no pre-defined credentials, users and passwords can be sent via HTTP Basic Authentication or Digest Authentication.

Users are specified in a `users.xml` file, which is stored in the database directory (see User Management for more information).

With cURL, and most browsers, you can specify the user name and password with each HTTP request within the request string as plain text, using the format `USER:PASSWORD@URL`. An example:

```
http://admin:admin@localhost:8984/
```

**Changelog**

**Version 8.0**

- Added: digest authentication
- Updated: user management
- Updated: default user/password disabled in web.xml

**Version 7.7**

- Added: service-specific permissions

**Version 7.5**

- Added: `jetty.xml`: configuration for Jetty Server
- Updated: server replaced with `httplocal` mode

**Version 7.3**

- Updated: `client` mode replaced with `server` mode

**Version 7.2**
• Web Application concept revised
Chapter 11. DBA

Read this entry online in the BaseX Wiki.

This page is part of the Getting Started Section.

The full distributions of BaseX are equipped with a simple browser-based database administration interface, the DBA. It allows you to create and administrate databases, evaluate queries in realtime, view log files and manage users. The server-side code is completely written in XQuery and RESTXQ.

These were our design goals:

• The code base is supposed to help and motivate you developing your own RESTXQ web applications.

• The whole DBA code consumes only 100 KB. It uses very simple Javascript code that should run with nearly every browser. The interface is functional, but limited in terms of flashiness and interactivity.

• We tried to make the DBA features as self-explanatory as possible. All functionalities are also available via Commands, XQuery Modules or the Java GUI.

• The dba sub-directory can simply be copied and moved to any other place. All URL paths point to the same directory; it should be straightforward to adjust the RESTXQ path.

Please be aware that it is an obvious security risk if you simply put DBA online along with your web page. At the very least, you should change the DBA path and the password of the admin user!

The DBA has just been released, and it is in beta stage. We are looking forward to your reports and feature requests! Patches and extensions are even more welcome.

Startup

• Download the ZIP Archive or the Windows Installer from the download page

• Start the BaseX HTTP Server

• Open a browser and visit the URL http://localhost:8984/dba

First Steps

On the welcome page, you will need to authenticate yourself by entering a username and password. The default user is admin/admin; after the first login, it should be changed in the Users panel. By specifying a host and port, you can communicate with a remote BaseX server instance. If you keep the field empty, you will be connected to the local BaseX instance.

The main page of the DBA interface contains a list of all databases on the left. On the right, the global and local options are listed, along with some system information.

With the "Create..." button, a new database can be created. Existing database can be viewed, optimized, and dropped.

Editor

In the editor panel, you can execute XQuery expressions. If evaluation takes too long, or if it consumes too much memory, it will be interrupted. You need to choose if your query is updating. Inside the editor area, you can press Ctrl-Enter to execute the query.
You can press Shift-Ctrl-Enter to run your XQuery expression as updating query (or non-updating, if "Updating" is chosen in the dropdown menu). The realtime mode was removed.

**Changelog**

Version 8.4

- Added: Editor: Key combination 'Shift-Ctrl-Enter', realtime mode removed.

Introduced with Version 8.0.
Part III. General Info
Chapter 12. Databases

In BaseX, a database is a pretty light-weight concept and can be compared to a collection. It contains an arbitrary number of resources, addressed by their unique database path. Resources can either be XML documents or raw files (binaries). Some information on binary data can be found on an extra page.

Create Databases

New databases can be created via commands, in the GUI, or with any of our APIs. If some input is specified along with the create operation, it will be added to the database in a bulk operation:

- **Console**: 
  
  ```
  CREATE DB db /path/to/resources
  ```
  
  will add initial documents to a database

- **GUI**: Go to Database → New, press Browse to choose an initial file or directory, and press OK

Database must follow the valid names constraints. Various parsers can be chosen to influence the database creation, or to convert different formats to XML.

**Note**: A main-memory only database can be created using the the SET MAINMEM true command before calling CREATE DB (see below for more).

Access Resources

Stored resources and external documents can be accessed in different ways:

**XML Documents**

Various XQuery functions exist to access XML documents in databases:

<table>
<thead>
<tr>
<th>Function</th>
<th>Example</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>db:open</td>
<td><code>db:open(&quot;db&quot;, &quot;path/to/docs&quot;)</code></td>
<td>Returns all documents that are found in the database db at the (optional) path path/to/docs.</td>
</tr>
<tr>
<td>fn:collection</td>
<td><code>collection(&quot;db/path/to/docs&quot;)</code></td>
<td>Returns all documents at the location path/to/docs in the database db. If no path is specified after the database, all documents in the database will be returned. If no argument is specified, all documents of the database will be returned that has been opened in the global context.</td>
</tr>
<tr>
<td>fn:doc</td>
<td><code>doc(&quot;db/path/to/doc.xml&quot;)</code></td>
<td>Returns the document at the location path/to/docs in the database db. An error is raised if the specified yields zero or more than one document.</td>
</tr>
</tbody>
</table>

You can access multiple databases in a single query:

```
for $i in 1 to 100
return db:open('books' || $i)//book/title
```
If the DEFAULTDB option is turned on, the path argument of the fn:doc or fn:collection function will first be resolved against the globally opened database.

Two more functions are available for retrieving information on database nodes:

<table>
<thead>
<tr>
<th>Function</th>
<th>Example</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>db:name</td>
<td>db:name($node)</td>
<td>Returns the name of the database in which the specified $node is stored.</td>
</tr>
<tr>
<td>db:path</td>
<td>db:path($node)</td>
<td>Returns the path of the database document in which the specified $node is stored.</td>
</tr>
</tbody>
</table>

The fn:document-uri and fn:base-uri functions return URIs that can also be reused as arguments for the fn:doc and fn:collection functions. As a result, the following example query always returns true:

```
every $c in collection('anyDB')
satisfies doc-available(document-uri($c))
```

If the argument of fn:doc or fn:collection does not start with a valid database name, or if the addressed database does not exist, the string is interpreted as URI reference, and the documents found at this location will be returned. Examples:

- `doc("http://web.de")`: retrieves the addressed URI and returns it as a main-memory document node.
- `doc("myfile.xml")`: retrieves the given file from the file system and returns it as a main-memory document node. Note that updates to main-memory nodes are not automatically written back to disk unless the WRITEBACK option is set.
- `collection("/path/to/docs")`: returns a main-memory collection with all XML documents found at the addressed file path.

**Raw Files**

The RETRIEVE command and the db:retrieve function can be used to return files in their native byte representation.

If the API you use does not support binary output (this is e.g. the case for various Client language bindings), you need to convert your binary data to its string representation before returning it to the client:

```
string(db:retrieve('multimedia', 'sample.avi'))
```

**HTTP Services**

- With REST and WebDAV, all database resources can be requested in a uniform way, no matter if they are well-formed XML documents or binary files.

**Update Resources**

Once you have created a database, additional commands exist to modify its contents:

- XML documents can be added with the ADD command.
- Raw files are added with STORE.
- Existing resources can be replaced with the REPLACE command.
- Resources can be deleted via DELETE.
The **AUTOFLUSH** option can be turned off before *bulk operations* (i.e. before a large number of new resources is added to the database).

The **ADDCACHE** option will first cache the input before adding it to the database. This is helpful when the input documents to be added are expected to eat up too much main memory.

The following commands create an empty database, add two resources, explicitly flush data structures to disk, and finally delete all inserted data:

```sql
CREATE DB example
SET AUTOFLUSH false
ADD example.xml
SET ADDCACHE true
ADD /path/to/xml/documents
STORE TO images/ 123.jpg
FLUSH
DELETE /
```

You may as well use the BaseX-specific [XQuery Database Functions](https://www.basex.com/doc/ref-guide/xquery) to create, add, replace, and delete XML documents:

```sql
let $root := "/path/to/xml/documents/
for $file in file:list($root)
return db:add("database", $root || $file)
```

Last but not least, XML documents can also be added via the GUI and the **Database** menu.

**Export Data**

All resources stored in a database can be *exported*, i.e., written back to disk. This can be done in several ways:

- **Commands:** `EXPORT` writes all resources to the specified target directory
- **GUI:** Go to **Database** → **Export**, choose the target directory and press **OK**
- **WebDAV:** Locate the database directory (or a sub-directory of it) and copy all contents to another location

**In Memory Database**

- In the standalone context, a main-memory database can be created (using `CREATE DB`), which can then be accessed by subsequent commands.
- If a BaseX server instance is started, and if a database is created in its context (using `CREATE DB`), other BaseX client instances can access (and update) this database (using `OPEN`, `db:open`, etc.) as long as no other database is opened/created by the server.

**Note:** main-memory database instances are also created by the invocation of `doc(...)` or `collection(...)`, if the argument is not a database (no matter which value is set for `MAINMEM`). In other words: the same internal representation is used for main-memory databases and documents/collections generated via XQuery.

**Changelog**

Version 8.4

- Updated: **Raw Files**: Items of binary type can be output without specifying the obsolete `raw` serialization method.

Version 7.2.1
• Updated: \texttt{fn:document-uri} and \texttt{fn:base-uri} now return strings that can be reused with \texttt{fn:doc} or \texttt{fn:collection} to reopen the original document.
Chapter 13. Binary Data

Read this entry online in the BaseX Wiki.

This page is linked from the Database page.

The BaseX store also provides support for raw files (binary data). A database may contain both XML documents and raw files. XML and binary data is handled in a uniform way: a unique database path serves as key, and the contents can be retrieved via database commands, XQuery, or the various APIs.

Storage

XML documents are stored in a proprietary format to speed up XPath axis traversals and update operations, and raw data is stored in its original format in a dedicated sub-directory (called raw). Several reasons exist why we did not extend our existing storage to binary data:

- **Good Performance**: the file system generally performs very well when it comes to the retrieval and update of binary files.
- **Key/Value Stores**: we do not want to compete with existing key/value database solutions. Again, this is not what we are after.
- **Our Focus**: our main focus is the efficient storage of hierarchical data structures and file formats such as XML or (more and more) JSON. The efficient storage of arbitrary binary resources would introduce many new challenges that would distract us from more pressing tasks.

For some use cases, the chosen database design may bring along certain limitations:

- **Performance Limits**: most file system are not capable of handling thousands or millions of binary resources in a single directory in an efficient way. The same problem happens if you have a large number of XML documents that need to imported in or exported from a BaseX database. The general solution to avoid this bottleneck is to distribute the relevant binaries in additional sub-directories.
- **Keys**: if you want to use arbitrary keys for XML and binary resources, which are not supported by the underlying file system, you may either add an XML document in your database that contains all key/path mappings.

In the latter case, a key/value store might be the better option anyway.

Usage

More information on how to store, retrieve, update and export binary data is found in the general Database documentation.
Chapter 14. Parsers

This article is part of the Getting Started Section. It presents the available parsers that can be used to import various data sources in BaseX databases. Please visit the Serialization article if you want to know how to export data.

XML Parsers

BaseX provides two parsers to import XML data:

- By default, the internal, built-in XML parser is used, which is more fault-tolerant than Java’s XML parser. It supports standard HTML entities out-of-the-box, and is faster in most cases. In turn, it does not support all oddities specified by DTDs, and cannot resolve catalogs.

- Java’s SAXParser can also be selected for parsing XML documents. This parser is stricter than the built-in parser, but it refuses to process some large documents.

GUI

Go to Menu Database → New, then choose the Parsing tab and (de)activate Use internal XML parser. The parsing of DTDs can be turned on/off by selecting the checkbox below.

Command Line

To turn the internal XML parser and DTD parsing on/off, modify the INTPARSE and DTD options:

```
SET INTPARSE true
SET DTD true
```

XQuery

The `db:add` and `db:replace` functions can also be used to add new XML documents to the database. The following example query uses the internal XML parser and adds all files to the database DB that are found in the directory 2Bimported:

```
for $file in file:list("2Bimported")
return db:add('DB', $file, '', map { 'intparse': true() })
```

HTML Parser

With TagSoup, HTML can be imported in BaseX without any problems. TagSoup ensures that only well-formed HTML arrives at the XML parser (correct opening and closing tags, etc.). Hence, if TagSoup is not available on a system, there will be a lot of cases where importing HTML fails, no matter whether you use the GUI or the standalone mode.

Installation

Downloads

TagSoup is already included in the full BaseX distributions (`BaseX.zip`, `BaseX.exe`, etc.). It can also be manually downloaded and embedded on the appropriate platforms.

Maven

An easy way to add TagSoup to your own project is to follow these steps:

1. visit MVN TagSoup Repository
2. click on the version you want

3. you can see on the first tab called Maven a XML like this:

```xml
<dependency>
  <groupId>org.ccil.cowan.tagsoup</groupId>
  <artifactId>tagsoup</artifactId>
  <version>1.2.1</version>
</dependency>
```

4. copy that in your own maven project's pom.xml under the <dependencies> tag.

5. don't forget to run mvn jetty:run again

**Debian**

With Debian, TagSoup will be automatically detected and included after it has been installed via:

```
apt-get install libtagsoup-java
```

**TagSoup Options**

TagSoup offers a variety of options to customize the HTML conversion. For the complete list please visit the TagSoup website. BaseX supports most of these options with a few exceptions:

- **encoding**: BaseX tries to guess the input encoding but this can be overwritten by the user if necessary.
- **files**: not supported as input documents are piped directly to the XML parser.
- **method**: set to 'xml' as default. If this is set to 'html' ending tags may be missing for instance.
- **version**: dismissed, as TagSoup always falls back to 'version 1.0', no matter what the input is.
- **standalone**: deactivated.
- **pyx, pyxin**: not supported as the XML parser can't handle this kind of input.
- **output-encoding**: not supported, BaseX already takes care of that.
- **reuse, help**: not supported.

**GUI**

Go to Menu Database → New and select "HTML" in the input format combo box. There's an info in the "Parsing" tab about whether TagSoup is available or not. The same applies to the "Resources" tab in the "Database Properties" dialog.

These two dialogs come with an input field 'Parameters' where TagSoup options can be entered.

**Command Line**

Turn on the HTML Parser before parsing documents, and set a file filter:

```
SET PARSER html
SET HTMLPARSER method=xml,nons=true,nodatad=true,nodefaults=true,nobogons=true,nocolons=true,ignoreable=true
SET CREATEFILTER *.html
```

**XQuery**

The HTML Module provides a function for converting HTML to XML documents.

Documents can also be converted by specifying the parser and additional options as function arguments:
BaseX can also import JSON documents. The resulting format is described in the documentation for the XQuery JSON Module:

**GUI**

Go to Menu Database → New and select "JSON" in the input format combo box. You can set the following options for parsing JSON documents in the "Parsing" tab:

- **Encoding**: Choose the appropriate encoding of the JSON file.
- **JsonML**: Activate this option if the incoming file is a JsonML file.

**Command Line**

Turn on the JSON Parser before parsing documents, and set some optional, parser-specific options and a file filter:

```
SET PARSER json
SET JSONPARSER encoding=utf-8, jsonml=true
SET CREATEFILTER *.json
```

**XQuery**

The JSON Module provides functions for converting JSON objects to XML documents.

---

**CSV Parser**

BaseX can be used to import CSV documents. Different alternatives how to proceed are shown in the following:

**GUI**

Go to Menu Database → New and select "CSV" in the input format combo box. You can set the following options for parsing CSV documents in the "Parsing" tab:

- **Encoding**: Choose the appropriate encoding of the CSV file.
- **Separator**: Choose the column separator of the CSV file. Possible: comma, semicolon, tab or space or an arbitrary character.
- **Header**: Activate this option if the incoming CSV files have a header line.

**Command Line**

Turn on the CSV Parser before parsing documents, and set some optional, parser-specific options and a file filter. Unicode code points can be specified as separators; 32 is the code point for spaces:

```
SET PARSER csv
SET CSVPARSER encoding=utf-8, lines=true, header=false, separator=space
SET CREATEFILTER *.csv
```

**XQuery**

The CSV Module provides a function for converting CSV to XML documents.
Documents can also be converted by specifying the parser in an XQuery function. The following example query adds all CSV files that are located in the directory `2Bimported` to the database `DB` and interprets the first lines as column headers:

```xml
for $file in file:list("2Bimported", false(), ",*.csv")
return db:add("DB", $file, ",", map {
   'parser': 'csv',
   'csvparser': map { 'header': true() }
})
```

### Text Parser

Plain text can be imported as well:

#### GUI

Go to Menu `Database → New` and select "TEXT" in the input format combobox. You can set the following option for parsing text documents in the "Parsing" tab:

- **Encoding**: Choose the appropriate encoding of the text file.
- **Lines**: Activate this option to create a `<line>...</line>` element for each line of the input text file.

#### Command Line

Turn on the CSV Parser before parsing documents and set some optional, parser-specific options and a file filter:

```bash
SET PARSER text
SET TEXTPARSER lines=yes
SET CREATEFILTER *
```

### XQuery

Similar to the other formats, the text parser can also be specified via XQuery:

```xml
for $file in file:list("2Bimported", true(), ",*.txt")
return db:add("DB", $file, ",", map { 'parser': 'text' })
```

### Changelog

**Version 7.8**

- Updated: parser options

**Version 7.7.2**

- Removed: CSV option "format"

**Version 7.3**

- Updated: the CSV `SEPARATOR` option may now be assigned arbitrary single characters

**Version 7.2**

- Updated: Enhanced support for TagSoup options
Chapter 15. Commands

Read this entry online in the BaseX Wiki.

This article is part of the Getting Started Section. It lists all database commands supported by BaseX.

Commands can be executed from the Command Line, as part of Scripts, via the Clients, REST, the input field in the GUI, and other ways. If the GUI is used, all commands that are triggered by the GUI itself will show up in the Info View. The Permission fields indicate which rights are required by a user to perform a command in the client/server architecture.

Basics

Command Scripts

Multiple commands can be written down in a single line – separated by semicolons – or put into a command script: Database commands in both string and XML syntax can be placed in a text file and stored as file with the BaseX command script suffix .bxs. If the path to a script file is passed on to BaseX on command-line, or if it is opened in the GUI editor, it will be recognized and evaluated as such.

String Syntax

Lines starting with # are interpreted as comments and are skipped. With the following script, a database is created, two documents are added to it, and a query is performed:

```
CREATE DB test
ADD TO embedded.xml <root>embedded</root>
# run query
XQUERY <hits>{ count(//text()) } </hits>
CLOSE
```

XML Syntax

The string syntax is limited when XML snippets need to be embedded in a command, or when complex queries are to be specified.

The XML syntax provides more flexibility here. Multiple commands can be enclosed by a <commands/> root element. Some commands, such as ADD, allow you to directly embed XML documents. If you want to embed XML in XQuery expressions, entities should be encoded, or the CDATA syntax should be used:

```
<commands>
  <create-db name='test'/>
  <add path='embedded.xml'><root>embedded</root></add>
  <!-- run query -->
  <xquery><![CDATA[<hits>{ count(//text()) } ]]></xquery>
  <close/>
</commands>
```

Glob Syntax

Some commands support the glob syntax to address more than one database or user. Question marks and asterisks can be used to match one or more characters, and commas can be used to separate multiple patterns. Some examples:

- AB? addresses all names with the characters AB and one more character.
• *AB addresses all names ending with the characters AB.

• X*, Y*, Z* addresses all names starting with the characters X, Y, or Z.

Valid Names

Database and user names follow the same naming constraints: Names are restricted to ASCII characters. They must at least have one character, and they may contain letters, numbers and any of the special characters !#$ %&'()+-=@[^_`{}~. The following characters are reserved for other features:

• *,?*: glob syntax

• ;: Separator for multiple database commands on the command line

• \/: Directory path separators

• .: hidden folders (e.g. the .logs directory)

• :*?"<>|: invalid filename characters on Windows

Aliases

In all commands, the DB keyword can be replaced by DATABASE.

Database Operations

CREATE DB

Syntax CREATE DB [name] ([input])

XML Syntax <create-db name='...'>({input})</create-db>

Permission CREATE

Summary Creates a new database with the specified name and, optionally, an initial input, and opens it. An existing database will be overwritten. The input may either be a reference to a single XML document, a directory, a remote URL, or a string containing XML:

• name must be a valid database name

• database creation can be controlled by setting Create Options

Errors The command fails if a database with the specified name is currently used by another process, if one of the documents to be added is not well-formed or if it cannot be parsed for some other reason.

Examples

• CREATE DB input creates an empty database input.

• CREATE DB xmark http://files.basex.org/xml/xmark.xml creates the database xmark, containing a single initial document called xmark.xml.

• CREATE DATABASE coll /path/to/input creates the database coll with all documents found in the input directory.

• SET INTPARSE false; CREATE DB input.input.xml creates a database input with input.xml as initial XML document, which will be parsed with Java's default XML parser.

• <create-db name='simple'><hello>Universe</hello></create-db> creates a database named simple with an initial document <hello>Universe</hello>.

OPEN

Syntax OPEN [name] ([path])
<table>
<thead>
<tr>
<th>Command</th>
<th>Syntax</th>
<th>XML Syntax</th>
<th>Permission</th>
<th>Summary</th>
<th>Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commands</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OPEN</td>
<td>OPEN</td>
<td>&lt;open name='...' (path='...')/&gt;</td>
<td>READ</td>
<td>Opens the database specified by name. The documents to be opened can be specified by the [path] argument.</td>
<td>The command fails if the specified database does not exist, is currently being updated by another process or cannot be opened for some other reason.</td>
</tr>
<tr>
<td>CHECK</td>
<td>CHECK [input]</td>
<td>&lt;check input='...'/&gt;</td>
<td>READ/CREATE</td>
<td>This convenience command combines OPEN and CREATE_DB: if a database with the name input exists, it is opened. Otherwise, a new database is created; if the specified input points to an existing resource, it is stored as initial content.</td>
<td>The command fails if the addressed database could neither be opened nor created.</td>
</tr>
<tr>
<td>CLOSE</td>
<td>CLOSE</td>
<td>&lt;close/&gt;</td>
<td>READ</td>
<td>Closes the currently opened database.</td>
<td>The command fails if the database files could not be closed for some reason.</td>
</tr>
<tr>
<td>EXPORT</td>
<td>EXPORT [path]</td>
<td>&lt;export path='...'/&gt;</td>
<td>CREATE</td>
<td>Exports all documents in the database to the specified file path, using the serializer options specified by the EXPORTER option.</td>
<td>The command fails if no database is opened, if the target path points to a file or is invalid, if the serialization parameters are invalid, or if the documents cannot be serialized for some other reason.</td>
</tr>
<tr>
<td>CREATE INDEX</td>
<td>CREATE INDEX [TEXT</td>
<td>ATTRIBUTE</td>
<td>TOKEN</td>
<td>FULLTEXT]</td>
<td>&lt;create-index type='text</td>
</tr>
<tr>
<td>DROP INDEX</td>
<td>DROP INDEX [TEXT</td>
<td>ATTRIBUTE</td>
<td>TOKEN</td>
<td>FULLTEXT]</td>
<td></td>
</tr>
</tbody>
</table>

*Updated with Version 8.4: Token index added.*
Commands

**Commands**

**XML Syntax**: `<drop-index type='text|attribute|token|fulltext'/>`

**Permission**: WRITE

**Summary**: Drops the specified Value Index.

**Errors**: The command fails if no database is opened, if the specified index is unknown, or if it could not be deleted for some other reason.

**Administration**

**ALTER DB**

**Syntax**: ALTER DB [name] [newname]

**XML Syntax**: `<alter-db name='...' newname='...'/>`

**Permission**: CREATE

**Summary**: Renames the database specified by name to newname. newname must be a valid database name.

**Errors**: The command fails if the target database already exists, if the source database does not exist or is currently locked, or if it could not be renamed for some other reason.

**Examples**

- ALTER DB db tempdb renames the database db into tempdb.

**DROP DB**

**Syntax**: DROP DB [name]

**XML Syntax**: `<drop-db name='...'/>`

**Permission**: CREATE

**Summary**: Drops the database with the specified name. The Glob Syntax can be used to address more than one database.

**Errors**: The command fails if the specified database does not exist or is currently locked, or if the database could not be deleted for some other reason.

**CREATE BACKUP**

**Syntax**: CREATE BACKUP [name]

**XML Syntax**: `<create-backup name='...'/>`

**Permission**: CREATE

**Summary**: Creates a zipped backup of the database specified by name. The backup file will be suffixed with the current timestamp and stored in the database directory. The Glob Syntax can be used to address more than one database.

**Errors**: The command fails if the specified database does not exist, or if it could not be zipped for some other reason.

**Examples**

- BACKUP db creates a zip archive of the database db (e.g. db-2014-04-01-12-27-28.zip) in the database directory.

**RESTORE**

**Syntax**: RESTORE [name]

**XML Syntax**: `<restore name='...'/>`

**Permission**: CREATE

**Summary**: Restores a database with the specified name. The name may include the timestamp of the backup file.
The command fails if the specified backup does not exist, if the database to be restored is currently locked, or if it could not be restored for some other reason.

**INSPECT**

**Syntax**  
INSPECT

**XML Syntax**  
<inspect/>

**Permission**  
READ

**Summary**  
Performs some integrity checks on the opened database and returns a brief summary.

**DROP BACKUP**

**Syntax**  
DROP BACKUP [name]

**XML Syntax**  
<drop-backup name='...'/>

**Permission**  
CREATE

**Summary**  
Drops all backups of the database with the specified name. The Glob Syntax can be used to address more than one database.

**Examples**  
* DROP BACKUP abc* deletes the backups of all databases starting with the characters abc.

**SHOW BACKUPS**

**Syntax**  
SHOW BACKUPS

**XML Syntax**  
<show-backups/>

**Permission**  
CREATE

**Summary**  
Shows all database backups.

**COPY**

**Syntax**  
COPY [name] [newname]

**XML Syntax**  
<copy name='...' newname='...'/>

**Permission**  
CREATE

**Summary**  
Creates a copy of the database specified by name. newname must be a valid database name.

**Errors**  
The command fails if the target database already exists, or if the source database does not exist.

**INFO DB**

**Syntax**  
INFO DB

**XML Syntax**  
<info-db/>

**Permission**  
READ

**Summary**  
Shows information on the currently opened database.

**Errors**  
The command fails if no database is opened.

**INFO INDEX**

**Syntax**  
INFO INDEX ([ELEMMNAME|ATTRNAME|PATH|TEXT|ATTRIBUTE|TOKEN|FULLTEXT])

**XML Syntax**  
<info-index type='elemmname|attrname|path|text|attribute|token|fulltext'/>

**Permission**  
READ
<table>
<thead>
<tr>
<th>Summary</th>
<th>Shows information on the existing index structures. The output can be optionally limited to the specified index.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Errors</td>
<td>The command fails if no database is opened, or if the specified index is unknown.</td>
</tr>
</tbody>
</table>

### INFO STORAGE

<table>
<thead>
<tr>
<th>Syntax</th>
<th>INFO STORAGE [start end]</th>
</tr>
</thead>
<tbody>
<tr>
<td>XML Syntax</td>
<td><code>&lt;info-storage (start='...') (end='...')/&gt;</code></td>
</tr>
<tr>
<td>Permission</td>
<td><code>READ</code></td>
</tr>
<tr>
<td>Summary</td>
<td>Shows the internal main table of the currently opened database. An integer range may be specified as argument.</td>
</tr>
<tr>
<td>Errors</td>
<td>The command fails if no database is opened, or if one of the specified arguments is invalid.</td>
</tr>
</tbody>
</table>

### Querying

### LIST

<table>
<thead>
<tr>
<th>Syntax</th>
<th>LIST ([name] ([path]))</th>
</tr>
</thead>
<tbody>
<tr>
<td>XML Syntax</td>
<td><code>&lt;list (name='...') (path='...')/&gt;'</code></td>
</tr>
<tr>
<td>Permission</td>
<td><code>NONE</code></td>
</tr>
<tr>
<td>Summary</td>
<td>Lists all available databases. If name is specified, the resources of a database are listed. The output can be further restricted to the resources matching the specified path. If database resources are listed, the size is either the number of nodes (for XML resources) or the number of bytes (for binary resources).</td>
</tr>
<tr>
<td>Errors</td>
<td>The command fails if the optional database cannot be opened, or if the existing databases cannot be listed for some other reason.</td>
</tr>
</tbody>
</table>

### XQUERY

<table>
<thead>
<tr>
<th>Syntax</th>
<th>XQUERY [query]</th>
</tr>
</thead>
<tbody>
<tr>
<td>XML Syntax</td>
<td><code>&lt;xquery&gt;[query]&lt;/xquery&gt;</code></td>
</tr>
<tr>
<td>Permission</td>
<td>depends on query</td>
</tr>
<tr>
<td>Summary</td>
<td>Runs the specified query and prints the result.</td>
</tr>
<tr>
<td>Errors</td>
<td>The command fails if the specified query is invalid.</td>
</tr>
</tbody>
</table>

#### Examples
- XQUERY 1 to 10 returns the sequence (1, 2, 3, 4, 5, 6, 7, 8, 9, 10).
- SET RUNS 10; XQUERY 1 to 10 returns the results after having run the query 10 times.
- SET XMLPLAN true; XQUERY 1 to 10 returns the result and prints the query plan as XML.

### RETRIEVE

<table>
<thead>
<tr>
<th>Syntax</th>
<th>RETRIEVE [path]</th>
</tr>
</thead>
<tbody>
<tr>
<td>XML Syntax</td>
<td><code>&lt;retrieve path='...'/&gt;</code></td>
</tr>
<tr>
<td>Permission</td>
<td><code>READ</code></td>
</tr>
<tr>
<td>Summary</td>
<td>Retrieves a raw file from the opened database at the specified path.</td>
</tr>
<tr>
<td>Errors</td>
<td>The command fails if no database is opened, if the source path is invalid or if the data cannot not be retrieved for some other reason.</td>
</tr>
</tbody>
</table>
### FIND

<table>
<thead>
<tr>
<th>Syntax</th>
<th>FIND [query]</th>
</tr>
</thead>
<tbody>
<tr>
<td>XML Syntax</td>
<td>&lt;find&gt;[query]&lt;/find&gt;</td>
</tr>
<tr>
<td>Permission</td>
<td>READ</td>
</tr>
<tr>
<td>Summary</td>
<td>Builds and runs a query for the specified query terms. Keywords can be enclosed in quotes to look for phrases. The following modifiers can be used to further limit search: = looks for exact text nodes~ looks for approximate hits@= looks for exact attribute values@ looks for attributes</td>
</tr>
<tr>
<td>Errors</td>
<td>The command fails if no database is opened.</td>
</tr>
</tbody>
</table>

### TEST

<table>
<thead>
<tr>
<th>Syntax</th>
<th>TEST [path]</th>
</tr>
</thead>
<tbody>
<tr>
<td>XML Syntax</td>
<td>&lt;test path='...'/&gt;</td>
</tr>
<tr>
<td>Permission</td>
<td>ADMIN</td>
</tr>
<tr>
<td>Summary</td>
<td>Runs all XQUnit tests in the specified path. The path can point to a single file or a directory. Unit testing can also be triggered via -t on command line.</td>
</tr>
<tr>
<td>Errors</td>
<td>The command fails if at least one test fails.</td>
</tr>
<tr>
<td>Examples</td>
<td>TEST project/tests runs all tests in the directory project/tests.</td>
</tr>
</tbody>
</table>

### REPO INSTALL

<table>
<thead>
<tr>
<th>Syntax</th>
<th>REPO INSTALL [path]</th>
</tr>
</thead>
<tbody>
<tr>
<td>XML Syntax</td>
<td>&lt;repo-install path='...'/&gt;</td>
</tr>
<tr>
<td>Permission</td>
<td>CREATE</td>
</tr>
<tr>
<td>Summary</td>
<td>Installs the package with path path.</td>
</tr>
<tr>
<td>Errors</td>
<td>The command fails in the following cases:</td>
</tr>
<tr>
<td></td>
<td>• The package to be installed is not a xar file.</td>
</tr>
<tr>
<td></td>
<td>• The package to be installed does not exist or is already installed.</td>
</tr>
<tr>
<td></td>
<td>• The package descriptor is with invalid syntax.</td>
</tr>
<tr>
<td></td>
<td>• The package to be installed depends on a package which is not installed.</td>
</tr>
<tr>
<td></td>
<td>• The package is not supported by the current version of BaseX.</td>
</tr>
<tr>
<td></td>
<td>• A component of the package is already installed as part of another package.</td>
</tr>
</tbody>
</table>

### REPO LIST

<table>
<thead>
<tr>
<th>Syntax</th>
<th>REPO LIST</th>
</tr>
</thead>
<tbody>
<tr>
<td>XML Syntax</td>
<td>&lt;repo-list/&gt;</td>
</tr>
<tr>
<td>Permission</td>
<td>READ</td>
</tr>
<tr>
<td>Summary</td>
<td>Lists all installed packages.</td>
</tr>
</tbody>
</table>

### REPO DELETE

| Syntax   | REPO DELETE [name] |
**XML Syntax** `<repo-delete name='...'/>

**Permission** CREATE

**Summary** Deletes the package with name name, optionally followed by a version.

**Errors** The command fails if the package to be deleted participates in a dependency.

## Updates

### ADD

**Syntax** ADD (TO [path]) [input]

**XML Syntax** `<add (path='...')>[input]</add>`

**Permission** WRITE

**Summary** Adds a file, directory or XML string specified by input to the currently opened database at the specified path:

- input may either be a single XML document, a directory, a remote URL or a plain XML string.
- A document with the same path may occur than once in a database. If this is unwanted, the REPLACE command can be used.
- If a file is too large to be added in one go, its data structures will be cached to disk first. Caching can be enforced by turning the ADDCACHE option on.

If files are to be added to an empty database, it is usually faster to use the CREATE_DB command and specify the initial input as argument.

**Errors** The command fails if no database is opened, if one of the documents to be added is not well-formed, or if it could not be parsed for some other reason.

**Examples**
- ADD input.xml adds the file input.xml to the database.
- ADD TO temp/one.xml input.xml adds input.xml to the database and moves it to temp/one.xml.
- ADD TO target/ xmldir adds all files from the xmldir directory to the database in the target path.

### DELETE

**Syntax** DELETE [path]

**XML Syntax** `<delete path='...'/>`

**Permission** WRITE

**Summary** Deletes all documents from the currently opened database that start with the specified path.

**Errors** The command fails if no database is opened.

### RENAME

**Syntax** RENAME [path] [newpath]

**XML Syntax** `<rename path='...' newpath='...'/>

**Permission** WRITE

**Summary** Renames all document paths in the currently opened database that start with the specified path. The command may be used to either rename single documents or directories.

**Errors** The command fails if no database is opened, or if the target path is empty.
Examples

• RENAME one.xml two.xml renames the document one.xml to two.xml.

• RENAME / TOP moves all documents to a TOP root directory.

REPLACE

Syntax

REPLACE [path] [input]

XML Syntax

<replace path='...'>[input]</replace>

Permission WRITE

Summary

Replaces a document in the currently opened database, addressed by path, with the file or XML string specified by input, or adds a new document if the resource does not exist yet.

Errors

The command fails if no database is opened, if the specified path points to a database directory, or if the input is not found.

Examples

• REPLACE one.xml input.xml replaces the document one.xml with the contents of the file input.xml.

• REPLACE top.xml <xml/> replaces the document top.xml with the document <xml/>.

STORE

Syntax

STORE (TO [path]) [input]

XML Syntax

<store (path='...')>[input]</store>

Permission WRITE

Summary

Stores a raw file in the opened database to the specified path. input may either be a file reference, a remote URL, or a plain string. If the path denotes a directory, it needs to be suffixed with a slash (/).

Errors

The command fails if no database is opened, if the specified resource is not found, if the target path is invalid or if the data cannot not be written for some other reason.

OPTIMIZE

Syntax

OPTIMIZE (ALL)

XML Syntax

<optimize/> <optimize-all/>

Permission WRITE

Summary

Optimizes the index structures, meta data and statistics of the currently opened database:

• If ALL is specified, all database structures are completely reconstructed. The database size will be reduced, and all orphaned data will be deleted.

• Without ALL, only the outdated index structures and database statistics will be updated. If the database is completely up-to-date, nothing will be done.

• Database options will be adopted from the original database. Only AUTOOPTIMIZE and (if ALL is specified) UPDINDEX will be adopted from the current options.

Errors

The command fails if no database is opened, or if the currently opened database is a main-memory instance.

FLUSH

Syntax

FLUSH

XML Syntax

<flush/>

Permission WRITE
### Commands

| Summary | Explicitly flushes the buffers of the currently opened database to disk. This command is applied if AUTOFLUSH has been set to false. |
| Errors | The command fails if no database is opened. |

### Server Administration

#### SHOW SESSIONS

| Syntax | SHOW SESSIONS |
| XML Syntax | <show-sessions/> |
| Permission | ADMIN |
| Summary | Shows all sessions that are connected to the current server instance. |

#### SHOW USERS

| Syntax | SHOW USERS (ON [database]) |
| XML Syntax | <show-users (database='...')/> |
| Permission | ADMIN |
| Summary | Shows all users that are registered in the database. If a database is specified, all user will be shown for which a pattern was specified that matches the database name. |
| Errors | The command fails if the optional database could not be opened. |

#### KILL

| Syntax | KILL [target] |
| XML Syntax | <kill target='...'/> |
| Permission | ADMIN |
| Summary | Kills sessions of a user or an IP:port combination, specified by target. The Glob Syntax can be used to address more than one user. |
| Errors | The command fails if a user tried to kill his/her own session. |

### User Management

#### CREATE USER

| Syntax | CREATE USER [name] ([password]) |
| XML Syntax | <create-user name='...'>([password])</create-user> |
| Permission | ADMIN |
| Summary | Creates a user with the specified name and password. If no password is specified, it is requested via the chosen frontend (GUI or bash). |
| Errors | The command fails if the specified user already exists. |

#### ALTER USER

<p>| Syntax | ALTER USER [name] ([newname]) |
| XML Syntax | &lt;alter-user name='...' newname='...'/&gt; |
| Permission | ADMIN |
| Summary | Renames the user with the specified name to newname. |</p>
<table>
<thead>
<tr>
<th>Commands</th>
</tr>
</thead>
<tbody>
<tr>
<td>Errors</td>
</tr>
</tbody>
</table>

**ALTER PASSWORD**

| Syntax | ALTER PASSWORD [name] ([password]) |
|------------------------|
| XML Syntax | <alter-password name='...'>([password])</alter-password> |
| Permission | ADMIN |
| Summary | Alters the password of the user with the specified name. If no password is specified, it is requested via the chosen frontend (GUI or bash). |
| Errors | The command fails if the specified user does not exist. |

**DROP USER**

| Syntax | DROP USER [name] (ON [pattern]): |
|------------------------|
| XML Syntax | <drop-user name='...' (pattern='...')/> |
| Permission | ADMIN |
| Summary | Drops the user with the specified name. The Glob Syntax can be used to address more than one database or user. If a glob pattern is specified, only the pattern will be removed. |
| Errors | The command fails if admin is specified as user name, or if the specified user does not exist or is currently logged in. |

**GRANT**

| Syntax | GRANT [NONE|READ|WRITE|CREATE|ADMIN] (ON [pattern]) TO [user] |
|------------------------|
| XML Syntax | <grant name='...' permission='none|read|write|create|admin' (pattern='...')/> |
| Permission | ADMIN |
| Summary | Grants the specified permission to the specified user. The Glob Syntax can be used to address more than one user. If a glob pattern is specified, the permission will be applied to all databases that match this pattern. |
| Errors | The command fails if admin is specified as user name or if the specified user does not exist. |
| Examples | • GRANT READ TO JoeWinson grants READ permission to the user JoeWinson.  
• GRANT WRITE ON Wiki TO editor* grants WRITE permissions on the Wiki database to all users starting with the characters editor*. |

**PASSWORD**

| Syntax | PASSWORD ([password]) |
|------------------------|
| XML Syntax | <password> ([password]) </password> |
| Permission | NONE |
| Summary | Changes the password of the current user. If no password is specified, it is requested via the chosen frontend (GUI or bash). |

**General Commands**

**RUN**

<p>| Syntax | RUN [file] |
|------------------------|
| XML Syntax | &lt;run file='...'/&gt; |</p>
<table>
<thead>
<tr>
<th>Command</th>
<th>Syntax</th>
<th>XML Syntax</th>
<th>Permission</th>
<th>Summary</th>
<th>Errors</th>
<th>Examples</th>
</tr>
</thead>
</table>
| Commands | EXECUTE [input] | <execute>[input]</execute> | depends on input | Evaluates the specified input as command script. This command can be used to run several commands in a single transaction. | The command fails if the syntax of the specified input is invalid. It will be canceled as soon as one of the executed commands fails. | • EXECUTE "create db db1; create db db2"
• EXECUTE "<commands><create-db name='db1'/><create-db name='db2'/></commands>" both commands will create two databases db1 and db2 in a single transaction. |
| GET | GET [option] | <get (option='...')/> | NONE | Returns the current value of the Option specified via option. Global options can only be requested by users with ADMIN permissions. | The command fails if the specified option is unknown. |
| SET | SET [option] ([value]) | <set option='...'>([value])</set> | NONE | Sets the Option specified by option to a new value. Only local options can be modified. If no value is specified, and if the value is boolean, it will be inverted. | The command fails if the specified option is unknown or if the specified value is invalid. |
| INFO | INFO | <info/> | READ | Shows global information. | |
HELP

Syntax   HELP ([command])

XML Syntax  <help> ([command]) </help>

Permission  NONE

Summary  If command is specified, information on the specific command is printed; otherwise, all commands are listed.

Errors  The command fails if the specified command is unknown.

EXIT

Syntax   EXIT

XML Syntax  <exit/>

Permission  NONE

Summary  Exits the console mode.

QUIT

Syntax   QUIT

XML Syntax  <quit/>

Permission  NONE

Summary  Exits the console mode (alias of EXIT).

Changelog

Version 8.4

• Updated: CREATE INDEX, DROP INDEX, INFO INDEX: token index added

• Updated: INFO STORAGE: Query argument removed, start/end added to XML syntax.

• Updated: INFO INDEX: Token index added; index TAG renamed to ELEMNAME; index ATTNAME renamed to ATTRNAME

• Updated: OPTIMIZE: adopt original index options

Version 8.2

• Removed: CREATE EVENT, DROP EVENT and SHOW EVENTS command

Version 8.0

• Updated: commands for User Management

• Updated: OPEN: path argument added

• Removed: CS command

• Added: QUIT

Version 7.9

• Added: TEST runs XQUnit tests.

Version 7.7
Commands

- Updated: syntax of valid names.

Version 7.5
- Added: EXECUTE executes a command script.
- Added: INSPECT performs integrity checks.
- Added: automatic detection of Command Scripts.
- Removed: SHOW DATABASES; information is also returned by SHOW SESSIONS.
- Removed: OPEN: path argument.

Version 7.3
- Added: XML Syntax added.
- Updated: CHECK can now be used to create empty databases.
- Updated: Names and paths in OPEN and LIST are now specified as separate arguments.

Version 7.2.1
- Updated: permissions for GET and SET changed from READ to NONE.

Version 7.2
- Updated: CREATE INDEX, DROP INDEX (PATH argument removed. Path summary is always available now and updated with OPTIMIZE).
- Updated: permissions for REPO DELETE, REPO INSTALL and REPO LIST.

Version 7.1
- Updated: KILL (killing sessions by specifying IP:port)

Version 7.0
- Added: FLUSH, RETRIEVE, STORE.
- Updated: ADD: simplified arguments.
Chapter 16. Options

Read this entry online in the BaseX Wiki.

This page is linked from the Getting Started Section.

The options listed on this page influence the way how database commands are executed and XQuery expressions are evaluated. Options are divided into global options, which are valid for all BaseX instances, and local options, which are specific to a client or session. Values of options are either strings, numbers or booleans.

The .basex configuration file is parsed by every new local BaseX instance. It contains all global options and, optionally, local options at the end of the file.

Various ways exist to access and change options:

• The current value of an option can be requested with the GET command. Local options can be changed via SET. All values are static: They stay valid until they are changed once again by another operation. If an option is of type boolean, and if no value is specified, its current value will be inverted.

• Initial values for global options can also be specified via system properties, which can e.g. be passed on with the -D flag on command line, or using System.setProperty() before creating a BaseX instance. The specified keys need to be prefixed with org.basex. An example:

```java
java -Dorg.basex.CHOP=false -cp basex.jar org.basex.BaseX -c"get chop"
```

ə

• If using the Mac OS X packaged application then global options can be set within the Info.plist file within the Contents folder of the application package. For example:

```xml
<key>JVMOptions</key>
<array>
  <string>-Dorg.basex.CHOP=false</string>
</array>
```

• In XQuery, local options can be set via option declarations and pragmas (see XQuery Extensions).

If options are implicitly changed by operations in the GUI, the underlying commands will be listed in the Info View.

Global Options

Global options are constants. They can only be set in the configuration file or via system properties (see above). One exception is the DEBUG option, which can also be changed at runtime by users with admin permissions.

General Options

DEBUG

<table>
<thead>
<tr>
<th>Signature</th>
<th>DEBUG [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>false</td>
</tr>
<tr>
<td>Summary</td>
<td>Sends internal debug info to STDERR. This option can be turned on to get additional information for development and debugging purposes. It can also be triggered on command line via -d.</td>
</tr>
</tbody>
</table>

DBPATH

| Signature      | DBPATH [path] |
### Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Signature</th>
<th>Default</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Default</strong></td>
<td>(home)/BaseXData or (home)/data</td>
<td></td>
<td>Points to the directory in which all databases are located.</td>
</tr>
<tr>
<td><strong>REPOPATH</strong></td>
<td>REPOPATH [path]</td>
<td>(home)/BaseXRepo</td>
<td>Points to the Repository, in which all XQuery modules are located.</td>
</tr>
<tr>
<td><strong>LANG</strong></td>
<td>LANG [language]</td>
<td>English</td>
<td>Specifies the interface language. Currently, seven languages are available: 'English', 'German', 'French', 'Dutch', 'Italian', 'Japanese', and 'Vietnamese'.</td>
</tr>
<tr>
<td><strong>LANGKEY</strong></td>
<td>LANGKEY [boolean]</td>
<td>false</td>
<td>Prefixes all texts with the internal language keys. This option is helpful if BaseX is translated into another language, and if you want to see where particular texts are displayed.</td>
</tr>
<tr>
<td><strong>GLOBALLOCK</strong></td>
<td>GLOBALLOCK [boolean]</td>
<td>false</td>
<td>Controls if local (database) or global (process) locking will be used for managing read and write operations. The article on Transaction Management provides more details on concurrency control.</td>
</tr>
</tbody>
</table>

### Client/Server Architecture

<table>
<thead>
<tr>
<th>Option</th>
<th>Signature</th>
<th>Default</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HOST</strong></td>
<td>HOST [host]</td>
<td>localhost</td>
<td>This host name is used by the client when connecting to a server. This option can also be changed when running the client on command line via −n.</td>
</tr>
<tr>
<td><strong>PORT</strong></td>
<td>PORT [port]</td>
<td>1984</td>
<td>This port is used by the client when connecting to a server. This option can also be changed when running the client on command line via −p.</td>
</tr>
<tr>
<td><strong>SERVERPORT</strong></td>
<td>SERVERPORT [port]</td>
<td>1984</td>
<td></td>
</tr>
</tbody>
</table>
### Summary
This is the port the database server will be listening to. This option can also be changed when running the server on command line via `-p`.

### USER

<table>
<thead>
<tr>
<th>Signature</th>
<th>USER [name]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>empty</td>
</tr>
</tbody>
</table>
| Summary         | Represents a user name, which is used for accessing the server or an HTTP service:
  - The default value will be overwritten if a client specifies its own credentials.
  - If the default value is empty, login will only be possible if the client specifies credentials.
  - The option can also be changed on command line via `-U`. |

### PASSWORD

<table>
<thead>
<tr>
<th>Signature</th>
<th>PASSWORD [password]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>empty</td>
</tr>
</tbody>
</table>
| Summary         | Represents a password, which is used for accessing the server or an HTTP service:
  - The default value will be overwritten if a client specifies its own credentials.
  - If the default value is empty, login will only be possible if the client specifies credentials.
  - The option can also be changed on command line via `-P`.
  - Please note that it is a security risk to specify your password in plain text. |

### AUTHMETHOD

<table>
<thead>
<tr>
<th>Signature</th>
<th>AUTHMETHOD [method]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>Basic</td>
</tr>
<tr>
<td>Summary</td>
<td>Specifies the HTTP Authentication, which will be proposed by the HTTP server if a client sends an unauthorized request. Allowed values are Basic and Digest.</td>
</tr>
</tbody>
</table>

### SERVERHOST

| Signature       | SERVERHOST [host|ip] |
|-----------------|------------------|
| Default         | empty            |
| Summary         | This is the host name or ip address the server is bound to. If the option is set to an empty string (which is the default), the server will be open to all clients. |

### PROXYHOST

<table>
<thead>
<tr>
<th>Signature</th>
<th>PROXYHOST [host]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>empty</td>
</tr>
<tr>
<td>Summary</td>
<td>This is the host name of a proxy server. If the value is an empty string, it will be ignored.</td>
</tr>
</tbody>
</table>

### PROXYPORT

<table>
<thead>
<tr>
<th>Signature</th>
<th>PROXYPORT [port]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>0</td>
</tr>
<tr>
<td>Summary</td>
<td>This is the port number of a proxy server. If the value is set to 0, it will be ignored.</td>
</tr>
</tbody>
</table>
**Options**

### NONPROXYHOSTS

**Signature** NONPROXYHOSTS [hosts]

**Default** empty

**Summary** This is a list of hosts that should be directly accessed. If the value is an empty string, it will be ignored.

### IGNORECERT

**Signature** IGNORECERT [boolean]

**Default** false

**Summary** This option can be turned on to ignore untrusted certificates when connecting to servers. Please use this option carefully.

### TIMEOUT

**Signature** TIMEOUT [seconds]

**Default** 30

**Summary** Specifies the maximum time a read-only transaction may take. If an operation takes longer than the specified timeout, it will be aborted. Write operations will not be affected by this timeout, as this would corrupt the integrity of the database. The timeout is deactivated if the timeout is set to 0. It is ignored for operations with admin permissions.

### KEEPALIVE

**Signature** KEEPALIVE [seconds]

**Default** 600

**Summary** Specifies the maximum time a client will be remembered by the server. If there has been no interaction with a client for a longer time than specified by this timeout, it will be disconnected. Running operations will not be affected by this option. The keepalive check is deactivated if the value is set to 0.

### PARALLEL

**Signature** PARALLEL [number]

**Default** 8

**Summary** Denotes the maximum allowed number of parallel transactions. Note that a higher number of parallel operations may increase disk activity and thus slow down queries. In some cases, a single transaction may even give you better results than any parallel activity. The main reason for allowing parallel operations is to prevent slow transactions from blocking all other operations.

### LOG

**Signature** LOG [boolean]

**Default** true

**Summary** Turns Logging of server operations and HTTP requests on/off. This option can also be changed when running the server on command line via \( -z \).

### LOGMSGMAXLEN

**Signature** LOGMSGMAXLEN [length]

**Default** 1000
**Summary**
Specifies the maximum length of a single **log message**.

### HTTP Services
If BaseX is run as web servlet, the HTTP options must be specified in the `webapp/WEB-INF` directory and the `jetty.xml` and `web.xml` configuration files.

#### WEBPATH

**Signature**
WEBPATH [path]

**Default**
(home)/BaseXWeb or (home)/webapp

**Summary**
Points to the directory in which all the **Web Application** contents are stored, including XQuery, Script, RESTXQ and configuration files. This option is ignored if BaseX is deployed as **web servlet**.

#### RESTXQPATH

**Signature**
RESTXQPATH [path]

**Default**
empty

**Summary**
Points to the directory which contains the **RESTXQ** modules of a web application. Relative paths will be resolved against the **WEBPATH** directory.

#### CACHERESTXQ

**Signature**
CACHERESTXQ

**Default**
false

**Summary**
Caches **RESTXQ** modules once when starting the web application. The option is helpful in productive environments with a high load, but files should not be replaced while the web server is running.

#### RESTPATH

**Signature**
RESTPATH [path]

**Default**
empty

**Summary**
Points to the directory which contains XQuery files and command scripts, which can be evaluated via the **REST run operation**. Relative paths will be resolved against the **WEBPATH** directory.

#### HTTPLOCAL

**Signature**
HTTPLOCAL [boolean]

**Default**
false

**Summary**
By default, if BaseX is run as **Web Application**, a database server instance will be started as soon as the first HTTP service is called. The server can then be addressed by other BaseX clients in parallel to the HTTP services. If the option is set to **false**, the database server will be disabled.

#### STOPPORT

**Signature**
STOPPORT [port]

**Default**
8985

**Summary**
This is the port on which the **HTTP Server** can be locally closed:
- The listener for stopping the web server will only be started if the specified value is greater than 0.
- The option is ignored if BaseX is used as a **Web Application** or started via Maven.
Options

- This option can also be changed when running the HTTP server on command line via -s.

Create Options

General

**MAINMEM**

<table>
<thead>
<tr>
<th>Signature</th>
<th>MAINMEM [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>false</td>
</tr>
<tr>
<td>Summary</td>
<td>If this option is turned on, new databases will be exclusively created in main memory. Most queries will be evaluated faster in main memory mode, but all data is lost if BaseX is shut down. The value of this option will be assigned once to a new database, and cannot be changed after that.</td>
</tr>
</tbody>
</table>

**ADDCACHE**

<table>
<thead>
<tr>
<th>Signature</th>
<th>ADDCACHE [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>false</td>
</tr>
<tr>
<td>Summary</td>
<td>If this option is activated, data structures of documents will first be cached to disk before being added to the final database. This option is helpful when larger documents need to be added, and if the existing heuristics cannot estimate the input size (e.g. when adding directories or sending input streams).</td>
</tr>
</tbody>
</table>

Parsing

**CREATEFILTER**

<table>
<thead>
<tr>
<th>Signature</th>
<th>CREATEFILTER [filter]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>*.xml</td>
</tr>
<tr>
<td>Summary</td>
<td>File filter in the Glob Syntax, which is applied whenever new databases are created, or resources are added to a database.</td>
</tr>
</tbody>
</table>

**ADDDRICHIVES**

<table>
<thead>
<tr>
<th>Signature</th>
<th>ADDARCHIVES [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>true</td>
</tr>
<tr>
<td>Summary</td>
<td>If this option is set to true, files within archives (ZIP, GZIP, TAR, TGZ, DOCX, etc.) are parsed whenever new databases are created or resources are added to a database.</td>
</tr>
</tbody>
</table>

**ARCHIVENAME**

<table>
<thead>
<tr>
<th>Signature</th>
<th>ARCHIVENAME [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>false</td>
</tr>
<tr>
<td>Summary</td>
<td>If this option is set to true, the file name of parsed archives will be included in the document paths.</td>
</tr>
</tbody>
</table>

**SKIPCORRUPT**

<table>
<thead>
<tr>
<th>Signature</th>
<th>SKIPCORRUPT [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>false</td>
</tr>
<tr>
<td>Summary</td>
<td>Skips corrupt (i.e., not well-formed) files while creating a database or adding new documents. If this option is activated, document updates are slowed down, as all files will be parsed twice. Next, main memory consumption will be higher as parsed files will be cached in main memory.</td>
</tr>
<tr>
<td>Options</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>ADDRAW</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Signature</strong></td>
<td>ADDRAW [boolean]</td>
</tr>
<tr>
<td><strong>Default</strong></td>
<td>false</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>If this option is activated, and if new resources are added to a database, all files that are not filtered by the CREATEFILTER option will be added as raw files (i.e., in their binary representation).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PARSER</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Signature</strong></td>
<td>PARSER [type]</td>
</tr>
<tr>
<td><strong>Default</strong></td>
<td>XML</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>Defines a parser for importing new files to the database. Currently, 'XML', 'JSON', 'CSV', 'TEXT', 'HTML' are available as parsers. HTML will be parsed as normal XML files if Tagsoup is not found in the classpath.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CSVPARSER</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Signature</strong></td>
<td>CSVPARSER [options]</td>
</tr>
<tr>
<td><strong>Default</strong></td>
<td>empty</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>Specifies the way how CSV data will be parsed. The available options are listed in the CSV Module.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>JSONPARSER</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Signature</strong></td>
<td>JSONPARSER [options]</td>
</tr>
<tr>
<td><strong>Default</strong></td>
<td>empty</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>Specifies the way how JSON data will be parsed. The available options are listed in the JSON Module.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HTMLPARSER</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Signature</strong></td>
<td>HTMLPARSER [options]</td>
</tr>
<tr>
<td><strong>Default</strong></td>
<td>empty</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>Specifies the way how HTML data will be parsed. Available options are listed in the Parsers article.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TEXTPARSER</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Signature</strong></td>
<td>TEXTPARSER [options]</td>
</tr>
<tr>
<td><strong>Default</strong></td>
<td>empty</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>Specifies the way how TEXT data will be parsed. Available options are listed in the Parsers article.</td>
</tr>
</tbody>
</table>

| XML Parsing |  |

<table>
<thead>
<tr>
<th>CHOP</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Signature</strong></td>
<td>CHOP [boolean]</td>
</tr>
<tr>
<td><strong>Default</strong></td>
<td>true</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>Many XML documents include whitespaces that have been added to improve readability. This option controls the white-space processing mode of the XML parser:</td>
</tr>
</tbody>
</table>

- With the default value true, leading and trailing whitespaces from text nodes will be chopped and all empty text nodes will be discarded. |
- The flag should be turned off if a document contains mixed content. |
The flag can also be turned off on command line via `-w`.

If the `xml:space="preserve"` attribute is attached to an element, chopping will be turned off for all descendant text nodes. In the following example document, the whitespaces in the text nodes of the `text` element will not be chopped:

```xml
<xml>
  <title>
    Demonstrating the CHOP flag
  </title>
  <text xml:space="preserve">To <b>be</b>, or not to <b>be</b>, that is the question.</text>
</xml>
```

**STRIPNS**

<table>
<thead>
<tr>
<th>Signature</th>
<th>STRIPNS [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>false</td>
</tr>
<tr>
<td>Summary</td>
<td>Strips all namespaces from an XML document and all elements while parsing.</td>
</tr>
</tbody>
</table>

**INTPARSE**

<table>
<thead>
<tr>
<th>Signature</th>
<th>INTPARSE [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>false</td>
</tr>
<tr>
<td>Summary</td>
<td>Uses the internal XML parser instead of the standard Java XML parser. The internal parser is faster, more fault tolerant and supports common HTML entities out-of-the-box, but it does not support all features needed for parsing DTDs.</td>
</tr>
</tbody>
</table>

**DTD**

<table>
<thead>
<tr>
<th>Signature</th>
<th>DTD [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>false</td>
</tr>
<tr>
<td>Summary</td>
<td>Parses referenced DTDs and resolves XML entities. By default, this option is switched to <code>false</code>, as many DTDs are located externally, which may completely block the process of creating new databases. The <code>CATFILE</code> option can be changed to locally resolve DTDs.</td>
</tr>
</tbody>
</table>

**XINCLUDE**

<table>
<thead>
<tr>
<th>Signature</th>
<th>XINCLUDE [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>true</td>
</tr>
<tr>
<td>Summary</td>
<td>Resolves XInclude inclusion tags and merges referenced XML documents. By default, this option is switched to <code>true</code>. This option is only available if the standard Java XML Parser is used (see <code>INTPARSE</code>).</td>
</tr>
</tbody>
</table>

**CATFILE**

<table>
<thead>
<tr>
<th>Signature</th>
<th>CATFILE [path]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td><code>empty</code></td>
</tr>
<tr>
<td>Summary</td>
<td>Specifies a catalog file to locally resolve DTDs. See the entry on Catalog Resolvers for more details.</td>
</tr>
</tbody>
</table>

**Indexing**

The following options control the creation of index structures. The current values will be considered if a new database is created. See Indexes for more details.
<table>
<thead>
<tr>
<th>Option</th>
<th>Signature</th>
<th>Default</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEXTINDEX</td>
<td>TEXTINDEX [boolean]</td>
<td>true</td>
<td>Creates a text index whenever a new database is created. A text index speeds up queries with equality comparisons on text nodes. See Text Index for more details.</td>
</tr>
<tr>
<td>ATTRINDEX</td>
<td>ATTRINDEX [boolean]</td>
<td>true</td>
<td>Creates an attribute index whenever a new database is created. An attribute index speeds up queries with equality comparisons on attribute values. See Attribute Index for more details.</td>
</tr>
<tr>
<td>TOKENINDEX</td>
<td>TOKENINDEX [boolean]</td>
<td>true</td>
<td>Creates a token index whenever a new database is created. A token index speeds up searches for single tokens in attribute values. See Token Index for more details.</td>
</tr>
<tr>
<td>FTINDEX</td>
<td>FTINDEX [boolean]</td>
<td>false</td>
<td>Creates a full-text index whenever a new database is created. A full-text index speeds up queries with full-text expressions. See Full-Text Index for more details.</td>
</tr>
<tr>
<td>TEXTINCLUDE</td>
<td>TEXTINCLUDE [names]</td>
<td>empty</td>
<td>Defines name patterns for the parent elements of texts that are indexed. By default, all text nodes will be indexed. Name patterns are separated by commas. See Selective Indexing for more details.</td>
</tr>
<tr>
<td>ATTRINCLUDE</td>
<td>ATTRINCLUDE [names]</td>
<td>empty</td>
<td>Defines name patterns for the attributes to be indexed. By default, all attribute nodes will be indexed. Name patterns are separated by commas. See Selective Indexing for more details.</td>
</tr>
<tr>
<td>TOKENINCLUDE</td>
<td>TOKENINCLUDE [names]</td>
<td>empty</td>
<td>Defines name patterns for the attributes to be indexed. By default, tokens in all attribute nodes will be indexed. Name patterns are separated by commas. See Selective Indexing for more details.</td>
</tr>
<tr>
<td>FTINCLUDE</td>
<td>FTINCLUDE [names]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Options**

<table>
<thead>
<tr>
<th>Option</th>
<th>Signature</th>
<th>Default</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>empty</td>
<td></td>
<td>Defines name patterns for the parent elements of texts that are indexed. By default, all text nodes will be indexed. Name patterns are separated by commas. See <a href="#">Selective Indexing</a> for more details.</td>
</tr>
<tr>
<td>MAXLEN</td>
<td>MAXLEN [int]</td>
<td>96</td>
<td>Specifies the maximum length of strings that are to be indexed by the name, path, value, and full-text index structures. The value of this option will be assigned once to a new database, and cannot be changed after that.</td>
</tr>
<tr>
<td>MAXCATS</td>
<td>MAXCATS [int]</td>
<td>100</td>
<td>Specifies the maximum number of distinct values (categories) that will be stored together with the element/attribute names or unique paths in the Name Index or Path Index. The value of this option will be assigned once to a new database, and cannot be changed after that.</td>
</tr>
<tr>
<td>UPDINDEX</td>
<td>UPDINDEX [boolean]</td>
<td>false</td>
<td>If turned on, incremental indexing will be enabled:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• The current value of this option will be assigned to new databases. It can be changed for existing databases by running <code>OPTIMIZE</code> with the <code>ALL</code> keyword or <code>db:optimize($db, true())</code>.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• After each update, the value indexes will be refreshed as well. Incremental updates are currently not available for the full-text index and database statistics.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Find more details in the article on <a href="#">Index Structures</a>.</td>
</tr>
<tr>
<td>AUTOOPTIMIZE</td>
<td>AUTOOPTIMIZE [boolean]</td>
<td>false</td>
<td>If turned on, auto optimization will be applied to new databases:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• With each update, outdated indexes and database statistics will be recreated.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• As a result, the index structures will always be up-to-date.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• However, updates can take much longer, so this option should only be activated for medium-sized databases.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• The value of this option will be assigned once to a new database. It can be reassigned by running <code>OPTIMIZE</code> or <code>db:optimize</code>.</td>
</tr>
<tr>
<td>SPLITSIZE</td>
<td>SPLITSIZE [num]</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
### Options

| Summary | This option affects the construction of new value indexes. It controls the number of index build operations that are performed before writing partial index data to disk:  
|         | • The larger the assigned value is, the less splits will take place, and the more main memory will be required.  
|         | • By default, if the value is set to 0, some heuristics are applied, based on the current memory consumption. |

### Full-Text

#### STEMMING

<table>
<thead>
<tr>
<th>Signature</th>
<th>STEMMING [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>false</td>
</tr>
</tbody>
</table>

| Summary | If true, all tokens will be stemmed during full-text indexing, using a language-specific stemmer implementation. By default, tokens will not be stemmed. See Full-Text Index for more details. |

#### CASESENS

<table>
<thead>
<tr>
<th>Signature</th>
<th>CASESENS [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>false</td>
</tr>
</tbody>
</table>

| Summary | If true, the case of tokens will be preserved during full-text indexing. By default, case will be ignored (all tokens will be indexed in lower case). See Full-Text Index for more details. |

#### DIACRITICS

<table>
<thead>
<tr>
<th>Signature</th>
<th>DIACRITICS [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>false</td>
</tr>
</tbody>
</table>

| Summary | If set to true, diacritics will be preserved during full-text indexing. By default, diacritics will be removed. See Full-Text Index for more details. |

#### LANGUAGE

<table>
<thead>
<tr>
<th>Signature</th>
<th>LANGUAGE [lang]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>en</td>
</tr>
</tbody>
</table>

| Summary | The specified language will influence the way how an input text will be tokenized. This option is mainly important if tokens are to be stemmed, or if the tokenization of a language differs from Western languages. See Full-Text Index for more details. |

#### STOPWORDS

<table>
<thead>
<tr>
<th>Signature</th>
<th>STOPWORDS [path]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>empty</td>
</tr>
</tbody>
</table>

| Summary | A new full-text index will drop tokens that are listed in the specified stopword list. A stopword list may decrease the size of the full text index. See Full-Text Index for more details. |

### Query Options

#### QUERYINFO

<table>
<thead>
<tr>
<th>Signature</th>
<th>QUERYINFO [boolean]</th>
</tr>
</thead>
</table>


Options

| Default | false |
| Summary | Prints more information on internal query rewritings, optimizations, and performance. By default, this info is shown in the Info View in the GUI. It can also be activated on command line via `-V`. |

**XQUERY3**

| Signature | XQUERY3 |
| Default | true |
| Summary | Enables all XQuery 3.0 features supported by BaseX. If this option is set to false, the XQuery parser will only accept expressions of the XQuery 1.0 specification. |

**MIXUPDATES**

| Signature | MIXUPDATES |
| Default | false |
| Summary | Allows queries to both contain updating and non-updating expressions. All updating constraints will be turned off, and nodes to be returned will be copied before they are modified by an updating expression. – By default, this option is set to false, because the XQuery Update Facility does not allow an updating query to return results. |

**BINDINGS**

| Signature | BINDINGS [vars] |
| Default | empty |
| Summary | Contains external variables to be bound to a query. The string must comply with the following rules:
* Variable names and values must be separated by equality signs.
* Multiple variables must be delimited by commas.
* Commas in values must be duplicated.
* Variables may optionally be introduced with a leading dollar sign.
* If a variable uses a namespace different to the default namespace, it can be specified with the Clark Notation or Expanded QName Notation.

This option can also be used on command line with the flag `-b`.

**Examples**

* `$a=1,$b=2` binds the values 1 and 2 to the variables $a and $b
* `a=1,,2` binds the value 1,2 to the variable $a
* `{URI}a=x` binds the value x to the variable $a with the namespace URI.
* In the following Command Script, the value `hello world!` is bound to the variable `$GREETING`:

```
SET BINDINGS GREETING="hello world!"
XQUERY declare variable $GREETING external; $GREETING
```

**QUERYPATH**

| Signature | QUERYPATH [path] |
| Default | empty |
Summary
Contains the path (base URI) to the executed query (default: empty). This directory will be used to resolve relative paths to documents, query modules, and other resources addressed in a query.

**INLINELIMIT**

**Signature** | INLINELIMIT
---|---
**Default** | 100
**Summary** | The XQuery compiler inlines functions to speed up query evaluation. Inlining will only take place if a function body is not too large (i.e., if it does not contain too many expressions). With this option, this maximum number of expressions can be specified. Function inlining can be turned off by setting the value to 0. The limit can be locally overridden via the `%basex:inline` annotation.

**TAILCALLS**

**Signature** | TAILCALLS
---|---
**Default** | 256
**Summary** | Specifies how many stack frames of tail-calls are allowed on the stack at any time. When this limit is reached, tail-call optimization takes place and some call frames are eliminated. The feature can be turned off by setting the value to -1.

**DEFAULTDB**

**Signature** | DEFAULTDB
---|---
**Default** | false
**Summary** | If this option is turned on, paths specified in the `fn:doc` and `fn:collection` functions will first be resolved against a database that has been opened in the global context outside the query (e.g. by the `OPEN` command). If the path does not match any existing resources, it will be resolved as described in the article on accessing database resources.

**FORCERECREATE**

**Signature** | FORCERECREATE [boolean]
---|---
**Default** | false
**Summary** | By activating this option, database instances will be created with the XQuery functions `fn:doc` and `fn:collection`.

**CHECKSTRINGS**

**Signature** | CHECKSTRINGS [boolean]
---|---
**Default** | true
**Summary** | If this option is turned off, strings from external sources will be adopted as is, i.e., without being checked for valid XML characters:

- This option affects Java Bindings and the string conversion and input functions `archive:create`, `archive:extract-text`, `archive:update`, `convert:binary-to-string`, `fetch:text`, `file:read-text`, and `zip:text-entry`.
- Please be aware that an inconsiderate use of this option may cause unexpected behavior when storing or outputting strings.

**LSERROR**

**Signature** | LSERROR [error]
### Options

<table>
<thead>
<tr>
<th>Default</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Summary</strong></td>
<td>This option specifies the maximum Levenshtein error for the BaseX-specific fuzzy match option. See the page on Full-Texts for more information on fuzzy querying.</td>
</tr>
</tbody>
</table>

---

#### RUNQUERY

<table>
<thead>
<tr>
<th><strong>Signature</strong></th>
<th>RUNQUERY [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Default</strong></td>
<td>true</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>Specifies if a query will be executed or parsed only. This option can also be changed on command line via (-R).</td>
</tr>
</tbody>
</table>

---

#### RUNS

<table>
<thead>
<tr>
<th><strong>Signature</strong></th>
<th>RUNS [num]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Default</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>Specifies how often a query will be evaluated. The result is serialized only once, and the measured times are averages of all runs. This option can also be changed on command line via (-r).</td>
</tr>
</tbody>
</table>

---

#### Serialization Options

##### SERIALIZE

<table>
<thead>
<tr>
<th><strong>Signature</strong></th>
<th>SERIALIZE [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Default</strong></td>
<td>true</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>Results of XQuery expressions will be serialized if this option is turned on. For debugging purposes and performance measurements, this option can be set to false. It can also be turned off on command line via (-z).</td>
</tr>
</tbody>
</table>

##### SERIALIZER

<table>
<thead>
<tr>
<th><strong>Signature</strong></th>
<th>SERIALIZER [params]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Default</strong></td>
<td>empty</td>
</tr>
</tbody>
</table>
| **Summary** | Parameters for serializing query results. The string must comply with the following rules:  
- Variable names and values must be separated by equality signs.  
- Multiple variables must be delimited by commas.  
- Commas in values must be duplicated.  

The option can also be used on command line with the flag \(-s\). |

| **Examples** |  
- encoding=US-ASCII,omit-xml-declaration=no : sets the encoding to US-ASCII and prints the XML declaration.  
- item-separator=,, : separates serialized items by a single comma. |

---

##### EXPORTER

<table>
<thead>
<tr>
<th><strong>Signature</strong></th>
<th>EXPORTER [params]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Default</strong></td>
<td>empty</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>Contains parameters for exporting all resources of a database. Keys and values are separated by equality signs, multiple parameters are delimited by commas. See Serialization for more details.</td>
</tr>
</tbody>
</table>
### Options

#### XMLPLAN

<table>
<thead>
<tr>
<th>Signature</th>
<th>XMLPLAN [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>false</td>
</tr>
<tr>
<td>Summary</td>
<td>Prints the execution plan of an XQuery expression in its XML representation. This option can also be activated on command line via <code>-x</code>.</td>
</tr>
</tbody>
</table>

#### COMPPLAN

<table>
<thead>
<tr>
<th>Signature</th>
<th>COMPPLAN [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>true</td>
</tr>
<tr>
<td>Summary</td>
<td>Generates the query plan, which can be activated via XMLPLAN, before or after query compilation. This option can also be activated on command line via <code>-X</code>.</td>
</tr>
</tbody>
</table>

#### DOTPLAN

<table>
<thead>
<tr>
<th>Signature</th>
<th>DOTPLAN [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>false</td>
</tr>
<tr>
<td>Summary</td>
<td>Saves the query plan of an XQuery expression as <code>.dot</code> file in the current working directory, using the graph description language. The output file can e.g. be visualized with Graphviz.</td>
</tr>
</tbody>
</table>

#### DOTCOMPACT

<table>
<thead>
<tr>
<th>Signature</th>
<th>DOTCOMPACT [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>false</td>
</tr>
<tr>
<td>Summary</td>
<td>Chooses a compact dot representation.</td>
</tr>
</tbody>
</table>

#### Other Options

#### AUTOFLUSH

<table>
<thead>
<tr>
<th>Signature</th>
<th>AUTOFLUSH [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>true</td>
</tr>
<tr>
<td>Summary</td>
<td>Flushes database buffers to disk after each update. If this option is set to false, bulk operations (multiple single updates) will be evaluated faster. As a drawback, the chance of data loss increases if the database is not explicitly flushed via the FLUSH command.</td>
</tr>
</tbody>
</table>

#### WRITEBACK

<table>
<thead>
<tr>
<th>Signature</th>
<th>WRITEBACK [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>false</td>
</tr>
<tr>
<td>Summary</td>
<td>Propagates updates on main-memory instances of files that have been retrieved via <code>fn:doc</code> and <code>fn:collection</code> back to disk. This option can also be activated on command line via <code>-u</code>. Please note that, when turning this option on, your original files will not be backed up.</td>
</tr>
</tbody>
</table>

#### MAXSTAT

<table>
<thead>
<tr>
<th>Signature</th>
<th>MAXSTAT [num]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>30</td>
</tr>
<tr>
<td>Summary</td>
<td>Specifies the maximum number of index occurrences printed by the INFO INDEX command.</td>
</tr>
</tbody>
</table>
Options

Changelog

Version 8.4
- Added: TOKENINDEX, TOKENINCLUDE
- Added: SPLITSIZE (replacing INDEXSPLITSIZE and FTINDEXSPLITSIZE)
- Removed: INDEXSPLITSIZE, FTINDEXSPLITSIZE

Version 8.3
- Added: CACHERESTXQ, TEXTINCLUDE, ATTRINCLUDE, FTINCLUDE, ARCHIVENAME

Version 8.2
- Removed: EVENTPORT, CACHEQUERY

Version 8.1
- Added: IGNORECERT, RESTPATH

Version 8.0
- Added: MIXUPDATES, AUTOOPTIMIZE, AUTHMETHOD, XINCLUDE
- Updated: PROXYPORT: default set to 0; will be ignored. PROXYHOST, NONPROXYHOSTS: empty strings will be ignored.

Version 7.8.1
- Updated: ADDARCHIVES: parsing of TAR and TGZ files.

Version 7.8
- Added: CSVPARSER, JSONPARSER, TEXTPARSER, HTMLPARSER, INLINELIMIT, TAILCALLS, DEFAULTDB, RUNQUERY
- Updated: WRITEBACK only applies to main-memory document instances.
- Updated: DEBUG option can be changed at runtime by users with admin permissions.
- Updated: default of INTPARSE is now false.
- Removed: HTMLOPT (replaced with HTMLPARSER), PARSEROPT (replaced with parser-specific options), DOTDISPLAY, DOTTY

Version 7.7
- Added: ADDCACHE, CHECKSTRINGS, FTINDEXSPLITSIZE, INDEXSPLITSIZE

Version 7.6
- Added: GLOBALLOCK
- Added: store local options in configuration file after # Local Options comments.

Version 7.5
- Added: options can now be set via system properties
- Added: a pragma expression can be used to locally change database options
Options

- **Added**: USER, PASSWORD, LOG, LOGMSGMAXLEN, WEBPATH, RESTXQPATH HTTPLOCAL, CREATEONLY, STRIPNS

- **Removed**: HTTPPATH; HTTPPORT: jetty.xml configuration file is used instead

- **Removed**: global options cannot be changed anymore during the lifetime of a BaseX instance

Version 7.3

- **Updated**: KEEPALIVE, TIMEOUT: default values changed

- **Removed**: WILDCARDS; new index supports both fuzzy and wildcard queries

- **Removed**: SCORING; new scoring model will focus on lengths of text nodes and match options

Version 7.2

- **Added**: PROXYHOST, PROXYPORy, NONPROXYHOSTS, HTMLOPT

- **Updated**: TIMEOUT: ignore timeout for admin users

Version 7.1

- **Added**: ADDRAW, MAXLEN, MAXCATS, UPDINDEX

- **Updated**: BINDINGS

Version 7.0

- **Added**: SERVERHOST, KEEPALIVE, AUTOFLUSH, QUERYPATH
Part IV. Integration
Chapter 17. Integrating oXygen

Read this entry online in the BaseX Wiki.

This tutorial is part of the Getting Started Section. It describes how to access BaseX from the oXygen XML Editor. Currently, there are two variants how to use BaseX in oXygen:

- Resources in databases can be opened and modified.
- XPath/XQuery expressions can be run by the query processor of BaseX.

Access Database Resources

Preparations

- First, start the BaseX WebDAV service.

Configuration

1. Go to menu Options → Preferences → Data Sources
2. In the Connections panel, click the New button (+)
3. Enter "BaseX-WebDAV" as connection name
4. Select "WebDAV" in the Data Source combo box
5. Fill in the appropriate connection details. Below, the default values are shown:
   - Set the URL to http://localhost:8984/webdav
   - Set the user name to admin
   - Set the password to admin
6. Now press OK, and your Data Source is ready for use

You can now open single database files as follows:

- Choose File → Open URL...

   - Enter the corresponding user name and password (if needed), the URL of the BaseX WebDAV Server, and then click "Browse".
Perform Queries

One-Time Setup

Preparations

1. Download one of the complete BaseX distributions (ZIP, EXE).

2. Charles Foster’s XQJ implementation provides a default (client/server) and a local driver. If you want to use the first flavor, you need to start a BaseX Server instance.

Configure Data Source

1. Start oXygen and select Options → Preferences → Data Sources.

2. In the Data Sources section, add a new data source using the New button (+).

3. Enter "BaseX-XQJ" as Name and select XQuery API for Java (XQJ) as Type.

4. Add the following JAR files (downloaded in Preparations procedure) with the Add Files Button. The versions of the JAR files may differ.
   
   - lib/xqj-api-1.0.jar
   - lib/xqj2-0.2.0.jar
   - lib/basex-xqj-1.3.0.jar
   - basex.jar, if you want to use the local XQJ driver

5. Under "Driver class", choose the default or the local XQJ driver (net.xqj.basex.BaseXXQDataSource vs. net.xqj.basex.local.BaseXXQDataSource).

6. Click OK.
Configure Connection

1. In the Connections section (in the same Preferences dialog from the Configure Data Source procedure), click New (+).
2. Enter "BaseX" as Name and select "BaseX-XQJ" as Data Source.
3. If you use the default driver, enter these values in the Connection Details section:
   • port: 1984
   • serverName: localhost
   • user: admin
   • password: admin
4. Click OK to complete the connection configuration.
5. Click OK again to close the Preferences dialog.

Configure New Transformation Scenario

1. Select Window → Show View → Transformation Scenarios.
2. Select the XQuery transformation tree entry, and click + to add a new scenario.
   If this entry does not appear in the tree, click + and select XQuery transformation in the dropdown list.
3. Enter a Name and an optional XML and XQuery URL (e.g. your query document/file).
4. Choose "BaseX" as Transformer from the combo box.
5. Click OK to complete the scenario configuration.

Execute Query

After the one-time setup steps are complete, execute a query using the transformation scenario. Start the transformation by clicking the red Apply associated scenarios button in the Transformation Scenarios window.

The results should immediately occur in the result panel.

--CG 14:45, 29 January 2015 (CET)
Chapter 18. Integrating Eclipse

Read this entry online in the BaseX Wiki.

This article is part of the Getting Started Section. It describes how to run XPath/XQuery code from within the Eclipse IDE.

Another article describes how to compile and run BaseX with Eclipse.

Installation

The following steps apply to all operating systems:

• Install Version 3.7 (Indigo) of Eclipse: http://www.eclipse.org. Please note that more recent versions may work as well, but haven’t been tested so far.

• download your favorite BaseX distribution (JAR, ZIP, EXE): http://basex.org/download/

Windows

It should be sufficient to install the official XQuery Development Tools Plugin (XQDT): http://www.xqdt.org/
Update Site: http://download.eclipse.org/webtools/incubator/repository/xquery/milestones/

Linux

• First, install the Dynamic Languages Toolkit (DLTK) Update Site: http://download.eclipse.org/releases/indigo/

• Next, install Marklogic’s XQDT Dropin

Mac OS X

• Install Marklogic’s XQDT Dropin
Setting up

Use BaseX as query processor in Eclipse. You can set up the XQuery interpreter as standalone or client version, as shown on the screenshot:

Setting up as Standalone

1. Start Eclipse and go to Preferences → XQuery → Interpreters.
2. Add a new Interpreter with the Add button.
3. Enter "BaseX" as name and choose "Java XQuery Engine" as Interpreter type.
4. Point Interpreter JAR/WAR to the BaseX JAR archive.
5. Choose org.basex.BaseX as Main class.

Setting up as Client

1. Start Eclipse and go to Preferences → XQuery → Interpreters.
2. Add a new Interpreter with the Add button.
3. Enter "BaseX" as name and choose "Java XQuery Engine" as Interpreter type.
4. Point Interpreter JAR/WAR to the BaseX JAR archive.
5. Choose org.basex.BaseXClient as Main class.
Usage

The query execution works as follows:

1. Create a new XQuery Project with File $\rightarrow$ New $\rightarrow$ XQuery Project.
2. Add a new XQuery Module with File $\rightarrow$ New $\rightarrow$ XQuery Module.
3. Edit your XQuery Module and execute it with Run.
4. The results are displayed in the Console window of Eclipse.
Part V. Query Features
**Chapter 19. XQuery**

Welcome to the Query Portal, which is one of the Main Sections of this documentation. BaseX provides an implementation of the W3 XPath and XQuery languages, which are tightly coupled with the underlying database store. However, the processor is also a flexible general purpose processor, which can access local and remote sources. High conformance with the official specifications is one of our main objectives, as the results of the XQuery Test Suite demonstrate. This section contains information on the query processor and its extensions:

**XQuery 3.0 and XQuery 3.1**

Features of the new XQuery Recommendations.

**XQuery Extensions**

Specifics of the BaseX XQuery processor.

**Module Library**

Additional functions included in the internal modules.

**Java Bindings**

Accessing and calling Java code from XQuery.

**Repository**

Install and manage XQuery and Java modules.

**Full-Text**

How to use BaseX as a full-fledged full-text processor.

**Updates**

Updating databases and local resources via XQuery Update.

**Indexes**

Available index structures and their utilization.

**Serialization**

Serialization parameters supported by BaseX.

**Errors**

Errors raised by XQuery expressions.
Chapter 20. XQuery 3.0

Read this entry online in the BaseX Wiki.

This article is part of the XQuery Portal. It summarizes the most interesting features of the XQuery 3.0 Recommendations. XQuery 3.0 is fully supported by BaseX.

Enhanced FLWOR Expressions

Most clauses of FLWOR expressions can now be specified in an arbitrary order: additional let and for clauses can be put after a where clause, and multiple where, order by and group by statements can be used. This means that many nested loops can now be rewritten to a single FLWOR expression.

Example:

```xquery
for $country in db:open('factbook')//country
  where $country/@population > 100000000
for $city in $country//city[population > 100000]
group by $name := $country/name[1]
count $id
return <country id='{ $id }' name='{ $name }'>{ $city/name }</country>
group by
```

FLWOR expressions have been extended to include the group by clause, which is well-established in SQL. group by can be used to apply value-based partitioning to query results:

XQuery:

```xquery
for $ppl in doc('xmark')//people/person
let $ic := $ppl/profile/@income
let $income := if($ic < 30000) then
  "challenge"
else if($ic >= 30000 and $ic < 100000) then
  "standard"
else if($ic >= 100000) then
  "preferred"
else
  "na"
group by $income
order by $income
return element { $income } { count($ppl) }
```

This query is a rewrite of Query #20 contained in the XMark Benchmark Suite to use group by. The query partitions the customers based on their income.

Result:

```xml
<challenge>4731</challenge>
<na>12677</na>
<preferred>314</preferred>
<standard>7778</standard>
```

In contrast to the relational GROUP BY statement, the XQuery counterpart concatenates the values of all non-grouping variables that belong to a specific group. In the context of our example, all nodes in //people/person that belong to the preferred partition are concatenated in $ppl after grouping has finished. You can see this effect by changing the return statement to:
... return element { $income } { $ppl }

Result:

<challenge>
  <person id="person0"/>
    <name>Kasidit Treweek</name>
    ...
  <person id="personX"/>
  ...
</challenge>

Moreover, a value can be assigned to the grouping variable. This is shown in the following example:

XQuery:

```xml
let $data :=
  <xml>
    <person country='USA' name='John'/>
    <person country='USA' name='Jack'/>
    <person country='Germany' name='Johann'/>
  </xml>
for $person in $data/person
group by $country := $person/@country/string()
return element persons {
  attribute country { $country },
  $person/@name ! element name { data() }
}
```

Result:

<persons country="USA">
  <name>John</name>
  <name>Jack</name>
</persons>
<persons country="Germany">
  <name>Johann</name>
</persons>

**count**

A new `count` clause enhances the FLWOR expression with a variable that enumerates the iterated tuples.

```xml
for $n in (1 to 10)[. mod 2 = 1]
count $c
return <number count="{$c}" number="{$n}"/>
```

**allowing empty**

The `allowing empty` provides functionality similar to outer joins in SQL:

```xml
for $n allowing empty in ()
return 'empty? ' || empty($n)
```

**window**

Window clauses provide a rich set of variable declarations to process sub-sequences of iterated tuples. An example:
for tumbling window $w$ in (2, 4, 6, 8, 10, 12, 14)
  start at $s$ when fn:true()
  only end at $e$ when $e - $s eq 2
return <window>{$w$}</window>

More information on window clauses, and all other enhancements, can be found in the specification.

**Function Items**

One of the most distinguishing features added in XQuery 3.0 are function items, also known as lambdas or lambda functions. They make it possible to abstract over functions and thus write more modular code.

**Examples:**

Function items can be obtained in three different ways:

- Declaring a new inline function:
  
  let $f := function($x, $y) { $x + $y }
  return $f(17, 25)

  **Result:** 42

- Getting the function item of an existing (built-in or user-defined) XQuery function. The arity (number of arguments) has to be specified as there can be more than one function with the same name:
  
  let $f := math:pow#2
  return $f(5, 2)

  **Result:** 25

- Partially applying another function or function item. This is done by supplying only some of the required arguments, writing the placeholder `?` in the positions of the arguments left out. The produced function item has one argument for every placeholder.
  
  let $f := fn:substring(?, 1, 3)
  return ($f('foo123'), $f('bar456'))

  **Result:** foo bar

Function items can also be passed as arguments to and returned as results from functions. These so-called Higher-Order Functions like `fn:map` and `fn:fold-left` are discussed in more depth on their own Wiki page.

**Simple Map Operator**

The simple map operator ! provides a compact notation for applying the results of a first to a second expression: the resulting items of the first expression are bound to the context item one by one, and the second expression is evaluated for each item. The map operator may be used as replacement for FLWOR expressions:

**Example:**

```xquery
(: Simple map notation :)  
(1 to 10) ! element node { . },
(: FLWOR notation :)  
for $i$ in 1 to 10
return element node ( $i $ )
```

A map operator is defined to be part of a path expression, which may now mix path and map operators. In contrast to the path operator, the results of the map operator will not be made duplicate-free and returned in document order.
**Try/Catch**

The *try/catch* construct can be used to handle errors at runtime:

**Example:**

```xquery
try {
  1 + '2'
} catch err:XPTY0004 {
  'Typing error: ' || $err:description
} catch * {
  'Error [' || $err:code || ']: ' || $err:description
}
```

**Result:** Typing error: '+' operator: number expected, xs:string found.

Within the scope of the catch clause, a number of variables are implicitly declared, giving information about the error that occurred:

- `$err:code`: error code
- `$err:description`: error message
- `$err:value`: value associated with the error (optional)
- `$err:module`: URI of the module where the error occurred
- `$err:line-number`: line number where the error occurred
- `$err:column-number`: column number where the error occurred
- `$err:additional`: error stack trace

**Switch**

The *switch* statement is available in many other programming languages. It chooses one of several expressions to evaluate based on its input value.

**Example:**

```xquery
for $fruit in ("Apple", "Pear", "Peach")
return switch ($fruit)
  case "Apple" return "red"
  case "Pear" return "green"
  case "Peach" return "pink"
  default return "unknown"
```

**Result:** red green pink

The expression to evaluate can correspond to multiple input values.

**Example:**

```xquery
for $fruit in ("Apple", "Cherry")
return switch ($fruit)
  case "Apple"
  case "Cherry"
    return "red"
  case "Pear"
    return "green"
  case "Peach"
```
return "pink"
default
  return "unknown"

**Result:** red red

## Expanded QNames

A *QName* can now be directly prefixed with the letter "Q" and a namespace URI in the Clark Notation.

**Examples:**

- Q{http://www.w3.org/2005/xpath-functions/math}pi() returns the number $\pi$
- Q{java:java.io.FileOutputStream}new("output.txt") creates a new Java file output stream

## Namespace Constructors

New namespaces can now be created via so-called 'Computed Namespace Constructors'.

```
<element node { namespace pref { 'http://url.org/' } } />
```

## String Concatenations

Two vertical bars `||` (also named *pipe characters*) can be used to concatenate strings. This operator is a shortcut for the *fn:concat()* function.

```
'Hello' || '' || 'Universe'
```

## External Variables

Default values can now be attached to external variable declarations. This way, an expression can also be evaluated if its external variables have not been bound to a new value.

```
declare variable $user external := "admin";
"User:" || $user
```

## Serialization

Serialization parameters can now be defined in XQuery expressions. Parameters are placed in the query prolog and need to be specified as option declarations, using the *output* prefix.

**Example:**

```
declare namespace output = "http://www.w3.org/2010/xslt-xquery-serialization";
declare option output:omit-xml-declaration "no";
declare option output:method "xhtml";
<html/>
```

**Result:** `<?xml version="1.0" encoding="UTF-8"?><html/></html>`

In BaseX, the *output* prefix is statically bound and can thus be omitted. Note that all namespaces need to be specified when using external APIs, such as XQJ.

## Context Item

The context item can now be specified in the prolog of an XQuery expression:
Example:

```xml
declare context item := document {
    <xml>
        <text>Hello</text>
        <text>World</text>
    </xml>
};
for $t in .//text()
return string-length($t)
```

**Result:** 5 5

### Annotations

XQuery 3.0 introduces annotations to declare properties associated with functions and variables. For instance, a function may be declared `%public`, `%private`, or `%updating`.

**Example:**

```xml
declare %private function local:max($x1, $x2) {
    if($x1 > $x2) then $x1 else $x2
};
local:max(2, 3)
```

### Functions

The following functions have been added in the [XQuery 3.0 Functions and Operators Specification](#):

- fn:analyze-string
- fn:available-environment-variables
- fn:element-with-id
- fn:environment-variable
- fn:filter
- fn:fold-left
- fn:fold-right
- fn:for-each
- fn:for-each-pair
- fn:format-date
- fn:format-dateTime
- fn:format-integer
- fn:format-number
- fn:format-time
- fn:function-arity
- fn:function-lookup
- fn:function-name
- fn:generate-id
- fn:has-children
- fn:head
- fn:innermost
- fn:outermost
- fn:parse-xml
- fn:parse-xml-fragment
- fn:path
- fn:serialize
- fn:tail
- fn:unparsed-text
- fn:unparsed-text-available
- fn:unparsed-text-lines
- fn:uri-collection

New signatures have been added for the following functions:

- fn:document-uri
- fn:string-join
- fn:node-name
- fn:round
- fn:data

### Changelog

**Version 8.4**

- Added: `%non-deterministic`

**Version 8.0**

- Added: `%basex:inline, %basex:lazy`

**Version 7.7**

- Added: **Enhanced FLWOR Expressions**

**Version 7.3**
• Added: Simple Map Operator

Version 7.2

• Added: Annotations

• Updated: Expanded QNames

Version 7.1

• Added: Expanded QNames, Namespace Constructors

Version 7.0

• Added: String Concatenations
Chapter 21. Higher-Order Functions

This page talks about higher-order functions introduced with XQuery 3.0. The BaseX-specific Higher-Order Functions Module contains some additional useful functions.

Function Items

Probably the most important new feature in XQuery 3.0 are function items, i.e., items that act as functions, but can also be passed to and from other functions and expressions. This feature makes functions first-class citizens of the language.

The XQuery 3.0 page goes into details on how function items can be obtained.

Function Types

Like every XQuery item, function items have a sequence type. It can be used to specify the arity (number of arguments the function takes) and the argument and result types.

The most general function type is function(*). It's the type of all function items. The following query for example goes through a list of XQuery items and, if it is a function item, prints its arity:

```xml
for $item in (1, "foo", fn:concat#3, function($a) { 42 * $a })
where $item instance of function(*)
return fn:function-arity($item)
```

Result: 3 1

The notation for specifying argument and return types is quite intuitive, as it closely resembles the function declaration. The XQuery function

```xml
declare function local:char-at(
    $str as xs:string,
    $pos as xs:integer
) as xs:string {  
    fn:substring($str, $pos, 1)
};
```

for example has the type function(xs:string, xs:integer) as xs:string. It isn't possible to specify only the argument and not the result type or the other way round. A good place-holder to use when no restriction is wanted is item()*, as it matches any XQuery value.

Function types can also be nested. As an example we take local:on-sequences, which takes a function defined on single items and makes it work on sequences as well:

```xml
declare function local:on-sequences(
    $fun as function(item()) as item()*
) as function(item()*) as item()* {  
    fn:for-each($fun, ?)
};
```

We'll see later how fn:for-each(...) works. The type of local:on-sequences(...) on the other hand is easily constructed, if a bit long:

```xml
function(function(item()) as item()*) as function(item()) as item()*.  
```
Higher-Order Functions

A higher-order function is a function that takes other functions as arguments and/or returns them as results. \texttt{fn:for-each} and \texttt{local:on-sequences} from the last chapter are nice examples.

With the help of higher-order functions, one can extract common patterns of behaviour and abstract them into a library function.

Higher-Order Functions on Sequences

Some usage patterns on sequences are so common that the higher-order functions describing them are in the XQuery standard libraries. They are listed here, together with their possible XQuery implementation and some motivating examples.

\textbf{fn:for-each}

<table>
<thead>
<tr>
<th>Signatures</th>
<th>\texttt{fn:for-each($seq as item()<em>, $fun as function(item()) as item()</em>)} as item()*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Applies the function \texttt{$fun} to every element of the sequence \texttt{$seq} and returns all of the results as a sequence.</td>
</tr>
<tr>
<td>Examples</td>
<td>• Squaring all numbers from 1 to 10:</td>
</tr>
<tr>
<td></td>
<td>\texttt{fn:for-each(1 to 10, math:pow(?x, 2))}</td>
</tr>
<tr>
<td></td>
<td>\textit{Result:} 1 4 9 16 25 36 49 64 81 100</td>
</tr>
<tr>
<td></td>
<td>• Applying a list of functions to a string:</td>
</tr>
<tr>
<td></td>
<td>let $fs := (</td>
</tr>
<tr>
<td></td>
<td>\hspace{1em} \texttt{fn:upper-case#1},</td>
</tr>
<tr>
<td></td>
<td>\hspace{1em} \texttt{fn:substring(?x, 4)},</td>
</tr>
<tr>
<td></td>
<td>\hspace{1em} \texttt{fn:string-length#1}</td>
</tr>
<tr>
<td></td>
<td>)</td>
</tr>
<tr>
<td></td>
<td>\hspace{1em} return \texttt{fn:for-each($fs, function($f) { $f('foobar') }})</td>
</tr>
<tr>
<td></td>
<td>\textit{Result:} FOOBAR bar 6</td>
</tr>
</tbody>
</table>

\textbf{XQuery 1.0}

\begin{verbatim}
declare function local:for-each(  
    $seq as item()*,  
    $fun as function(item()) as item()*  
) as item()* {  
    for $s in $seq  
    return $fun($s)  
};
\end{verbatim}

\textbf{fn:filter}

<table>
<thead>
<tr>
<th>Signatures</th>
<th>\texttt{fn:filter($seq as item()<em>, $pred as function(item()) as xs:boolean)} as item()</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Applies the boolean predicate \texttt{$pred} to all elements of the sequence \texttt{$seq}, returning those for which it returns \texttt{true}().</td>
</tr>
<tr>
<td>Examples</td>
<td>• All even integers until 10:</td>
</tr>
<tr>
<td></td>
<td>\texttt{fn:filter(1 to 10, function($x) { $x mod 2 eq 0 })}</td>
</tr>
<tr>
<td></td>
<td>\textit{Result:} 2 4 6 8 10</td>
</tr>
</tbody>
</table>
• Strings that start with an upper-case letter:

```xquery
let $first-upper := function($str) {
  let $first := fn:substring($str, 1, 1)
  return $first eq fn:upper-case($first)
}
return fn:filter(('FooBar', 'foo', 'BAR'), $first-upper)
```

Result: FooBar BAR

• Inefficient prime number generator:

```xquery
let $is-prime := function($x) {
  $x gt 1 and (every $y in 2 to ($x - 1) satisfies $x mod $y ne 0)
} return filter(1 to 20, $is-prime)
```

Result: 2 3 5 7 11 13 17 19

Note
fn:filter can be easily implemented with fn:for-each:

```xquery
declare function local:filter($seq, $pred) {
  for-each($seq, function($x) {
    if($pred($x)) then $x else ()
  })
}
```

XQuery 1.0

```xquery
declare function local:filter($seq as item()*,$pred as function(item()) as xs:boolean) as item()* {
  $seq[$pred(.)]
}
```

fn:for-each-pair

Signatures
fn:for-each-pair($seq1 as item()*, $seq2 as item()*, $fun as function(item(), item()) as item()*) as item()*

Summary
zips the elements from the two sequences $seq1 and $seq2 together with the function $f. It stops after the shorter sequence ends.

Examples
• Adding one to the numbers at odd positions:

```xquery
fn:for-each-pair(
  fn:for-each(1 to 10, function($x) { $x mod 2 }),(1, 1, 1, 1, 1),
  function($a, $b) { $a + $b })
```

Result: 2 1 2 1 2

• Line numbering:

```xquery
let $number-lines := function($str) {
  fn:string-join(
    fn:for-each-pair(
      fn:for-each-
Higher-Order Functions

Folds

A fold, also called reduce or accumulate in other languages, is a very basic higher-order function on sequences. It starts from a seed value and incrementally builds up a result, consuming one element from the sequence at a time and combining it with the aggregate of a user-defined function.

Folds are one solution to the problem of not having state in functional programs. Solving a problem in imperative programming languages often means repeatedly updating the value of variables, which isn't allowed in functional languages.

Calculating the product of a sequence of integers for example is easy in Java:

```java
public int product(int[] seq) {
    int result = 1;
    for(int i : seq) {
        result = result * i;
    }
    return result;
}
```
Nice and efficient implementations using folds will be given below.

The linear folds on sequences come in two flavours. They differ in the direction in which they traverse the sequence:

**fn:fold-left**

**Signatures**

\[
\text{fn:fold-left}($\text{seq}$ as item()*$, \text{seed}$ as item()*, $\text{fun}$ as function(item()*, item()) as item()*) as item()*
\]

**Summary**

The left fold traverses the sequence from the left. The query \( \text{fn:fold-left}(1 \text{ to } 5, 0, \text{$f$}) \) for example would be evaluated as:

\[
\text{$f$}($\text{$f$}($\text{$f$}($\text{$f$}(0, 1), 2), 3), 4), 5$)
\]

**Examples**

- Product of a sequence of integers:

  \[
  \text{fn:fold-left}(1 \text{ to } 5, 1, \\
  \text{function($result$, $curr$) { $result \times $curr }} \\
  )
  \]

  **Result:** 120

- Illustrating the evaluation order:

  \[
  \text{fn:fold-left}(1 \text{ to } 5, '$\text{seed}$', \\
  \text{concat('$f$', ?, ', ', ?, ', ', '?')})
  \]

  **Result:** $\text{$f$}($\text{$f$}($\text{$f$}($\text{$f$}($\text{$f$}($\text{$seed$}, 1), 2), 3), 4), 5$}$

- Building a decimal number from digits:

  \[
  \text{let $from-digits := fold-left(?, 0,} \\
  \text{function($n$, $d$) { 10 * $n + $d }} \\
  ) \text{ return } ( \\
  $from-digits(1 \text{ to } 5), \\
  $from-digits((4, 2)) \\
  )
  \]

  **Result:** 12345 42

**XQuery 1.0**

As folds are more general than FLWOR expressions, the implementation isn't as concise as the former ones:

```xquery
declare function local:fold-left( 
  $seq$ as item()*, 
  $seed$ as item()*, 
  $fun$ as function(item()*, item()) as item()*
) as item() { 
  if(empty($seq)) then $seed 
  else local:fold-left( 
    fn:tail($seq), 
    $fun($seed, fn:head($seq)), 
    $fun 
  )
};
```


Higher-Order Functions

**fn:fold-right**

**Signatures**

\[
\text{fn:fold-right($seq$ as item()*}, \ $\text{seed$ as item()$}, \ $\text{fun$ as function(item()), item()$)} \text{as item()$}
\]

**Summary**

The **right fold** \( \text{fn:fold-right($seq$, $seed$, $fun$) } \) traverses the sequence from the right. The query \( \text{fn:fold-right(1 to 5, 0, $f$) } \) for example would be evaluated as:

\[
\text{$f(1, \ $f(2, \ $f(3, \ $f(4, \ $f(5, 0))))$)
}\]

**Examples**

- Product of a sequence of integers:

\[
\begin{align*}
\text{fn:fold-right(1 to 5, 1, } \& \\
\quad \text{function($curr$, $result$) } \{ \text{$result$ * $curr$ } \}
\end{align*}
\]

*Result:* 120

- Illustrating the evaluation order:

\[
\begin{align*}
\text{fn:fold-right(1 to 5, '$seed', } \& \\
\quad \text{concat('$f(', ?, ', ', ?, ')')}$
\end{align*}
\]

*Result:* \( \text{$f(1, \ $f(2, \ $f(3, \ $f(4, \ $f(5, $seed))))$) } \)

- Reversing a sequence of items:

\[
\begin{align*}
\text{let $reverse$ := fn:fold-right(?}, \ (), } \& \\
\quad \text{function($item$, $rev$) } \{ \text{$rev$, $item$ } \}
\end{align*}
\]

\[
\begin{align*}
\text{return $reverse(1 to 10)$}
\end{align*}
\]

*Result:* 10 9 8 7 6 5 4 3 2 1

**XQuery 1.0**

\[
\begin{align*}
\text{declare function local:fold-right(} \\
\quad \text{$seq$ as item()*}, \ $\text{seed$ as item()$}, \ $\text{fun$ as function(item()), item()$)} \text{as item()$}
\end{align*}
\]

\[
\begin{align*}
\text{if(empty($seq$)) then $seed$}
\end{align*}
\]

\[
\begin{align*}
\text{else $fun$}
\end{align*}
\]

\[
\begin{align*}
\text{fn:head($seq$),}
\end{align*}
\]

\[
\begin{align*}
\text{local:fold-right(tail($seq$), $seed$, $fun$)}
\end{align*}
\]

Note that the order of the arguments of $fun$ are inverted compared to that in \( \text{fn:fold-left(...)} \).
Chapter 22. XQuery 3.1

Read this entry online in the BaseX Wiki.

This article is part of the XQuery Portal. It summarizes the new features of the XQuery 3.1 and XQuery 3.1 Functions and Operators specifications that are already supported by BaseX.

Please note that not all of the features that are listed on this page are already found in the official specification documents today. You can check out the W3 Bugzilla Tracker to find out more about the most recently added features.

Maps

A map is a function that associates a set of keys with values, resulting in a collection of key/value pairs. Each key/value pair in a map is called an entry. A key is an arbitrary atomic value, and the associated value is an arbitrary sequence. Within a map, no two entries have the same key, when compared using the eq operator. It is not necessary that all the keys should be mutually comparable (for example, they can include a mixture of integers and strings).

Maps can be constructed as follows:

```xquery
map { },                                    (: empty map :) 
map { 'key': true(), 1984: (<a/>, <b/>) },  (: map with two entries :) 
map:merge( 
  for $i in 1 to 10 
  return map { $i: 'value' || $i } 
)  (: map with ten entries :) 
```

The function corresponding to the map has the signature function($key as xs:anyAtomicType) as item(). The expression $map($key) returns the associated value; the function call map:get($map, $key) is equivalent. For example, if $books-by-isbn is a map whose keys are ISBNs and whose associated values are book elements, then the expression $books-by-isbn("0470192747") returns the book element with the given ISBN. The fact that a map is a function item allows it to be passed as an argument to higher-order functions that expect a function item as one of their arguments. As an example, the following query uses the higher-order function fn:map($f, $seq) to extract all bound values from a map:

```xquery
let $map := map { 'foo': 42, 'bar': 'baz', 123: 456 } 
return fn:for-each(map:keys($map), $map)
```

This returns some permutation of (42, 'baz', 456).

Because a map is a function item, functions that apply to functions also apply to maps. A map is an anonymous function, so fn:function-name returns the empty sequence; fn:function-arity always returns 1.

Like all other values, maps are immutable. For example, the map:remove function creates a new map by removing an entry from an existing map, but the existing map is not changed by the operation. Like sequences, maps have no identity. It is meaningful to compare the contents of two maps, but there is no way of asking whether they are "the same map": two maps with the same content are indistinguishable.

Maps may be compared using the fn:deep-equal function. The Map Module describes the available set of map functions.

Arrays

An array is a function that associates a set of positions, represented as positive integer keys, with values. The first position in an array is associated with the integer 1. The values of an array are called its members. In the type
hierarchy, array has a distinct type, which is derived from function. In BaseX, arrays (as well as sequences) are based on an efficient Finger Tree implementation.

Arrays can be constructed in two ways. With the square bracket notation, the comma serves as delimiter:

```
[ ],             (: empty array :)
[ (1, 2) ],     (: array with single member :)
[ 1 to 2, 3 ]   (: array with two members; same as: [ (1, 2), 3 ] :)
```

With the `array` keyword and curly brackets, the inner expression is evaluated as usual, and the resulting values will be the members of the array:

```
array { },           (: empty array;              same as: array { () } :)
array { (1, 2) },    (: array with two members;   same as: array { 1, 2 } :)
array { 1 to 2, 3 }  (: array with three members; same as: array { 1, 2, 3 } :)
```

The function corresponding to the array has the signature `function($index as xs:integer) as item(*)`. The expression `$array($index)` returns an addressed member of the array. The following query returns the five array members 48 49 50 51 52 as result:

```
let $array := array { 48 to 52 }
for $i in 1 to array:size($array)
return $array($i)
```

Like all other values, arrays are immutable. For example, the `array:reverse` function creates a new array containing a re-ordering of the members of an existing array, but the existing array is not changed by the operation. Like sequences, arrays have no identity. It is meaningful to compare the contents of two arrays, but there is no way of asking whether they are "the same array": two arrays with the same content are indistinguishable.

**Atomization**

If an array is *atomized*, all of its members will be atomized. As a result, an atomized item may now result in more than one item. Some examples:

```
fn:data([1 to 2])        (: returns the sequence 1, 2 :)  
[ 'a', 'b', 'c' ] = 'b'  (: returns true :)  
<a>[ [ 1, 2 ] ]</a>      (: returns <a>1 2</a> :)  
array { 1 to 2 } + 3     (: error: the left operand returns two items :)
```

Atomization also applies to function arguments. The following query returns 5, because the array will be atomized to a sequence of 5 integers:

```
let $f := function($x as xs:integer*) { count($x) }
return $f([1 to 5])
```

However, the next query returns 1, because the array is already of the general type `item(*)`, and no atomization will take place:

```
let $f := function($x as item(*)*) { count($x) }
return $f([1 to 5])
```

Arrays can be compared with the `fn:deep-equal` function. The Array Module describes the available set of array functions.

**Lookup Operator**

The lookup operator provides some syntactic sugar to access values of maps or array members. It is introduced by the question mark (?) and followed by a specifier. The specifier can be:
1. A wildcard *,
2. The name of the key,
3. The integer offset, or
4. Any other parenthesized expression.

The following example demonstrates the four alternatives:

```
let $map := map { 'R': 'red', 'G': 'green', 'B': 'blue' }
return ($map?*          (: 1. returns all values; same as: map:keys($map) ! $map(.) :),
       $map?'R'          (: 2. returns the value associated with the key 'R'; same as: $map('R') :),
       $map?('G','B')  (: 3. returns the values associated with the key 'G' and 'B' :)
)
```

```
let $array := [ 'one', 'two', 'three' ]
return ($array?*         (: 1. returns all values; same as: (1 to array:size($array)) ! $array(.) :),
       $array?1         (: 2. returns the first value; same as: $array(1) :),
       $array?(2 to 3)  (: 3. returns the second and third values; same as: (1 to 2) ! $array(.) :)
)
```

The lookup operator can also be used without left operand. In this case, the context item will be used as input. This query returns Akureyri:

```
let $maps := (map { 'name': 'Guðrún', 'city': 'Reykjavík' },
               map { 'name': 'Hildur', 'city': 'Akureyri' })
return $maps[?name = 'Hildur'] ?city
```

### Arrow Operator

The arrow operator applies a function to a value. The value is used as the first argument to the function. It is introduced with the characters =>, and it is followed by the function to be called. If $v$ is a value and $f()$ is a function, then $v=>$f() is equivalent to $f($v$)$, and $v=>$f($j$) is equivalent to $f($v$, $j$). This is further illustrated by an example:

```
(: Returns 3 :) count(('A', 'B', 'C')),
   ('A', 'B', 'C') => count(),
   ('A', 'B', 'C') => (function( $sequence) { count( $sequence)})(),

(: Returns W-E-L-C-O-M-E :) string-join(tokenize(upper-case('w e l c o m e')), '-'),
   'w e l c o m e' => upper-case() => tokenize() => string-join('-'),

(: Returns xfmdpnnf :) codepoints-to-string( for $i in string-to-codepoints('welcome')
                        return $i + 1 ),
   (for $i in 'welcome' => string-to-codepoints() return $i + 1) => codepoints-to-string()
```

The syntax makes nested function calls more readable, as it is easy to see if parentheses are balanced.
String Constructor

The string constructor has been inspired by here document literals of the Unix shell and script languages. It allows you to generate strings that contain various characters that would otherwise be interpreted as XQuery delimiters.

The string constructors syntax uses two backticks and a square bracket for opening and closing a string:

```
(: Returns "This is a 'new' & 'flexible' syntax." :)```

```
``["This is a 'new' & 'flexible' syntax."]```

XQuery expressions can be embedded via backticks and a curly bracket. The evaluated results will be separated with spaces, and all strings will eventually be concatenated:

```
(: Returns »Count 1 2 3, and I will be there.« :)```

```
let $c := 1 to 3
return `[
"Count `{ $c }`, and I will be there.«`]
```

Serialization

Two Serialization methods have been added to the Serialization spec:

Adaptive Serialization

The adaptive serialization provides an intuitive textual representation for all XDM types, including maps and arrays, functions, attributes, and namespaces. All items will be separated by the value of the item-separator parameter, which by default is a newline character. It is utilized by the functions prof:dump and fn:trace.

Example:

```
declare option output:method 'adaptive';
<element id='id0'/><id,
xs:token("abc"),
map { 'key': 'value' },
true#0
```

Result:

```
id="id0"
xs:token("abc"),
map {
  "key": "value"
}
fn:true#0
```

JSON Serialization

The new json serialization output method can be used to serialize XQuery maps, arrays, atomic values and empty sequences as JSON.

The json output method has been introduced in BaseX quite a while ago. The implementation of this method now complies with the standard serialization rules and, at the same time, preserves the existing semantics:

• If an XML node of type element(json) is found, it will be serialized following the serialization rules of the JSON Module.

• Any other node or atomic value, map, array, or empty sequence will be serialized according to the rules in the specification.
The following two queries will both return the JSON snippet `{ "key": "value" }`

```
declare option output:method 'json';
map { "key": "value" }
```

```
declare option output:method 'json';
<json type='object'>
  <key>value</key>
</json>
```

## Functions

The following functions of the XQuery 3.1 Functions and Operators Working Draft have been added. Please be aware that the functions are still subject to change:

### Map Functions

The following map functions are now available:

- `map:merge`
- `map:size`
- `map:keys`
- `map:contains`
- `map:get`
- `map:entry`
- `map:put`
- `map:remove`
- `map:for-each`

Please check out the [Map Module](#) for more details.

### Array Functions

The following array functions are now available:

- `array:size`
- `array:append`
- `array:subarray`
- `array:remove`
- `array:insert-before`
- `array:head`
- `array:tail`
- `array:reverse`
- `array:join`
- `array:flatten`
- `array:for-each`
- `array:filter`
- `array:fold-left`
- `array:fold-right`
- `array:for-each-pair`

Please check out the [Array Module](#) for more details.

### JSON Functions

XQuery now provides native support for JSON objects. Strings and resources can be parsed to XQuery items and, as shown above, serialized back to their original form.

#### `fn:parse-json`

**Signatures**

- `fn:parse-json($input as xs:string) as item()?`
- `fn:parse-json($input as xs:string, $options as map(*)) as item()?`

Parses the supplied string as JSON text and returns its item representation. The result may be a map, an array, a string, a double, a boolean, or an empty sequence. The allowed options can be looked up in the specification.

```
parse-json('{ "name": "john" }')   (: yields { "name": "json" } :),
parse-json('[ 1, 2, 4, 8, 16 ]')    (: yields [ 1, 2, 4, 8, 16 ] :)
```

#### `fn:json-doc`

**Signatures**

- `fn:json-doc($uri as xs:string) as item()?`
XQuery 3.1

- **fn:json-doc($uri as xs:string, $options as map(*)) as item()?**

Retrieves the text from the specified URI, parses the supplied string as JSON text and returns its item representation (see fn:parse-json for more details).

```
json-doc("http://ip.jsontest.com/")('ip')  (: returns your IP address :)
```

**fn:json-to-xml**

Signatures

- **fn:json-to-xml($string as xs:string?) as node()?**

Converts a JSON string to an XML node representation. The allowed options can be looked up in the specification.

```
json-to-xml('{ "message": "world" }')
(: result:
<map xmlns="http://www.w3.org/2005/xpath-functions">
   <string key="message">world</string>
</map> :)
```

**fn:xml-to-json**

Signatures

- **fn:xml-to-json($node as node()?) as xs:string?**

Converts an XML node, whose format conforms to the results created by fn:json-to-xml, to a JSON string representation. The allowed options can be looked up in the specification.

```
(: returns "JSON" ;)
xm-to-json(<string xmlns="http://www.w3.org/2005/xpath-functions">JSON</string>)
```

**fn:sort**

Signatures

- **fn:sort($input as item()* as item()***

- **fn:sort($input as item()* as item()* as item()*)**

Returns a new sequence with sorted $input items. If a sort $key function is given, it will be applied on all items. The items of the resulting values will be sorted using the semantics of the lt expression.

```
sort(reverse(1 to 3))                (: yields 1, 2, 3 :),
reverse(sort(1 to 3))                (: returns the sorted order in descending order ;),
sort((3, -2, 1), abs#1)             (: yields 1, -2, 3 :),
sort((1,2,3), function($x) { -$x }) (: yields 3, 2, 1 :),
sort((1, 'a'))                       (: yields an error, as strings and integers cannot be compared ;)
```

**fn:contains-token**

Signatures

- **fn:contains-token($input as xs:string*, $token as string) as xs:boolean**
• \texttt{fn:contains-token($input as xs:string*, $token as string, $collation as xs:string) as xs:boolean}

The supplied strings will be tokenized at whitespace boundaries. The function returns \texttt{true} if one of the strings equals the supplied token, possibly under the rules of a supplied collation:

\begin{verbatim}
contains-token(("a", "b c", "d"), "c")  (: yields \texttt{true} :)  
<xml class="one two"></contains-token(@class, 'one')}  (: yields \texttt{true} :)
\end{verbatim}

**\texttt{fn:parse-ietf-date}**

Signature

• \texttt{fn:parse-ietf-date($input as xs:string?) as xs:string?}

Parses a string in the IETF format (which is widely used on the Internet) and returns a \texttt{xs:dateTime} item:

\begin{verbatim}
fn:parse-ietf-date('28-Feb-1984 07:07:07')"  (: yields 1984-02-28T07:07:07 :)  ,
fn:parse-ietf-date('Wed, 01 Jun 2001 23:45:54 +02:00')"  (: yields 2001-06-01T23:45:54+02:00 :)  
\end{verbatim}

**\texttt{fn:apply}**

Signatures

• \texttt{fn:apply($function as function(*), $array as array(*)) as item()}*

A supplied function is invoked with the arguments supplied by an array. The arity of the function must be the same as the size of the array.

Example:

\begin{verbatim}
fn:apply(concat#5, array { 1 to 5 })  (: 12345 :)  ,
fn:apply(function($a) { sum($a) }, [ 1 to 5 ])  (: 15 :)  ,
fn:apply(count#1, [ 1,2 ])  (: error (the array has two members) :)  
\end{verbatim}

**\texttt{fn:random-number-generator}**

Signatures

• \texttt{fn:random-number-generator() as map(xs:string, item())}

• \texttt{fn:random-number-generator($seed as xs:anyAtomicType) as map(xs:string, item())}

Creates a random number generator, using an optional seed. The returned map contains three entries:

• \texttt{number} is a random double between 0 and 1

• \texttt{next} is a function that returns another random number generator

• \texttt{permute} is a function that returns a random permutation of its argument

The returned random generator is \textit{deterministic}: If the function is called twice with the same arguments and in the same execution scope, it will always return the same result.

Example:
let $rng := fn:random-number-generator()
let $number := $rng('number')               (: returns a random number :) 
let $next-rng := $rng('next')()             (: returns a new generator :) 
let $next-number := $next-rng('number')     (: returns another random number :) 
let $permutation := $rng('permute')(1 to 5) (: returns a random permutation of  
(1,2,3,4,5) :) 
return ($number, $next-number, $permutation)

fn:format-number

The function has been extended to support scientific notation:

format-number(1984.42, '00.0e0')  (: yields 19.8e2 :)

fn:tokenize

If no separator is specified as second argument, a string will be tokenized at whitespace boundaries:

fn:tokenize(" a b c d") (: yields "a", "b", "c", "d" :)

fn:trace

The second argument can now be omitted:

fn:trace(<xml/>, "Node: ")/node()   (: yields the debugging output "Node: <xml/  "> :),
fn:trace(<xml/>)/node()            (: returns the debugging output "<xml/>" :)

fn:string-join

The type of the first argument is now xs:anyAtomicType*, and all items will be implicitly cast to strings:

fn:string-join(1 to 3)  (: yields the string "123" :)

fn:default-language

Returns the default language used for formatting numbers and dates. BaseX always returns en.

Binary Data

Items of type xs:hexBinary and xs:base64Binary can now be compared against each other. The following queries all yield true:

xs:hexBinary('') < xs:hexBinary('bb'),
xs:hexBinary('aa') < xs:hexBinary('bb'),
max((xs:hexBinary('aa'), xs:hexBinary('bb'))) = xs:hexBinary('bb')

Collations

XQuery 3.1 provides a new default collation, which allows for a case-insensitive comparison of ASCII characters (A-Z = a-z). This query returns true:

declare default collation 'http://www.w3.org/2005/xpath-functions/collation/html-ascii-case-insensitive';
'HTML' = 'html'
If the ICU Library is downloaded and added to the classpath, the full Unicode Collation Algorithm features get available in BaseX:

```xml
(: returns 0 (both strings are compared as equal) :)
compare('a-b', 'ab', 'http://www.w3.org/2013/collation/UCA?alternate=shifted')
```

### Enclosed Expressions

*Enclosed expression* is the syntactical term for the expressions that are specified inside a function body, try/catch clauses, node constructors and some other expressions. In the following example expressions, it’s the empty sequence:

```xml
declare function local:x() { () };
try { () } catch * { () },
element x { () },
text { () }
```

With XQuery 3.1, the expression can be omitted. The following query is equivalent to the upper one:

```xml
declare function local:x() { };
try { } catch * { },
element x { }
text { }
```

### Pending Features

The following functions have not been implemented yet:

- fn:collation-key, fn:load-xquery-module, fn:transform

### Changelog

**Version 8.4**

- Added: String Constructors, fn:default-language, Enclosed Expressions
- Updated: Adaptive Serialization, fn:string-join

**Version 8.2**


**Version 8.1**

- Updated: arrays are now based on a Finger Tree implementation.

Introduced with Version 8.0.
Chapter 23. XQuery Extensions

Read this entry online in the BaseX Wiki.

This article is part of the XQuery Portal. It lists extensions and optimizations that are specific to the BaseX XQuery processor.

Option Declarations

Local database options can be set in the prolog of an XQuery main module. In the option declaration, options need to be bound to the Database Module namespace. All values will be reset after the evaluation of a query:

```
declare option db:chop 'false';
doc('doc.xml')
```

Pragmas

Local database options can be assigned locally via pragmas:

```
(# db:chop false #) { doc('doc.xml') }
```

Various optimizations can be disabled by marking an expression as non-deterministic:

```
count( (# basex:non-deterministic #) { 1 to 10 })
```

Annotations

The following implementation-defined annotations are available:

- `%basex:inline([limit])` enforces the inlining of a function. The inlining limit can be specified as argument. Inlining can be disabled by specifying 0.

Example:

```
declare %basex:inline(0) function local:e() { error() ];
local:e()
```

Result:

```
Stopped at query.xq, 1/53:
[FOER0000] Halted on error().

Stack Trace:
- query.xq, 2/9
```

In the next example, function inlining was disabled globally by assigning 0 to the INLINELIMIT option. However, annotation is enforced to a single function:

Example:

```
declare option db:inlinelimit '0';
declare %basex:inline function local:id($x) { $x ];
local:id(123)
```

The query will be optimized to 123.
XQuery Extensions

• %basex:lazy enforces the lazy evaluation of a global variable. Example:

Example:

```
declare %basex:lazy variable $january := doc('does-not-exist');
if(month-from-date(current-date()) == 1) then $january else ()
```

The annotation ensures that an error will only be thrown if the condition yields true. Without the annotation, the error will always be thrown, because the referenced document is not found.

**Serialization**

• basex is used as the default serialization method: nodes are serialized as XML, atomic values are serialized as string, and items of binary type are output in their native byte representation. Function items (including maps and arrays) are output just like with the adaptive method.

• csv allows you to output XML nodes as CSV data (see the CSV Module for more details).

For more information and some additional BaseX-specific parameters, see the article on Serialization.

**Non-determinism**

In XQuery, deterministic functions are “guaranteed to produce identical results from repeated calls within a single execution scope if the explicit and implicit arguments are identical”. In BaseX, many extension functions are non-deterministic or side-effecting. If an expression is internally flagged as non-deterministic, various optimizations that might change their execution order will not be applied.

```
(: QUERY A... :) 
let $n := 456
for $i in 1 to 2
return $n

(: ...will be optimized to :) 
for $i in 1 to 2
return 456

(: QUERY B will not be rewritten :) 
let $n := random:integer()
for $i in 1 to 2
return $n
```

In some cases, functions may contain non-deterministic code, but the query compiler may not be able to detect this statically. See the following example:

```
for $read in (file:read-text#1, file:read-binary#1)
let $ignored := non-deterministic $read('input.file')
return ()
```

Two non-deterministic functions will be bound to $read, and the result of the function call will be bound to $ignored. As the variable is not referenced in the subsequent code, the let clause would usually be discarded by the compiler. In the given query, however, execution will be enforced because of the BaseX-specific non-deterministic keyword.

**Miscellaneous**

Various other extensions are described in the articles on XQuery Full Text and XQuery Update.
Chapter 24. Module Library

Read this entry online in the BaseX Wiki.

This article is part of the XQuery Portal.

In addition to the standard XQuery Functions, BaseX offers further function modules, which are listed in the following table. The module namespaces are statically bound, which means that they need not (but may) be explicitly declared in the query prolog.

<table>
<thead>
<tr>
<th>Module</th>
<th>Description</th>
<th>Prefix</th>
<th>Namespace URI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Admin</td>
<td>Functions restricted to admin users.</td>
<td>admin</td>
<td><a href="http://basex.org/modules/admin">http://basex.org/modules/admin</a></td>
</tr>
<tr>
<td>Archive</td>
<td>Creating and processing ZIP archives.</td>
<td>archive</td>
<td><a href="http://basex.org/modules/archive">http://basex.org/modules/archive</a></td>
</tr>
<tr>
<td>Array</td>
<td>Functions for handling array arrays.</td>
<td>array</td>
<td><a href="http://www.w3.org/2005/xpath-functions/array">http://www.w3.org/2005/xpath-functions/array</a></td>
</tr>
<tr>
<td>Binary</td>
<td>Processing binary data.</td>
<td>bin</td>
<td><a href="http://expath.org/ns/binary">http://expath.org/ns/binary</a></td>
</tr>
<tr>
<td>Client</td>
<td>Executing commands and queries on remote BaseX servers.</td>
<td>client</td>
<td><a href="http://basex.org/modules/client">http://basex.org/modules/client</a></td>
</tr>
<tr>
<td>Conversion</td>
<td>Converting data (binary, numeric) to other formats.</td>
<td>convert</td>
<td><a href="http://basex.org/modules/convert">http://basex.org/modules/convert</a></td>
</tr>
<tr>
<td>Cryptography</td>
<td>Cryptographic functions, based on the EXPath Cryptographic module.</td>
<td>crypto</td>
<td><a href="http://expath.org/ns/crypto">http://expath.org/ns/crypto</a></td>
</tr>
<tr>
<td>CSV</td>
<td>Functions for processing CSV input.</td>
<td>csv</td>
<td><a href="http://basex.org/modules/csv">http://basex.org/modules/csv</a></td>
</tr>
<tr>
<td>Database</td>
<td>Functions for accessing and updating databases.</td>
<td>db</td>
<td><a href="http://basex.org/modules/db">http://basex.org/modules/db</a></td>
</tr>
<tr>
<td>Fetch</td>
<td>Functions for fetching resources identified by URIs.</td>
<td>fetch</td>
<td><a href="http://basex.org/modules/fetch">http://basex.org/modules/fetch</a></td>
</tr>
<tr>
<td>File</td>
<td>File handling, based on the latest draft of the EXPath File module.</td>
<td>file</td>
<td><a href="http://expath.org/ns/file">http://expath.org/ns/file</a></td>
</tr>
<tr>
<td>Full-Text</td>
<td>Functions for performing full-text operations.</td>
<td>ft</td>
<td><a href="http://basex.org/modules/ft">http://basex.org/modules/ft</a></td>
</tr>
<tr>
<td>Hashing</td>
<td>Cryptographic hash functions.</td>
<td>hash</td>
<td><a href="http://basex.org/modules/hash">http://basex.org/modules/hash</a></td>
</tr>
<tr>
<td>Higher-Order</td>
<td>Additional higher-order functions that are not in the standard libraries.</td>
<td>hof</td>
<td><a href="http://basex.org/modules/hof">http://basex.org/modules/hof</a></td>
</tr>
<tr>
<td>HTML</td>
<td>Functions for converting HTML input to XML documents.</td>
<td>html</td>
<td><a href="http://basex.org/modules/html">http://basex.org/modules/html</a></td>
</tr>
<tr>
<td>Module Library</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HTTP</strong></td>
<td>Sending HTTP requests, http based on the EXPath HTTP module.</td>
<td><a href="http://expat.org/ns/http-client">http://expat.org/ns/http-client</a></td>
<td></td>
</tr>
<tr>
<td><strong>Index</strong></td>
<td>Functions for requesting index details on database indexes.</td>
<td><a href="http://basex.org/modules/index">http://basex.org/modules/index</a></td>
<td></td>
</tr>
<tr>
<td><strong>Inspection</strong></td>
<td>Functions for extracting inspect internal module information.</td>
<td><a href="http://basex.org/modules/inspect">http://basex.org/modules/inspect</a></td>
<td></td>
</tr>
<tr>
<td><strong>JSON</strong></td>
<td>Parsing and serializing json JSON documents.</td>
<td><a href="http://basex.org/modules/json">http://basex.org/modules/json</a></td>
<td></td>
</tr>
<tr>
<td><strong>Map</strong></td>
<td>Functions for handling map maps (key/value pairs).</td>
<td><a href="http://www.w3.org/2005/xpath-functions/map">http://www.w3.org/2005/xpath-functions/map</a></td>
<td></td>
</tr>
<tr>
<td><strong>Math</strong></td>
<td>Mathematical operations, math extending the W3C Working Draft.</td>
<td><a href="http://www.w3.org/2005/xpath-functions/math">http://www.w3.org/2005/xpath-functions/math</a></td>
<td></td>
</tr>
<tr>
<td><strong>Output</strong></td>
<td>Functions for simplifying output formatted output.</td>
<td><a href="http://basex.org/modules/out">http://basex.org/modules/out</a></td>
<td></td>
</tr>
<tr>
<td><strong>Process</strong></td>
<td>Executing system commands from XQuery.</td>
<td><a href="http://basex.org/modules/proc">http://basex.org/modules/proc</a></td>
<td></td>
</tr>
<tr>
<td><strong>Profiling</strong></td>
<td>Functions for profiling code snippets.</td>
<td><a href="http://basex.org/modules/prof">http://basex.org/modules/prof</a></td>
<td></td>
</tr>
<tr>
<td><strong>Random</strong></td>
<td>Functions for creating random random numbers.</td>
<td><a href="http://basex.org/modules/random">http://basex.org/modules/random</a></td>
<td></td>
</tr>
<tr>
<td><strong>Repository</strong></td>
<td>Installing, deleting and listing packages.</td>
<td><a href="http://basex.org/modules/repo">http://basex.org/modules/repo</a></td>
<td></td>
</tr>
<tr>
<td><strong>SQL</strong></td>
<td>JDBC bridge to access sql relational databases.</td>
<td><a href="http://basex.org/modules/sql">http://basex.org/modules/sql</a></td>
<td></td>
</tr>
<tr>
<td><strong>Streaming</strong></td>
<td>Functions for handling stream streamable items.</td>
<td><a href="http://basex.org/modules/stream">http://basex.org/modules/stream</a></td>
<td></td>
</tr>
<tr>
<td><strong>Strings</strong></td>
<td>Functions for performing stream string computations.</td>
<td><a href="http://basex.org/modules/strings">http://basex.org/modules/strings</a></td>
<td></td>
</tr>
<tr>
<td><strong>Unit</strong></td>
<td>Unit testing framework.</td>
<td>unit <a href="http://basex.org/modules/unit">http://basex.org/modules/unit</a></td>
<td></td>
</tr>
<tr>
<td><strong>User</strong></td>
<td>Creating and administering database users.</td>
<td><a href="http://basex.org/modules/user">http://basex.org/modules/user</a></td>
<td></td>
</tr>
<tr>
<td><strong>Validation</strong></td>
<td>Validating documents: validate DTDs, XML Schema, RelaxNG.</td>
<td><a href="http://basex.org/modules/validate">http://basex.org/modules/validate</a></td>
<td></td>
</tr>
<tr>
<td><strong>Web</strong></td>
<td>Convenience functions for building web applications.</td>
<td><a href="http://basex.org/modules/web">http://basex.org/modules/web</a></td>
<td></td>
</tr>
<tr>
<td><strong>XQuery</strong></td>
<td>Evaluating new XQuery expressions at runtime.</td>
<td><a href="http://basex.org/modules/xquery">http://basex.org/modules/xquery</a></td>
<td></td>
</tr>
<tr>
<td><strong>XSLT</strong></td>
<td>Stylesheet transformations, xslt based on Java’s and Saxon’s XSLT processor.</td>
<td><a href="http://basex.org/modules/xslt">http://basex.org/modules/xslt</a></td>
<td></td>
</tr>
</tbody>
</table>
ZIP functionality, based on the EXPath ZIP module (soon obsolete).

For the following web application modules, the basex-api package must be included in the classpath and the modules must be imported in the query prolog. This is automatically the case if you use one of the complete distributions (zip, exe, war) of BaseX:

<table>
<thead>
<tr>
<th>Module</th>
<th>Description</th>
<th>Prefix</th>
<th>Namespace URI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geo</td>
<td>Functions for processing geo</td>
<td>geo</td>
<td><a href="http://expath.org/ns/geo">http://expath.org/ns/geo</a></td>
</tr>
<tr>
<td>Request</td>
<td>Server-side functions for request handling HTTP Request data.</td>
<td>request</td>
<td><a href="http://exquery.org/ns/request">http://exquery.org/ns/request</a></td>
</tr>
<tr>
<td>RESTXQ</td>
<td>Helper functions for the RESTXQ API.</td>
<td>rest</td>
<td><a href="http://exquery.org/ns/restxq">http://exquery.org/ns/restxq</a></td>
</tr>
<tr>
<td>Session</td>
<td>Functions for handling server-side HTTP Sessions.</td>
<td>session</td>
<td><a href="http://basex.org/modules/session">http://basex.org/modules/session</a></td>
</tr>
<tr>
<td>Sessions</td>
<td>Functions for managing all server-side HTTP Sessions.</td>
<td>sessions</td>
<td><a href="http://basex.org/modules/sessions">http://basex.org/modules/sessions</a></td>
</tr>
</tbody>
</table>
Chapter 25. Java Bindings

Read this entry online in the BaseX Wiki.

This article is part of the XQuery Portal. It demonstrates different ways to invoke Java code from XQuery, and it presents extensions to access the current query context from Java.

The Java Binding feature is an extensibility mechanism which enables developers to directly access Java variables and execute code from XQuery. Addressed Java code must either be contained in the Java classpath, or it must be located in the Repository.

Identification

Classes

A Java class is identified by a namespace URI. The original URI is rewritten as follows:

1. The URI Rewriting steps are applied to the URI.
2. Slashes in the resulting URI are replaced with dots.
3. The last path segment of the URI is capitalized and rewritten to camel case.

The normalization steps are skipped if the URI is prefixed with java:. See the following examples:

- java:java.lang.String → java.lang.String

Functions and Variables

Java functions and variables can be referenced and evaluated by the existing XQuery function syntax:

- The namespace of the function name identifies the Java class.
- The local part of the name, which is rewritten to camel case, identifies a variable or function of that class.
- The middle dot character (·) is a valid character in XQuery names, but not in Java. It can be used to append exact Java parameter types to the function name. Class types must be referenced by their full path.

<table>
<thead>
<tr>
<th>Type</th>
<th>XQuery</th>
<th>Java</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
<td>Q{java.lang.Integer}MIN_VALUE</td>
<td>Integer.MIN_VALUE</td>
</tr>
<tr>
<td>Function</td>
<td>Q{java.lang.Object}hashCode()</td>
<td>object.hashCode()</td>
</tr>
<tr>
<td>Function with types</td>
<td>Q{java.lang.String}split(java.lang.String:::3)</td>
<td>java.lang.String:::3</td>
</tr>
</tbody>
</table>

As XQuery and Java have different type systems, XQuery arguments are converted to equivalent Java values, and the result of a Java function is converted back to an XQuery value (see Data Types).

If a Java function is not found, XQuery values may need to be cast the target type. For example, if a Java function expects a primitive int value, you will need to convert your XQuery integers to xs:int.

Namespace Declarations

In the following example, Java’s Math class is referenced. When executed, the query returns the cosine of an angle by calling the static method cos(), and the value of π by addressing the static variable via PI():

```java
Math.cos(Math.PI)
```
declare namespace math = "java:java.lang.Math";
math:cos(xs:double(0)), math:PI()

With the Expanded QName notation of XQuery 3.0, the namespace can directly be embedded in the function call:

Q{java:java.lang.Math}cos(xs:double(0))

The constructor of a class can be invoked by calling the virtual function new(). Instance methods can then be called by passing on the resulting Java object as first argument. In the following example, 256 bytes are written to the file output.txt. First, a new FileWriter instance is created, and its write() function is called in the next step:

declare namespace fw = "java.io.FileWriter";
let $file := fw:new('output.txt')
return (for $i in 0 to 255
    return fw:write($file, xs:int($i)),
    fw:close($file))

If the result of a Java call contains invalid XML characters, it will be rejected. The validity check can be disabled by setting the CHECKSTRINGs option to false. The following query writes a file with a single 00-byte, which will then be successfully read via Java functions:

declare namespace br = 'java.io.BufferedReader';
declare namespace fr = 'java.io.FileReader';
declare option db:checkstrings 'false';
file:write-binary('00.bin', xs:hexBinary('00')),
br:new(fr:new('00.bin')) ! (br:readLine(.), br:close(.))

Note that Java code cannot be pre-compiled, and will as such be evaluated slower than optimized XQuery code.

Module Imports

Java code can also be integrated by importing classes as modules. A new instance of the addressed class is created, which can then be accessed in the query body.

The following, side-effecting example returns the number of distinct values added to a hash set (the boolean values returned by set:add() will be swallowed):

import module namespace set = "java.util.HashSet";
prof:void(
  for $s in ("one", "two", "one")
    return set:add($s)
),
set:size()

The advantages of this approach is that imported code is executed faster than instances created at runtime via new(). A drawback is that no arguments can be passed on to the class constructor. As a consequence, the import only works if the class provides a constructor with no arguments.

Context-Awareness

Java classes can be coupled more closely to the BaseX core library. If a class inherits the abstract QueryModule class, the two variables queryContext and staticContext get available, which provide access to the global and static context of a query. Additionally, the default properties of functions can be changed via annotations:

- Java functions can only be executed by users with Admin permissions. You may annotate a function with @Requires(permissions) to also make it accessible to users with less privileges.
• Java code is treated as non-deterministic, as its behavior cannot be predicted by the XQuery processor. You may annotate a function as @Deterministic if you know that it will have no side-effects and will always yield the same result.

• Java code is treated as context-independent. If a function accesses the query context, it should be annotated as @ContextDependent.

• Java code is treated as focus-independent. If a function accesses the current context item, position or size, it should be annotated as @FocusDependent.

The QueryResource interface can be implemented to enforce finalizing operations, such as the closing of opened connections or resources in a module. Its close() method will be called after a query has been fully evaluated.

The following XQuery code invokes two Java methods. The first Java function retrieves information from the static query context, and the second one throws a query exception:

```xquery
import module namespace context = 'org.basex.examples.query.ContextModule';

element user {
    context:user()
},
element to-int {
    try { context:to-int('abc') }
    catch * { 'Error in line', $err:line-number }
}
```

The imported Java class is shown below:

```java
package org.basex.examples.query;

import org.basex.query.*;
import org.basex.query.value.item.*;
import org.basex.util.*;

/**<n*
* This example inherits the {@link QueryModule} class and
* implements the QueryResource interface.
*/
public class ContextModule extends QueryModule implements QueryResource {
    /**<n*
* Returns the name of the logged in user.
* @return user
*/
@Requires(Permission.NONE)
@Deterministic
@ContextDependent
public String user() {
    return queryContext.context.user.name;
}

/**<n*
* Converts the specified string to an integer.
* @param value string representation
* @return integer
* @throws QueryException query exception
*/
@Requires(Permission.NONE)
@Deterministic
public int toInt(final String value) throws QueryException {
    try {
        return Integer.parseInt(value);
    } catch(NumberFormatException ex) {
```
throw new QueryException(ex.getMessage());
}

@Override
public void close() {
   // see description above
}

The result will look as follows:

<user>admin</admin>
<to-int>Error in line 6</to-int>

Please visit the XQuery 3.0 specification if you want to get more insight into function properties.

### Locking

By default, a Java function will be executed in parallel with other code. However, if a Java function performs sensitive write operations, it is advisable to explicitly lock the code. This can be realized via locking annotations:

```java
@Lock(write = { "HEAVYIO" })
public void write() {
   // ...
}

@Lock(read = { "HEAVYIO" })
public void read() {
   // ...
}
```

If an XQuery expression is run which calls the Java `write()` function, every other query that calls `write()` or `read()` needs to wait for the query to be finished. If a query calls the `read()` function, only those queries are queued that call `write()`, because this function is only annotated with a `read` lock. More details on parallel query execution can be found in the article on Transaction Management.

### Data Types

The following table lists the mappings of XQuery and Java types:

<table>
<thead>
<tr>
<th>XQuery Type</th>
<th>Java Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>xs:string</td>
<td><code>String</code>, <code>char</code>, <code>Character</code></td>
</tr>
<tr>
<td>xs:boolean</td>
<td><code>boolean</code>, <code>Boolean</code></td>
</tr>
<tr>
<td>xs:byte</td>
<td><code>byte</code>, <code>Byte</code></td>
</tr>
<tr>
<td>xs:short</td>
<td><code>short</code>, <code>Short</code></td>
</tr>
<tr>
<td>xs:int</td>
<td><code>int</code>, <code>Integer</code></td>
</tr>
<tr>
<td>xs:long</td>
<td><code>long</code>, <code>Long</code></td>
</tr>
<tr>
<td>xs:float</td>
<td><code>float</code>, <code>Float</code></td>
</tr>
<tr>
<td>xs:double</td>
<td><code>double</code>, <code>Double</code></td>
</tr>
<tr>
<td>xs:decimal</td>
<td><code>java.math.BigDecimal</code></td>
</tr>
<tr>
<td>xs:integer</td>
<td><code>java.math.BigInteger</code></td>
</tr>
<tr>
<td>xs:QName</td>
<td><code>javax.xml.namespace.QName</code></td>
</tr>
<tr>
<td>xs:anyURI</td>
<td><code>java.net.URI</code>, <code>java.net.URL</code></td>
</tr>
</tbody>
</table>
empty sequence | null

# Changelog

Version 8.4
- Updated: Rewriting rules

Version 8.0
- Added: QueryResource interface, called after a query has been fully evaluated.

Version 7.8
- Added: Java locking annotations
- Updated: context variable has been split into queryContext and staticContext.

Version 7.2.1
- Added: import of Java modules, context awareness
Chapter 26. Repository

Read this entry online in the BaseX Wiki.

This article is part of the XQuery Portal. It describes how external XQuery modules and Java code can be installed in the XQuery module repository, and how new packages are built and deployed.

Introduction

One of the reasons why languages such as Java or Perl have been so successful is the vast amount of libraries that are available to developers. As XQuery comes with only 150 pre-defined functions, which cannot meet all requirements, there is some need for additional library modules – such as FunctX – that extend the language with new features.

BaseX offers the following mechanisms to make modules accessible to the XQuery processor:

1. The default Packaging mechanism will install single XQuery and Java modules in the repository.
2. The EXPath Packaging system provides a generic mechanism for adding XQuery modules to query processors.

A package is defined as a .xar archive, which encapsulates one or more extension libraries.

Accessing Modules

Library modules can be imported with the import module statement, followed by a freely choosable prefix and the namespace of the target module. The specified location may be absolute or relative; in the latter case, it is resolved against the location (i.e., static base URI) of the calling module. Import module statements must be placed at the beginning of a module:

Main Module HelloUniverse.xq:

```xml
import module namespace m = 'http://basex.org/modules/Hello' at 'HelloWorld.xqm';
m:hello("Universe")
```

Library Module HelloWorld.xqm (in the same directory):

```xml
module namespace m = 'http://basex.org/modules/Hello';
declare function m:hello($world) {
    'Hello ' || $world 
};
```

Repository modules are stored in a directory named BaseXRepo or repo, which resides in your home directory. XQuery modules can be manually copied to the repository directory or installed and deleted via commands.

If a module has been placed in the repository (see below), there is no need to specify the location. The following example calls a function from the FunctX module:

```xml
import module namespace functx = 'http://www.functx.com';
functx:capitalize-first('test')
```

Commands

There are various ways to handle your packages:

- Execute BaseX REPO commands (listed below)
- Call XQuery functions of the Repository Module
• Use the GUI (Options → Packages)

You can even manually add and remove packages in the repository directory; all changes will automatically be
detected by the query processor.

Installation

A module or package can be installed with the REPO INSTALL command. The path to the file has to be given
as a parameter:

REPO INSTALL http://files.basex.org/modules/expath/functx-1.0.xar
REPO INSTALL hello-world.xqm

The installation will only succeed if the specified file conforms to the constraints described below. If you know
that your input is valid, you may as well copy the files directly to the repository directory, or edit its contents in
the repository without deleting and reinstalling them.

Listing

All currently installed packages can be listed with the REPO LIST command. It will return the names of all
packages, their version, and the directory in which they are installed:

<table>
<thead>
<tr>
<th>URI</th>
<th>Version</th>
<th>Directory</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://www.functx.com">http://www.functx.com</a></td>
<td>1.0</td>
<td>http-www.functx.com-1.0</td>
</tr>
</tbody>
</table>

1 package(s).

Removal

A package can be deleted with the command REPO DELETE and an additional argument, containing its name or
the name suffixed with a hyphen and the package version:

REPO DELETE http://www.functx.com ...or...
REPO DELETE http://www.functx.com-1.0

Packaging

XQuery

If an XQuery file is specified as input for the install command, it will be parsed as XQuery library module. If
parsing was successful, the module URI will be rewritten to a file path and attached with the .xqm file suffix, and
the original file will be renamed and copied to that path into the repository.

Example:

Installation (the original file will be copied to the org/basex/modules/Hello sub-directory of the
repository):

REPO INSTALL http://files.basex.org/modules/org/basex/modules/Hello/HelloWorld.xqm

Importing the repository module:

import module namespace m = 'http://basex.org/modules/Hello';
m:hello("Universe")

Java

Suitable JAR archives may contain one or more class files. One of them will be chosen as main class, which must
be specified in a Main-Class entry in the manifest file (META-INF/MANIFEST.MF). This fully qualified
Java class name will be rewritten to a file path by replacing the dots with slashes and attaching with the .jar file suffix, and the original file will be renamed and copied to that path into the repository.

The public functions of this class can then be addressed from XQuery, using the class or file path as namespace URI, or an alternative writing that can be rewritten to the module file path. Moreover, a class may extend the QueryModule class to get access to the current query context and to be enriched by some helpful annotations (please consult Context Awareness of Java Bindings for more information).

**Example:**

Structure of the HelloWorld.jar archive:

```
META-INF/
  MANIFEST.MF
org/basex/modules/
  Hello.class
```

Contents of the file MANIFEST.mf (the whitespaces are obligatory):

```
Manifest-Version: 1.0
Main-Class: org.basex.modules.Hello
```

Contents of the file Hello.java (comments removed):

```
package org.basex.modules;
public class Hello {
    public String hello(final String world) {
        return "Hello " + world;
    }
}
```

Installation (the file will be copied to org/basex/modules/Hello.jar):

```
REPO INSTALL HelloWorld.jar
```

XQuery file HelloUniverse.xq (same as above):

```
import module namespace m = 'http://basex.org/modules/Hello';
m:hello("Universe")
```

After installing the module, all of the following URIs can be used in XQuery to import this module or call its functions:

```
http://basex.org/modules/Hello
org/basex/modules/Hello
org.basex.modules.Hello
```

Please be aware that the execution of Java code can cause side effects that conflict with the functional nature of XQuery, or may introduce new security risks. The article on Java Bindings gives more insight on how Java code is handled from the XQuery processor.

**EXPath Packaging**

The EXPath specification defines how the structure of a .xar archive shall look like. The package contains at its root a package descriptor named expath-pkg.xml. This descriptor presents some meta data about the package as well as the libraries which it contains and their dependencies on other libraries or processors.

**XQuery**

Apart from the package descriptor, a .xar archive contains a directory which includes the actual XQuery modules. For example, the FunctX XAR archive is packaged as follows:
Java

In case you want to extend BaseX with a Java archive, some additional requirements have to be fulfilled:

• Apart from the package descriptor expath-pkg.xml, the package has to contain a descriptor file at its root, defining the included jars and the binary names of their public classes. It must be named basex.xml and must conform to the following structure:

```xml
<package xmlns="http://expath.org/ns/pkg">
  <jar>...</jar>
  ....
  <class>...</class>
  <class>...</class>
  ....
</package>
```

• The jar file itself along with an XQuery file defining wrapper functions around the java methods has to reside in the module directory. The following example illustrates how java methods are wrapped with XQuery functions:

**Example:** Suppose we have a simple class Printer having just one public method print():

```java
package test;

public final class Printer {
  public String print(final String s) {
    return new Writer(s).write();
  }
}
```

We want to extend BaseX with this class and use its method. In order to make this possible we have to define an XQuery function which wraps the print method of our class. This can be done in the following way:

```xml
import module namespace j="http://basex.org/lib/testJar";
declare namespace p="java:test.Printer";
declare function j:print($str as xs:string) as xs:string {
  let $printer := p:new()
  return p:print($printer, $str)
};
```

As it can be seen, the class Printer is declared with its binary name as a namespace prefixed with "java" and the XQuery function is implemented using the Java Bindings offered by BaseX.

On our file server, you can find some example libraries packaged as XML archives (xar files). You can use them to try our packaging API or just as a reference for creating your own packages.

**URI Rewriting**

If a module is looked up in the repository, the namespace URI is rewritten to a local file path:

1. If the URI is a URL:
   a. colons will be replaced with slashes,
b. in the URI authority, the order of all substrings separated by dots is reversed, and
c. dots in the authority and the path are replaced by slashes. If no path exists, a single slash is appended.

2. Otherwise, if the URI is a URN, colons will be replaced with slashes.

3. Characters other than letters, dots and slashes will be replaced with dashes.

4. If the resulting string ends with a slash, the index string is appended.

If the resulting path has no file suffix, it may point to either an XQuery module or a Java archive. The following examples show some rewritings:

• http://basex.org/modules/hello/World → org/basex/modules/hello/World
• http://www.example.com → com/example/www/index
• a/little/example → a/little/example
• a:b:c → a/b/c

Changelog

Version 8.2
• Added: URI Rewriting: support for URNs

Version 7.2.1
• Updated: Installation: existing packages will be replaced without raising an error
• Updated: Removal: remove specific version of a package
• Added: Packaging, URI Rewriting

Version 7.1
• Added: Repository Module

Version 7.0
• Added: EXPath Packaging
Chapter 27. Full-Text

Read this entry online in the BaseX Wiki.

This article is part of the XQuery Portal. It summarizes the full-text features of BaseX.

Full-text retrieval in XML documents is an essential requirement in many use cases. BaseX was the first query processor that supported the W3C XQuery Full Text 1.0 Recommendation, and it additionally comes with a powerful Full-Text Index, which allows you to evaluate full-text queries on large databases within milliseconds.

Introduction

The XQuery and XPath Full Text Recommendation (XQFT) is a feature-rich extension of the XQuery language. It can be used to both query XML documents and single strings for words and phrases. This section gives you a quick insight into the most important features of the language.

This is a simple example for a basic full-text expression:

"This is YOUR World" contains text "your world"

It yields true, because the search string is tokenized before it is compared with the tokenized input string. In the tokenization process, several normalizations take place. Many of those steps can hardly be simulated with plain XQuery: as an example, upper/lower case and diacritics (umlauts, accents, etc.) are removed and an optional, language-dependent stemming algorithm is applied. Beside that, special characters such as whitespaces and punctuation marks will be ignored. Thus, this query also yields true:

"Well... Done!" contains text "well, done"

The occurs keyword comes into play when more than one occurrence of a token is to be found:

"one and two and three" contains text "and" occurs at least 2 times

Varius range modifiers are available: exactly, at least, at most, and from ... to ....

Combining Results

In the given example, curly braces are used to combine multiple keywords:

for $country in doc('factbook')//country
where $country//religions[text() contains text { 'Sunni', 'Shia' } any]
return $country/name

The query will output the names of all countries with a religion element containing Sunni or Shia. The any keyword is optional; it can be replaced with:

• all: all strings need to be found
• any word: any of the single words within the specified strings need to be found
• all words: all single words within the specified strings need to be found
• phrase: all strings need to be found as a single phrase

The keywords ftand, ftor and ftnot can also be used to combine multiple query terms. The following query yields the same result as the last one does:
The keywords not in are special: they are used to find tokens which are not part of a longer token sequence:

```xml
for $text in ("New York", "new conditions")
return $text contains text "New" not in "New York"
```

Due to the complex data model of the XQuery Full Text spec, the usage of `ftand` may lead to a high memory consumption. If you should encounter problems, simply use the `all` keyword:

```xml
doc('factbook')//country[descendant::religions contains text { 'Christian', 'Jewish'} all]/name
```

### Positional Filters

A popular retrieval operation is to filter texts by the distance of the searched words. In this query…

```xml
<xml>
  <text>There is some reason why ...</text>
  <text>For some good yet unknown reason, ...</text>
  <text>The reason why some people ...</text>
</xml>//text[. contains text { "some", "reason" } all ordered distance at most 3 words]
```

…the two first texts will be returned as result, because there are at most three words between `some` and `reason`. Additionally, the `ordered` keyword ensures that the words are found in the specified order, which is why the third text is excluded. Note that `all` is required here to guarantee that only those hits will be accepted that contain all searched words.

The `window` keyword is related: it accepts those texts in which all keyword occur within the specified number of tokens. Can you guess what is returned by the following query?

```xml
("A C D", "A B C D E")[. contains text { "A", "E" } all window 3 words]
```

Sometimes it is interesting to only select texts in which all searched terms occur in the same sentence or paragraph (you can even filter for different sentences/paragraphs). This is obviously not the case in the following example:

```xml
"'I will survive!' This is what Mary told me." contains text { 'will', 'told' } all words same sentence
```

Sentences are delimited by end of line markers (., !, ?, etc.), and newline characters are treated as paragraph delimiters. By the way: in the examples above, the `word` unit has been used, but `sentences` and `paragraphs` are valid alternatives.

Last but not least, three specifiers exist to filter results depending on the position of a hit:

- `at start` expects tokens to occur at the beginning of a text
- `at end` expects tokens to occur at the text end
- `entire content` only accepts texts which have no other words at the beginning or end

### Match Options

As indicated in the introduction, the input and query texts are tokenized before they are compared with each other. During this process, texts are split into tokens, which are then normalized, based on the following matching options:
• If case is insensitive, no distinction is made between characters in upper and lower case. By default, the option is insensitive; it can also be set to sensitive:

"Respect Upper Case" contains text "Upper" using case sensitive

• If diacritics is insensitive, characters with and without diacritics (umlauts, characters with accents) are declared as identical. By default, the option is insensitive; it can also be set to sensitive:

"Äpfel" will not be found..." contains text "Apfel" diacritics sensitive

• If stemming is activated, words are shortened to a base form by a language-specific stemmer:

"catch" contains text "catches" using stemming,
"Haus" contains text "Häuser" using stemming using language 'de'

• With the stop words option, a list of words can be defined that will be ignored when tokenizing a string. This is particularly helpful if the full-text index takes too much space (a standard stopword list for English texts is provided in the directory etc/stopwords.txt in the full distributions of BaseX, and available online at http://files.basex.org/etc/stopwords.txt):

"You and me" contains text "you or me" using stop words ("and", "or"),
"You and me" contains text "you or me" using stop words at "http://files.basex.org/etc/stopwords.txt"

• Related terms such as synonyms can be found with the sophisticated Thesaurus option.

The wildcards option facilitates search operations similar to simple regular expressions:

• . matches a single arbitrary character.
• .? matches either zero or one character.
• .* matches zero or more characters.
• .+ matches one or more characters.
• .{min,max} matches min–max number of characters.

"This may be interesting in the year 2000" contains text { "interest.*", "2. {3,3}" } using wildcards

This was a quick introduction to XQuery Full Text; you are invited to explore the numerous other features of the language!

**BaseX Features**

**Languages**

The chosen language determines how the input text will be tokenized and stemmed. The basic code base and jar file of BaseX comes with built-in support for English and German. More languages are supported if the following libraries are found in the classpath:

• lucene-stemmers-3.4.0.jar : includes Snowball and Lucene stemmers and extends language support to the following languages: Bulgarian, Catalan, Czech, Danish, Dutch, Finnish, French, Greek, Hindi, Hungarian, Indonesian, Italian, Latvian, Lithuanian, Norwegian, Portuguese, Romanian, Russian, Spanish, Swedish, Turkish.
• **igo-0.4.3.jar**: An additional article explains how Igo can be integrated, and how Japanese texts are tokenized and stemmed.

The JAR files can also be found in the zip and exe distribution files of BaseX.

The following two queries, which both return **true**, demonstrate that stemming depends on the selected language:

```
"Indexing" contains text "index" using stemming,
"häuser" contains text "haus" using stemming using language "de"
```

**Scoring**

The XQuery Full Text Recommendation allows for the usage of scoring models and values within queries, with scoring being completely implementation-defined.

The scoring model of BaseX takes into consideration the number of found terms, their frequency in a text, and the length of a text. The shorter the input text is, the higher scores will be:

```
{: Score values: 1 0.62 0.45 :}
for $text score $score in ("A", "A B", "A B C")[.. contains text "A"]
order by $score descending
return <hit score='{ format-number($score, "0.00") }'>{ $text }</hit>
```

This simple approach has proven to consistently deliver good results, and in particular when little is known about the structure of the queried XML documents.

Please note that scores will only be computed if a parent expression requests them:

```
{: Computes and returns a scoring value. :}
let score $score := <x>Hello Universe</x> contains text "hello"
return $score

{: No scoring value will be computed here. :}
let $result := <x>Hello Universe</x> contains text "hello"
let score $score := $result
return $score
```

Scores will be propagated by the **and** and **or** expressions and in predicates. In the following query, all returned scores are equal:

```
let $text := "A B C"
let score $s1 := $text contains text "A" ftand "B C"
let score $s2 := $text contains text "A" ftand "B C"
let score $s3 := $text contains text "A" and $text contains text "B C"
let score $s4 := $text contains text "A" or $text contains text "B C"
let score $s5 := $text[.. contains text "A"][.. contains text "B C"]
return ($s1, $s2, $s3, $s4, $s5)
```

**Thesaurus**

BaseX supports full-text queries using thesauri, but it does not provide a default thesaurus. This is why queries such as:

```
'computers' contains text 'hardware'
using thesaurus default
```

will return **false**. However, if the thesaurus is specified, then the result will be **true**:
'computers' contains text 'hardware'
using thesaurus at 'XQFTTS_1_0_4/TestSources/usability2.xml'

The format of the thesaurus files must be the same as the format of the thesauri provided by the XQuery and XPath Full Text 1.0 Test Suite. It is an XML with structure defined by an XSD Schema.

**Fuzzy Querying**

In addition to the official recommendation, BaseX supports a fuzzy search feature. The XQFT grammar was enhanced by the FTMatchOption using fuzzy to allow for approximate results in full texts. Fuzzy search is also supported by the full-text index.

**Document 'doc.xml':**

```xml
<doc>
  <a>house</a>
  <a>hous</a>
  <a>haus</a>
</doc>
```

**Query:**

```
//a[text() contains text 'house' using fuzzy]
```

**Result:**

```xml
<a>house</a>
<a>hous</a>
```

Fuzzy search is based on the Levenshtein distance. The maximum number of allowed errors is calculated by dividing the token length of a specified query term by 4, preserving a minimum of 1 errors. A static error distance can be set by adjusting the LSERROR property (default: SET LSERROR 0). The query above yields two results as there is no error between the query term “house” and the text node “house”, and one error between “house” and “hous”.

**Mixed Content**

When working with so-called narrative XML documents, such as HTML, TEI, or DocBook documents, you typically have mixed content, i.e., elements containing a mix of text and markup, such as:

```xml
<p>This is only an illustrative <hi>example</hi>, not a <q>real</q> text.</p>
```

Since the logical flow of the text is not interrupted by the child elements, you will typically want to search across elements, so that the above paragraph would match a search for “real text”. For more examples, see XQuery and XPath Full Text 1.0 Use Cases.

To enable this kind of searches, whitespace chopping must be turned off when importing XML documents by setting the option CHOP to OFF (default: SET CHOP ON). In the GUI, you find this option in Database → New… → Parsing → Chop Whitespaces. A query such as `//p[. contains text 'real text']` will then match the example paragraph above. However, the full-text index will not be used in this query, so it may take a long time. The full-text index would be used for the query `//p[text() contains text 'real text']`, but this query will not find the example paragraph, because the matching text is split over two text nodes.

Note that the node structure is ignored by the full-text tokenizer: The `contains text` expression applies all full-text operations to the string value of its left operand. As a consequence, the `ft:mark` and `ft:extract` functions (see Full-Text Functions) will only yield useful results if they are applied to single text nodes, as the following example demonstrates:
BaseX does not support the `ignore option (without content)` of the W3C XQuery Full Text 1.0 Recommendation. If you want to ignore descendant element content, such as footnotes or other material that does not belong to the same logical text flow, you can build a second database from and exclude all information you do not want to search for. See the following example (visit XQuery Update to learn more about updates):

```xml
let $docs := db:open('docs')
return db:create('index-db',
    $docs update delete node (.//footnote),
    $docs/db:path(.),
    map { 'ftindex': true() })
```

## Functions

Some additional Full-Text Functions have been added to BaseX to extend the official language recommendation with useful features, such as explicitly requesting the score value of an item, marking the hits of a full-text request, or directly accessing the full-text index with the default index options.

## Collations

See XQuery 3.1 for standard collation features.

By default, string comparisons in XQuery are based on the Unicode codepoint order. The default namespace URI `http://www.w3.org/2003/05/xpath-functions/collation/codepoint` specifies this ordering. In BaseX, the following URI syntax is supported to specify collations:

```xml
http://basex.org/collation?lang=...;strength=...;decomposition=...
```

Semicolons can be replaced with ampersands; for convenience, the URL can be reduced to its query string component (including the question mark). All arguments are optional:

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>lang</td>
<td>A language code, selecting a Locale. It may be followed by a language variant. If no language is specified, the system’s default will be chosen. Examples: de, en-US.</td>
</tr>
<tr>
<td>strength</td>
<td>Level of difference considered significant in comparisons. Four strengths are supported: primary, secondary, tertiary, and identical. As an example, in German:</td>
</tr>
<tr>
<td></td>
<td>• “Ä” and “A” are considered primary differences,</td>
</tr>
<tr>
<td></td>
<td>• “Ä” and “ä” are secondary differences,</td>
</tr>
<tr>
<td></td>
<td>• “Ä” and “A&amp;#x308;” (see <a href="http://www.fileformat.info/info/unicode/char/308/index.htm">http://www.fileformat.info/info/unicode/char/308/index.htm</a>) are tertiary differences, and</td>
</tr>
<tr>
<td></td>
<td>• “A” and “A” are identical.</td>
</tr>
<tr>
<td>decomposition</td>
<td>Defines how composed characters are handled. Three decompositions are supported: none, standard, and full. More details are found in the JavaDoc of the JDK.</td>
</tr>
</tbody>
</table>

Some Examples:
• If a default collation is specified, it applies to all collation-dependent string operations in the query. The following expression yields true:

```sql
declare default collation 'http://baseX.org/collation?lang=de;strength=secondary';
'Straße' = 'Strasse'
```

• Collations can also be specified in order by and group by clauses of FLWOR expressions. This query returns à plutôt! bonjour!:

```sql
for $w in ("bonjour!", "à plutôt!") order by $w collation "?lang=fr" return $w
```

• Various string function exists that take an optional collation as argument: The following functions give us a and 1 2 3 as results:

```sql
distinct-values(("a", "á", "à"), "?lang=it-IT;strength=primary"),
index-of(("a", "á", "à"), "a", "?lang=it-IT;strength=primary")
```

## Changelog

### Version 8.0

• Updated: Scores will be propagated by the and and or expressions and in predicates.

### Version 7.7

• Added: Collations support.

### Version 7.3

• Removed: Trie index, which was specialized on wildcard queries. The fuzzy index now supports both wildcard and fuzzy queries.

• Removed: TF/IDF scoring was discarded in favor of the internal scoring model.
Chapter 28. Full-Text: Japanese

This article is linked from the Full-Text page. It gives some insight into the implementation of the full-text features for Japanese text corpora. The Japanese version is also available as PDF. Thank you to Toshio HIRAI for integrating the lexer in BaseX!

Introduction

The lexical analysis of Japanese documents is performed by Igo. Igo is a morphological analyser, and some of the advantages and reasons for using Igo are:

• compatible with the results of a prominent morphological analyzer "MeCab"
• it can use the dictionary distributed by the Project MeCab
• the morphological analyzer is implemented in Java and is relatively fast

Japanese tokenization will be activated in BaseX if Igo is found in the classpath. igo-0.4.3.jar of Igo is currently included in all distributions of BaseX.

In addition to the library, one of the following dictionary files must either be unzipped into the current directory, or into the etc sub-directory of the project’s Home Directory:

• IPA Dictionary: http://files.basex.org/etc/ipadic.zip
• NAIST Dictionary: http://files.basex.org/etc/naistdic.zip

Lexical Analysis

The example sentence "##########(I wrote a book.)" is analyzed as follows.

```
####    ####,####,####,
####    ####,####,####,####,
####    ####,####,####,####,####,
####    ####,####,####,####,####,####,
####    ####,####,####,####,####,####,####,
####    #,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,##,
• Filler
• Postpositional particle
• Auxiliary verb

Thus, in the example above, #, #, and ## will be passed to the indexer for each token.

**Token Processing**

"Fullwidth" and "Halfwidth" (which is defined by East Asian Width Properties) are not distinguished (this is the so-called ZENKAKU/HANKAKU problem). For example, ### and XML will be treated as the same word. If documents are hybrid, i.e. written in multiple languages, this is also helpful for some other options of the XQuery Full Text Specification, such as the Case or the Diacritics Option.

**Stemming**

Stemming in Japanese means to analyze the results of morphological analysis ("verbs" and "adjectives") that are processed using the "prototype".

If the stemming option is enabled, for example, the two statements "####### (I wrote the book)") and "#### (I write the book)" can be led back to the same prototype by analyzing their verb:

```
##    ##,##,*,*,########,###,#,#
##    ##,##,*,*,########,####,##
#      ###,*,*,*,####,###,#,#
```

Because the "auxiliary verb" is always excluded from the tokens, there is no need to consider its use. Therefore, the same result (true) is returned for the following two types of queries:

'#####' contains text '.##' using wildcards using language 'ja'
'#####' contains text '.##' using wildcards using language 'ja'

**Wildcards**

The Wildcard option in XQuery Full-Text is available for Japanese as well. The following example is based on '# # ###(AKUTAGAWA, Ryunosuke)', a prominent Japanese writer, the first name of whom is often spelled as "# ". The following two queries both return true:

'#####' contains text '.##' using wildcards using language 'ja'
'#####' contains text '.##' using wildcards using language 'ja'

However, there is a special case that requires attention. The following query will yield false:

'#####' contains text '##.##' using wildcards using language 'ja'

This is because the next word boundary metacharacters cannot be determined in the query. In this case, you may insert an additional whitespaces as word boundary:

'#####' contains text '## .##' using wildcards using language 'ja'

As an alternative, you may modify the query as follows:

'#####' contains text '## fortand '.##' using wildcards using language 'ja'
Chapter 29. XQuery Update

Read this entry online in the BaseX Wiki.

This article is part of the XQuery Portal. It summarizes the update features of BaseX.

BaseX offers a complete implementation of the XQuery Update Facility (XQUF). This article aims to provide a very quick and basic introduction to the XQUF. First, some examples for update expressions are given. After that, a few problems are addressed that frequently arise due to the nature of the language. These are stated in the Concepts paragraph.

Features

Updating Expressions

There are five new expressions to modify data. While insert, delete, rename and replace are basically self-explanatory, the transform expression is different, as modified nodes are copied in advance and the original databases remain untouched.

An expression consists of a target node (the node we want to alter) and additional information like insertion nodes, a QName, etc. which depends on the type of expression. Optional modifiers are available for some of them. You can find a few examples and additional information below.

insert

```
insert node (attribute ( 'a' ) { 5 }, 'text', <e/>) into /n
```

Insert enables you to insert a sequence of nodes into a single target node. Several modifiers are available to specify the exact insert location: insert into as first/as last, insert before/after and insert into.

Note: in most cases, as last and after will be evaluated faster than as first and before!

delete

```
delete node //n
```

The example query deletes all <n> elements in your database. Note that, in contrast to other updating expressions, the delete expression allows multiple nodes as a target.

replace

```
replace node /n with <a/>
```

The target element is replaced by the DOM node <a/>. You can also replace the value of a node or its descendants by using the modifier value of.

```
replace value of node /n with 'newValue'
```

All descendants of /n are deleted and the given text is inserted as the only child. Note that the result of the insert sequence is either a single text node or an empty sequence. If the insert sequence is empty, all descendants of the target are deleted. Consequently, replacing the value of a node leaves the target with either a single text node or no descendants at all.

rename

```
for $n in //originalNode
return rename node $n as 'renamedNode'
```
All originalNode elements are renamed. An iterative approach helps to modify multiple nodes within a single statement. Nodes on the descendant- or attribute-axis of the target are not affected. This has to be done explicitly as well.

Non-Updating Expressions

**transform**

```xquery
transform

copy $c := doc('example.xml')//originalNode[@id = 1]
modify rename node $c as 'copyOfNode'
return $c
```

The originalNode element with @id=1 is copied and subsequently assigned a new QName using the rename expression. Note that the transform expression is the only expression which returns an actual XDM instance as a result. You can therefore use it to modify results and especially DOM nodes. This is an issue beginners are often confronted with. More on this topic can be found in the XQUF Concepts section.

The following example demonstrates a common use case:

**Query:**

```xquery
copy $c :=
  <entry>
    <title>Transform expression example</title>
    <author>BaseX Team</author>
  </entry>
modify (replace value of node $c/author with 'BaseX',
  replace value of node $c/title with concat('Copy of: ', $c/title),
  insert node <author>Joey</author> into $c)
return $c
```

**Result:**

```xml
<entry>
  <title>Copy of: Transform expression example</title>
  <author>BaseX</author>
  <author>Joey</author>
</entry>
```

The `<entry>` element (here it is passed to the expression as a DOM node) can also be replaced by a database node, e.g.:

```xquery
copy $c := (db:open('example')//entry)[1]
...
```

In this case, the original database node remains untouched as well, as all updates are performed on the node copy.

Here is an example where we return an entire document, parts modified and all:

```xquery
copy $c := doc("zaokeng.kml")
modify (for $d in $c//*:Point
  return insert node (
    <extrude>1</extrude>,
    <altitudeMode>relativeToGround</altitudeMode>
  ) before $d/*/coordinates)
```
return $c

update

for $item in db:open('data')//item
return $item update delete node text()

The update expression is a convenience operator for writing simple transform expressions. Similar to the XQuery 3.0 map operator, the value of the first expression is bound as context item, and the second expression performs updates on this item. The updated item is returned as result.

Please note that update is not part of the official XQuery Update Facility yet. It is currently being discussed in the W3 Bug Tracker; your feedback is welcome.

Functions

Built-in Functions

fn:put() can be used to serialize XDM instances to secondary storage. It is executed at the end of a snapshot. Serialized documents therefore reflect all changes made effective during a query.

Additional database functions exist for performing updates on document and database level.

User-Defined Functions

If an updating function item is called, the function call must be prefixed with the keyword updating. This ensures that the query compiler can statically detect if an invoked function item will perform updates or not:

let $node := <node>TO-BE-DELETED</node>
let $delete-text := %updating function($node) {
  delete node $node//text()
}
return $node update {
  updating $delete-text(.)
}

As shown in the example, user-defined and anonymous functions can additionally be annotated as %updating.

Concepts

There are a few specialties around XQuery Update that you should know about. In addition to the simple expression, the XQUF adds the updating expression as a new type of expression. An updating expression returns only a Pending Update List (PUL) as a result which is subsequently applied to addressed databases and DOM nodes. A simple expression cannot perform any permanent changes and returns an empty or non-empty sequence.

Pending Update List

The most important thing to keep in mind when using XQuery Update is the Pending Update List (PUL). Updating statements are not executed immediately, but are first collected as update primitives within a set-like structure. At the end of a query, after some consistency checks and optimizations, the update primitives will be applied in the following order:

- **Backups** (1): db:create-backup()
- **XQuery Update**: insert before, delete, replace, rename, replace value, insert attribute, insert into first, insert into, insert into last, insert, insert after, put
- **Documents**: db:add(), db:store(), db:replace(), db:rename(), db:delete(), db:optimize(), db:flush()
• **Users**: `user:grant()`, `user:password()`, `user:drop()`, `user:alter()`, `user:create()`

• **Databases**: `db:copy()`, `db:drop()`, `db:alter()`, `db:create()`

• **Backups (2)**: `db:restore()`, `db:drop-backup()`

If an inconsistency is found, an error message is returned and all accessed databases remain untouched (atomicity). For the user, this means that updates are only visible after the end of a snapshot.

It may be surprising to see `db:create` in the lower part of this list. This means that newly created database cannot be accessed by the same query, which can be explained by the semantics of updating queries: all expressions can only be evaluated on databases that already exist while the query is evaluated. As a consequence, `db:create` is mainly useful in the context of **Command Scripts**, or **Web Applications**, in which a redirect to another page can be triggered after having created a database.

**Example**

The query…

```xquery
insert node <b/> into /doc,
for $n in /doc/child::node()
return rename node $n as 'justRenamed'
```

…applied on the document…

```
<doc> <a/> </doc>
```

…results in the following document:

```
<doc> <justRenamed/> <b/> </doc>
```

Despite explicitly renaming all child nodes of `<doc/>`, the former `<a/>` element is the only one to be renamed. The `<b/>` element is inserted within the same snapshot and is therefore not yet visible to the user.

**Returning Results**

By default, it is not possible to mix different types of expressions in a query result. The outermost expression of a query must either be a collection of updating or non-updating expressions. But there are two ways out:

• The BaseX-specific `db:output()` function bridges this gap: it caches the results of its arguments at runtime and returns them after all updates have been processed. The following example performs an update and returns a success message:

```xquery
db:output("Update successful."), insert node <c/> into doc('factbook')/mondial
```

• With the **MIXUPDATES** option, all updating constraints will be turned off. Returned nodes will be copied before they are modified by updating expressions. An error is raised if items are returned within a transform expression.

If you want to modify nodes in main memory, you can use the *transform expression*.

**Effects**

**Original Files**

In BaseX, all updates are performed on database nodes or in main memory. By default, update operations do not affect the original input file (the info string “Updates are not written back” appears in the query info to indicate this). The following solutions exist to write XML documents and binary resources to disk:
XQuery Update

- Updates on main-memory instances of files that have been retrieved via `fn:doc` or `fn:collection` will be propagated back to disk when the WRITEBACK option is turned on. This option can also be activated on command line via `-u`. Make sure you back up the original documents before running your queries.

- Functions like `fn:put` or `file:write` can be used to write single XML documents to disk. With `file:write-binary`, you can write binary resources.

- The EXports command can be used write all resources of a databases to disk.

## Indexes

Index structures are discarded after update operations when UPDINDEX is turned off (which is the default). More details are found in the article on Indexing.

## Error Messages

Along with the Update Facility, a number of new error codes and messages have been added to the specification and BaseX. All errors are listed in the XQuery Errors overview.

## Changelog

### Version 8.0

- Added: MIXUPDATES option for Returning Results in updating expressions
- Added: information message if files are not written back

### Version 7.8

- Added: update convenience operator
Chapter 30. Indexes

Read this entry online in the BaseX Wiki.

This article is part of the XQuery Portal. It contains information on the available index structures.

The query compiler tries to optimize and speed up queries by applying the index whenever it is possible and seems promising. To see how a query is rewritten, and if an index is used, you can turn on the Info View in the GUI or use the -V flag on the command line:

• A message like Applying text index for "Japan" indicates that the text index is applied to speed up the search of the shown string. The following message…

• Removing path with no index results indicates that a string in a path expression will never yield results. Because of that, the path does not need to be evaluated at all.

• If you cannot find any index optimization hints in the info output, it often helps if you rewrite and simplify your query.

Structural Indexes

Structural indexes are automatically created and cannot be dropped by the user:

Name Index

The name index contains all element and attribute names of a database, and the fixed-size index ids are stored in the main database table. If a database is updated, new names are automatically added. Furthermore, the index is enriched with statistical information, such as the distinct (categorical) or minimum and maximum values of its elements and attributes. The maximum number of categories to store per name can be changed via MAXCATS. The index are discarded after database updates.

The name index is e.g. applied to pre-evaluate location steps that will never yield results:

(: will be rewritten to an empty sequence :)
/doc('factbook.xml')/non-existing-name

The contents of the name indexes can be directly accessed with the XQuery functions index:element-names and index:attribute-names.

Path Index

The path index (also called path summary) stores all distinct paths of the documents in the database. It contains similar statistical information as the name index. The index are discarded after database updates.

The path index is applied to rewrite descendant steps to multiple child steps. Child steps can be evaluated faster, as fewer nodes have to be accessed:

doc('factbook.xml')//province,  
(: ...will be rewritten to... :)  
doc('factbook.xml')/mondial/country/province

The paths statistics are e.g. used to pre-evaluate the count function:

(: will be rewritten and pre-evaluated by the path index :)  
count( doc('factbook')//country )

The contents of the path index can be directly accessed with the XQuery function index:facets.
Indexes

Document Index

The document index contains references to all document nodes in a database. Once documents with specific paths are requested, the index will be extended to also contain document paths.

The index generally speeds up access to single documents and database paths. It will always be kept up-to-date.

Value Indexes

Value indexes can be created and dropped by the user. Four types of values indexes are available: a text and attribute index, and an optional token and full-text index. By default, the text and attribute index will automatically be created.

In the GUI, index structures can be managed in the dialog windows for creating new databases or displaying the database properties. On command-line, the commands `CREATE INDEX` and `DROP INDEX` are used to create and drop index structures. With `INFO INDEX`, you get some insight into the contents of an index structure, and `SET` allows you to change the index defaults for new databases:

- `OPEN factbook; CREATE INDEX fulltext`: Open database; create full-text index
- `OPEN factbook; INFO INDEX TOKEN`: Open database; show info on token index
- `SET ATTRINDEX true; SET ATTRINCLUDE id name; CREATE DB factbook.xml`: Enable attribute index; only index 'id' and 'name' attributes; create database

With XQuery, index structures can be created and dropped via `db:optimize`:

```
(db:optimize(
  'factbook',
  false(),
  map { 'ftindex': true(), 'ftinclude': 'p div' }
))
```

Text Index

Exact Queries

This index speeds up string-based equality tests on text nodes. The following queries will all be rewritten for index access:

```
:*[text() = 'Germany'],
(doc('factbook.xml')//name[. = 'Germany'],
for $c in db:open('factbook')//country
where $c//city/name = 'Hanoi'
return $c/name
```

Matching text nodes can be directly requested from the index with the XQuery function `db:text`. The index contents can be accessed via `index:text`.

The `UPDINDEX` option can be activated to keep this index up-to-date.

Range Queries

The text index also supports range queries based on string comparisons:
Indexes

Text nodes can directly be retrieved from the index via the XQuery function `db:text-range`.

Please note that the current index structures do not support queries for numbers and dates.

**Attribute Index**

Similar to the text index, this index speeds up string-based equality and range tests on attribute values. Additionally, the XQuery function `fn:id` takes advantage of the index whenever possible. The following queries will all be rewritten for index access:

```
(: 1st example :)
//country[@car_code = 'J'],
(: 2nd example :)
//province[@* = 'Hokkaido']//name,
(: 3rd example :)
//sea[@depth > '2100' and @depth < '4000']
(: 4th example :)
fn:id('f0_119', db:open('factbook'))
```

Attribute nodes can directly be retrieved from the index with the XQuery functions `db:attribute` and `db:attribute-range`. The index contents can be accessed with `index:attributes`.

The **UPDINDEX** option can be activated to keep this index up-to-date.

**Token Index**

In many XML dialects, such as HTML or DITA, multiple tokens are stored in attribute values. The token index can be created to speed up the retrieval of these tokens. The XQuery functions `fn:contains-token`, `fn:tokenize` and `fn:idref` are rewritten for index access whenever possible. If a token index exists, it will e.g. be utilized for the following queries:

```
(: 1st example :)
//div[contains-token(@class, 'row')],
(: 2nd example :)
//p[tokenize(@class) = 'row'],
(: 3rd example :)
doc('graph.xml')/idref('edge8')
```

Attributes with tokens can be directly retrieved from the index with the XQuery function `db:token`. The index contents can be accessed with `index:tokens`.

**Full-Text Index**

The **Full-Text** index contains the normalized tokens of text nodes of a document. It is utilized to speed up queries with the `contains text` expression, and it is capable of processing wildcard and fuzzy search operations. Three evaluation strategies are available: the standard sequential database scan, a full-text index based evaluation and a hybrid one, combining both strategies (see XQuery Full Text implementation in BaseX).

If the full-text index exists, the following queries will all be rewritten for index access:

```
(: 1st example :)
//country[name/text() contains text 'and'],
```
Indexes

The index provides support for the following full-text features (the values can be changed in the GUI or via the SET command):

• **Stemming**: tokens are stemmed before being indexed (see STEMMING)

• **Case Sensitive**: tokens are indexed in case-sensitive mode (see CASESENS)

• **Diacritics**: diacritics are indexed as well (see DIACRITICS)

• **Stopword List**: a stop word list can be defined to reduce the number of indexed tokens (see STOPWORDS)

• **Language**: see Languages for more details (see LANGUAGE)

The options that have been used for creating the full-text index will also be applied to the optimized full-text queries. However, the defaults can be overwritten if you supply options in your query. For example, if words were stemmed in the index, and if the query can be rewritten for index access, the query terms will be stemmed as well, unless stemming is not explicitly disabled. This is demonstrated in the following Command Script:

```xml
<commands>
  <!-- Create database with stemmed full-text index -->
  <set option='stemming'>true</set>
  <set option='ftindex'>true</set>
  <create-db name='test-db'> <text>house</text> </create-db>
  <!-- Index access: Query term will be stemmed -->
  <xquery> /text[. contains text { 'houses' }] </xquery>
  <!-- Disable stemming (query will not be evaluated by the index) -->
  <xquery> /text[. contains text { 'houses' } using no stemming] </xquery>
</commands>
```

Text nodes can be directly requested from the index via the XQuery function ft:search. The index contents can be accessed with ft:tokens.

### Selective Indexing

Value indexing can be restricted to specific elements and attributes. The nodes to be indexed can be restricted via the TEXTINCLUDE, ATTRINCLUDE, TOKENINCLUDE and FTINCLUDE options. The options take a list of name patterns, which are separated by commas. The following name patterns are supported:

• `*`: all names

• `name`: elements or attributes called name, which are in the empty default namespace

• `*:name`: elements or attributes called name, no matter which namespace

• `Q{uri}*`: all elements or attributes in the uri namespace

• `Q{uri}name`: elements or attributes called name in the uri namespace

The options can either be specified via the SET command or via XQuery. With the following operations, an attribute index is created for all id and name attributes:

```xml
SET ATTRINCLUDE id,name
CREATE DB factbook http://files.basex.org/xml/factbook.xml'
# Restore default
SET ATTRINCLUDE
```
Indexes

XQuery

```xquery
<query>
    <db:create('factbook', 'http://files.basex.org/xml/factbook.xml', '',
        map { 'attribinclude': 'id,name' })
</query>
```

With `CREATE INDEX` and `db:optimize`, new selective indexing options will be applied to an existing database.

**Custom Index Structures**

With XQuery, it is comparatively easy to create your own, custom index structures. The following query demonstrates how you can create a `factbook-index` database, which contains all texts of the original database in lower case:

```xquery
let $db := 'factbook'
let $index := <index>{
    for $nodes in db:open($db)//text()
        group by $text := lower-case($nodes)
        return <text string='{ $text }'>{
            for $node in $nodes
                return <id>{ db:node-id($node) }</id>
        }<text>
    }</index>
return db:create($db || '-index', $index, $db || '-index.xml')
```

In the following query, a text string is searched, and the text nodes of the original database are retrieved:

```xquery
let $db := 'factbook'
let $text := 'italian'
for $id in db:open($db || '-index')//*[@string = $text]/id
return db:open-id($db, $id)/..
```

With some extra effort, and if `UPDINDEX` is enabled for both your original and your index database (see below), your index database will support updates as well (try it, it’s fun!).

**Performance**

If main memory runs out while creating a value index, the current index structures will be partially written to disk and eventually merged. If the memory heuristics fail for some reason (i.e., because multiple index operations run at the same time, or because the applied JVM does not support explicit garbage collections), a fixed index split sizes may be chosen via the `SPLITSIZE` option.

If `DEBUG` is enabled, the command-line output might help you to find a good split size. The following example shows the output for creating a database for an XMark document with 1 GB, and with 128 MB assigned to the JVM:

```bash
> basex -d -c "SET FTINDEX ON; SET TOKENINDEX ON; CREATE DB xmark 1gb.xml"
Creating Database...                      76559.99 ms (29001 KB)
Indexing Text...                         ..|...|...|...|...| 9.81 M operations, 18576.92 ms (13523 KB). Recommended SPLITSIZE: 20.
Indexing Attribute Values...             ..........|....... 3.82 M operations, 9636.73 ms (10809 KB). Recommended SPLITSIZE: 10.
Indexing Tokens...                      ........|.|..|..|..|.. 3.82 M operations, 9636.73 ms (10809 KB). Recommended SPLITSIZE: 10.
Indexing Full-Text...
```

143
Indexes

The output can be interpreted as follows:

- The vertical bar | indicates that a partial index structure was written to disk.

- The mean value of the recommendations can be assigned to the SPLITSIZE option. Please note that the recommendation is only a vague proposal, so try different values if you get main-of-memory errors or indexing gets too slow. Greater values will require more main memory.

- In the example, the full-text index was split 12 times. 116 million tokens were indexed, processing time was 2.5 minutes, and final main memory consumption (after writing the index to disk) was 76 MB. A good value for the split size option could be 15.

Updates

Updates in BaseX are generally very fast, because the index structures will be invalidated by updates. As a result, subsequent queries may be executed more slowly than before the update. There are different alternatives to cope with this:

- After the execution of one or more update operations, the OPTIMIZE command or the db:optimize function can be called to rebuild the index structures.

- The UPDINDEX option can be activated before creating or optimizing the database. As a result, the text, attribute and token indexes will be incrementally updated after each database update. Please note that incremental updates are not available for the full-text index and database statistics. This is also explains why the up-to-date flag, which is e.g. displayed via INFO DB or db:info, will be set to false until the database will be optimized again.

- The AUTOOPTIMIZE option can be enabled before creating or optimizing the database. All outdated index structures and statistics will then be recreated after each database update. This option should only be used for small and medium-sized databases.

Changelog

Version 8.4

- Added: Token Index

Version 8.3

- Added: Selective Indexing

Version 8.0

- Added: AUTOOPTIMIZE option

Version 7.2.1

- Added: string-based range queries
Chapter 31. Serialization

Read this entry online in the BaseX Wiki.

This page is part of the XQuery Portal. Serialization parameters define how XQuery items and XML nodes are textually output, i.e., serialized. (For input, see Parsers.) They have been formalized in the W3C XQuery Serialization 3.1 document. In BaseX, they can be specified by...

- including them in the prolog of the XQuery expression,
- specifying them in the XQuery functions file:write() or fn:serialize(). The serialization parameters are specified as
  - children of an `<output:serialization-parameters/>` element, as defined for the `fn:serialize()` function, or as
  - map, which contains all key/value pairs: `map { "method": "xml", "cdata-section-elements": "div", ... }`.
- using the `-s` flag of the BaseX command-line clients,
- setting the SERIALIZER option before running a query,
- setting the EXPORTER option before exporting a database, or
- setting them as REST query parameters.

Parameters

The following table gives a brief summary of all serialization parameters recognized by BaseX. For details, please refer to official specification.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Allowed</th>
<th>Default</th>
</tr>
</thead>
<tbody>
<tr>
<td>method</td>
<td>Specifies the serialization method. xml, xhtml, html, text, json, and adaptive are adopted from the official specification. The methods <code>baseX</code> and <code>csv</code> are specific to BaseX (see XQuery Extensions).</td>
<td>xml, xhtml, html, text, json, adaptive, csv, baseX</td>
<td>baseX</td>
</tr>
<tr>
<td>version</td>
<td>Specifies the version of the serialization method.</td>
<td>xml/xhtml: 1.0, 1.1html: 4.0, 4.01, 5.0</td>
<td>1.0</td>
</tr>
<tr>
<td>html-version</td>
<td>Specifies the version of the HTML serialization method.</td>
<td>4.0, 4.01, 5.0</td>
<td>4.0</td>
</tr>
</tbody>
</table>
| item-separator| Determines a string to be used as item separator. If a separator is specified, the default separation of atomic values with single whitespaces will be skipped. | arbitrary strings, \n, \r
| encoding      | Encoding to be used for outputting the data.                                | all encodings supported by Java | UTF-8   |
| indent        | Adjusts whitespaces to make the output better readable.                      | yes, no                       | yes     |
### Serialization

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>cdata-section-elements</strong></td>
<td>List of elements to be output as CDATA, separated by whitespaces. Example:</td>
</tr>
<tr>
<td></td>
<td><code>&lt;text&gt;&lt;![CDATA[&lt; &gt;]]&gt;&lt;![CDATA[&gt;]]&lt;/text&gt;</code></td>
</tr>
<tr>
<td><strong>omit-xml-declaration</strong></td>
<td>Omits the XML declaration, which is serialized before the actual query result. Example:</td>
</tr>
<tr>
<td></td>
<td><code>&lt;?xml version=&quot;1.0&quot; encoding=&quot;UTF-8&quot;?&gt;</code></td>
</tr>
<tr>
<td><strong>standalone</strong></td>
<td>Prints or omits the &quot;standalone&quot; attribute in the XML declaration.</td>
</tr>
<tr>
<td><strong>doctype-system</strong></td>
<td>Introduces the output with a document type declaration and the given system identifier. Example:</td>
</tr>
<tr>
<td></td>
<td><code>&lt;!DOCTYPE x SYSTEM &quot;entities.dtd&quot;&gt;</code></td>
</tr>
<tr>
<td><strong>doctype-public</strong></td>
<td>If doctype-system is specified, adds a public identifier. Example:</td>
</tr>
<tr>
<td></td>
<td><code>&lt;!DOCTYPE HTML PUBLIC &quot;-//W3C//DTD HTML 4.01//EN&quot;http://www.w3.org/TR/html4/strict.dtd&gt;</code></td>
</tr>
<tr>
<td><strong>undeclare-prefixes</strong></td>
<td>Undeclares prefixes in XML 1.1.</td>
</tr>
<tr>
<td><strong>normalization-form</strong></td>
<td>Specifies a normalization form. BaseX supports Form C (NFC).</td>
</tr>
<tr>
<td><strong>media-type</strong></td>
<td>Specifies the media type.</td>
</tr>
<tr>
<td></td>
<td><code>application/xml</code></td>
</tr>
<tr>
<td><strong>parameter-document</strong></td>
<td>Parses the value as XML document with additional serialization parameters (see the Serialization Specification for more details).</td>
</tr>
<tr>
<td><strong>use-character-maps</strong></td>
<td>Defines character mappings. May only occur in documents parsed with parameter-document.</td>
</tr>
<tr>
<td><strong>byte-order-mark</strong></td>
<td>Prints a byte-order-mark before starting serialization.</td>
</tr>
<tr>
<td><strong>escape-uri-attributes</strong></td>
<td>Escapes URI information in certain HTML attributes. Example:</td>
</tr>
<tr>
<td></td>
<td><code>&lt;a href=&quot;%C3%A4%C3%B6%C3%BC&quot;&gt;äöü&lt;a&gt;</code></td>
</tr>
<tr>
<td><strong>include-content-type</strong></td>
<td>Inserts a meta content-type element into the head element if the result is output as HTML. Example:</td>
</tr>
<tr>
<td></td>
<td><code>&lt;head&gt;&lt;meta http-</code></td>
</tr>
</tbody>
</table>

Yes, no, omit
BaseX provides some additional serialization parameters:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Allowed</th>
<th>Default</th>
</tr>
</thead>
<tbody>
<tr>
<td>csv</td>
<td>Defines the way how data is serialized as CSV.</td>
<td>see CSV Module</td>
<td></td>
</tr>
<tr>
<td>json</td>
<td>Defines the way how data is serialized as JSON.</td>
<td>see JSON Module</td>
<td></td>
</tr>
<tr>
<td>tabulator</td>
<td>Uses tab characters (\t) instead of spaces for indenting elements.</td>
<td>yes, no</td>
<td>no</td>
</tr>
<tr>
<td>indents</td>
<td>Specifies the number of characters to be indented.</td>
<td>positive number</td>
<td>2</td>
</tr>
<tr>
<td>newline</td>
<td>Specifies the type of newline to be used as end-of-line marker.</td>
<td>\n, \r\n, \r</td>
<td>system dependent</td>
</tr>
<tr>
<td>limit</td>
<td>Stops serialization after the specified number of bytes has been serialized. If a negative number is specified, everything will be output.</td>
<td>positive number</td>
<td>-1</td>
</tr>
<tr>
<td>binary</td>
<td>Indicates if items of binary type are output in their native byte representation. Only applicable to the base serialization method.</td>
<td>yes, no</td>
<td>yes</td>
</tr>
</tbody>
</table>

The `csv` and `json` parameters are supplied with a list of options. Option names and values are combined with `=`, several options are separated by `,`.

**Query:**

```xml
(: The output namespace declaration is optional, because it is statically declared in BaseX) :)
declare namespace output = "http://www.w3.org/2010/xslt-xquery-serialization";
declare option output:method "csv";
declare option output:csv "header=yes, separator=semicolon";
<csv>
  <record>
    <Name>John</Name>
    <City>Newton</City>
  </record>
  <record>
    <Name>Jack</Name>
    <City>Oldtown</City>
  </record>
</csv>
```

**Result:**

| equiv="Content-Type" content="text/html; charset=UTF-8"></head>. The head element must already exist or nothing will be added. Any existing meta content-type elements will be removed. |  |
Name;City
John;Newton
Jack;Oldtown

Changelog

Version 8.4

- Added: Serialization parameter binary.
- Updated: New serialization method basex. By default, items of binary type are now output in their native byte representation. The method raw was removed.

Version 8.0

- Added: Support for use-character-maps and parameter-document.
- Added: Serialization method adaptive.
- Updated: adaptive is new default method (before: xml).
- Removed: format, wrap-prefix, wrap-uri.

Version 7.8.2

- Added: limit: Stops serialization after the specified number of bytes has been serialized.

Version 7.8

- Added: csv and json serialization parameters.
- Removed: separator option (use item-separator instead).

Version 7.7.2

- Added: csv serialization method.
- Added: temporary serialization methods csv-header, csv-separator, json-unescape, json-spec, json-format.

Version 7.5

- Added: official item-separator and html-version parameter.
- Updated: method=html5 removed; serializers updated with the latest version of the specification, using method=html and version=5.0.

Version 7.2

- Added: separator parameter.

Version 7.1

- Added: newline parameter.

Version 7.0

- Added: Serialization parameters added to REST API; JSON/JsonML/raw methods.
Chapter 32. XQuery Errors

Read this entry online in the BaseX Wiki.

This article is part of the XQuery Portal. It summarizes the codes of errors that are raised by the standard features and functions of XQuery. As the original specifications are pretty comprehensive, we tried our best to make this overview comprehensible to a wider range of readers.

The following tables list the error codes that are known to BaseX, a short description, and examples of queries raising that errors. Errors that are specific to BaseX can be found in the descriptions of the respective modules.

Original definitions of the error codes are found in the XQuery 3.0, XQuery 3.0 Functions, XQuery 1.0 Update, XQuery 1.0 Full Text, and EXPath HTTP Specifications.

Static Errors

Error Codes: XPST, XQST

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>XPST0003</td>
<td>An error occurred while parsing the query string (i.e., 1+ for i in // * return $i before the query could be compiled and executed). This error is the most common one, and may be accompanied by a variety of different error messages.</td>
<td>1+ for i in //* return $i</td>
</tr>
<tr>
<td>XPST0005</td>
<td>An expression will never return any results, no matter what input is provided.</td>
<td>doc('input')/..</td>
</tr>
<tr>
<td>XPST0008</td>
<td>A variable or type name is used that has not been defined in the current scope.</td>
<td>$a---element(*, x)</td>
</tr>
<tr>
<td>XPST0017</td>
<td>• The specified function is unknown, • it uses the wrong number of arguments, or, when calling Java functions: • there is more than one function with the same number of arguments.</td>
<td>unknown() count(1,2,3)</td>
</tr>
<tr>
<td>XPST0051</td>
<td>An unknown QName is used in a sequence type (e.g. in the target type of the cast expression).</td>
<td>1 instance of x &quot;test&quot; cast as xs:itr</td>
</tr>
<tr>
<td>XPST0080</td>
<td>xs:NOTATION or xs:anyAtomicType is used as castable as xs:NOTATION target type of cast or castable.</td>
<td>xs:NOTATION or xs:anyAtomicType is used as castable as xs:NOTATION target type of cast or castable.</td>
</tr>
<tr>
<td>XPST0081</td>
<td>• A QName uses a prefix that has not been bound to any namespace, or a pragma or option declaration has not been prefixed.</td>
<td>unknown:x (# pragma #) { 1 } namespace, or a pragma or option declaration has not been prefixed.</td>
</tr>
<tr>
<td>XQST0009</td>
<td>The query imports a schema (schema import is not supported by BaseX).</td>
<td>import schema &quot;x&quot;; ()</td>
</tr>
<tr>
<td>XQST0022</td>
<td>Namespace values must be constant strings.</td>
<td>&lt;elem xmlns=&quot;( 'dynamic' )&quot;/&gt;</td>
</tr>
<tr>
<td>XQST0031</td>
<td>The specified XQuery version is not specified.</td>
<td>xquery version &quot;9.9&quot;; ()</td>
</tr>
<tr>
<td>XQST0032</td>
<td>The base URI was declared more than once.</td>
<td>declare base-uri ...</td>
</tr>
<tr>
<td>XQST0033</td>
<td>A namespace prefix was declared more than once.</td>
<td>declare a=&quot;a&quot;; declare a=&quot;b&quot;; ()</td>
</tr>
<tr>
<td>XQST0034</td>
<td>A function was declared more than once.</td>
<td>declare function local:a() ( 1 ); declare function local:a() ( 2 ); local:a()</td>
</tr>
<tr>
<td>XQST0038</td>
<td>The default collation was declared more than once.</td>
<td>declare default collation ...</td>
</tr>
</tbody>
</table>
XQuery Errors

XQST0039 Two or more parameters in a user-defined function have the same name.

```xquery
declare function local:fun($a, $a) { $a * $a };local:fun(1,2)
```

XQDY0040 Two or more attributes in an element have the same node name.

```xml
<elem a="1" a="12"/>
```

XQDY0045 A user-defined function uses a reserved namespace.

```xquery
declare function fn:fun() { 1 };
```

XQST0047 A module was defined more than once.

```xquery
import module ...;
```

XQST0048 A module declaration does not match the namespace of the specified module.

```xquery
import module namespace invalid="uri";
```

XQST0049 A global variable was declared more than once.

```xquery
declare variable $a := 1; declare variable $a := 1; $a
```

XQST0054 A user-defined function has no namespace.

```xquery
declare default function namespace "";
```

XQST0055 A global variable depends on itself. This may be triggered by a circular variable definition.

```xquery
declare variable $a := local:a(); declare function local:a() { $a }; $a
```

XQST0057 The mode for copying namespaces was declared more than once.

```xquery
declare copy-namespaces ...
```

XQST0059 The schema or module with the specified namespace cannot be found or processed.

```xquery
import module "unknown";
```

XQST0060 A user-defined function has no namespace.

```xquery
declare default function name="";
```

XQST0065 The ordering mode was declared more than once.

```xquery
declare ordering ...
```

XQST0066 The default namespace mode for elements or functions was declared more than once.

```xquery
declare default element namespace ...
```

XQST0067 The construction mode was declared more than once.

```xquery
declare construction ...
```

XQST0068 The mode for handling boundary spaces was declared more than once.

```xquery
declare boundary-space ...
```

XQST0069 The default order for empty sequences was declared more than once.

```xquery
declare default order empty ...
```

XQST0070 A namespace declaration overwrites a reserved namespace.

```xquery
<xmlns="uri1" xmlns="uri2"/>
```

XQST0071 A namespace is declared more than once in an element constructor.

```xquery
<a xmlns="uri1" xmlns="uri2"/>
```

XQST0075 The query contains a validate expression (validation is not supported by BaseX).

```xquery
validate strict {}
```

XQST0076 A group by or order by clause specifies an unknown collation.

```xquery
for $i in 1 to 10 order by $i collation "unknown" return $i
```

XQST0079 A pragma was specified without the expression that is to be evaluated.

```xquery
(# xml:a #) {}
```

XQST0085 An empty namespace URI was specified.

```xquery
<pref:elem xmlns:pref=""/>
```

XQST0087 An unknown encoding was specified. Note that the encoding declaration is currently ignored in BaseX.

```xquery
xquery version "1.0" encoding "a b";
```

XQST0088 An empty module namespace was specified.

```xquery
import module "";
```

XQST0089 Two variables in a for or let clause have the same name.

```xquery
for $a at $a in 1 return $i
```

XQST0090 A character reference specifies an invalid character.
XQuery Errors

XQST0093 A module depends on itself. This may be triggered by a `import module` ...
circular module definition.

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Module References</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>XQST0094</td>
<td>group by references a variable that has not been declared before.</td>
<td>for $a in 1 group by $b return $a</td>
</tr>
</tbody>
</table>

XQST0097 A decimal-format property is invalid.

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>XQST0098</td>
<td>A single decimal-format character was assigned to multiple properties.</td>
<td>declare default decimal-format digit = &quot;xxx&quot;; 1</td>
</tr>
</tbody>
</table>

XQST0099 The context item was declared more than once.

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>XQST0106</td>
<td>declare %updating %updating function ...</td>
</tr>
</tbody>
</table>

XQST0108 Output declarations may only be specified in the main module.

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>XQST0109</td>
<td>declare option output:unknown &quot;...&quot;; 1</td>
</tr>
</tbody>
</table>

XQST0110 A serialization parameter was specified more than once in the output declarations.

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>XQST0111</td>
<td>declare decimal-format ...</td>
</tr>
</tbody>
</table>

XQST0113 Context item values may only be in the main module.

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>XQST0114</td>
<td>declare decimal-format EN NaN=&quot;!&quot; NaN=&quot;?&quot;; ()</td>
</tr>
</tbody>
</table>

Type Errors

Error Codes: XPTY, XQTY

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>XPTY0004</td>
<td>This error is raised if an expression has the wrong type, or cannot be cast into the specified type. It may be raised both statically (during query compilation) or dynamically (at runtime).</td>
<td>1 + &quot;A&quot; abs(&quot;a&quot;) 1 cast as xs:gYear</td>
</tr>
</tbody>
</table>

XPTY0018 The result of the last step in a path expression contains both nodes and atomic values.

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>XPTY0018</td>
<td>doc('input.xml')/*, 1</td>
</tr>
</tbody>
</table>

XPTY0019 The result of a step (other than the last step) in a path expression contains an atomic values.

Dynamic Errors

Error Codes: XPDY, XQDY

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>XPDY0002</td>
<td>No value has been defined for an external variable. declare variable $x external; or no context item has been set before the query was $x descendant::* executed.</td>
<td>declare variable $x external;</td>
</tr>
</tbody>
</table>
XQuery Errors

XPDY0050 • The operand type of a treat expression does not "string" treat as xs:int match the type of the argument, or the root of the "string"[/]

XQDY0025 Two or more attributes in a constructed element have the same node name.

XQDY0026 The content of a computed processing instruction contains "?x".

XQDY0041 The name of a processing instruction is invalid.

XQDY0044 The node name of an attribute uses reserved prefixes or namespaces.

XQDY0064 The name of a processing instruction equals "XML" (case insensitive).

XQDY0072 The content of a computed comment contains "--" or ends with "."

XQDY0074 The name of a computed attribute or element is invalid, or uses an unbound prefix.

XQDY0095 A sequence with more than one item was bound to a group by clause.

XQDY0096 The node name of an element uses reserved prefixes or namespaces.

XQDY0101 Invalid namespace declaration.

XQDY0102 Duplicate namespace declaration.

Functions Errors

Error Codes: FOAR, FOCA, FOCH, FODC, FODF, FODT, FOER, FOFD, FONS, FORG, FORX, FOTY, FOUT

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOAR0001</td>
<td>A value was divided by zero.</td>
<td>1 div 0</td>
</tr>
<tr>
<td>FOAR0002</td>
<td>A numeric declaration or operation causes an over- or underflow.</td>
<td>12345678901234567890 xs:double(&quot;-INF&quot;) idiv 1</td>
</tr>
<tr>
<td>FOCA0002</td>
<td>A float number cannot be converted to a decimal or integer value, or a function argument cannot be converted to a valid QName.</td>
<td>xs:int(xs:double(&quot;INF&quot;)) QName(&quot;&quot;, &quot;el em&quot;)</td>
</tr>
<tr>
<td>FOCA0003</td>
<td>A value is too large to be represented as integer.</td>
<td>xs:integer(99e100)</td>
</tr>
<tr>
<td>FOCA0005</td>
<td>&quot;NaN&quot; is supplied to duration operations.</td>
<td>xs:yearMonthDuration(&quot;PLY&quot;) * xs:double(&quot;NaN&quot;)</td>
</tr>
<tr>
<td>FOCH0001</td>
<td>A codepoint was specified that does not represent a valid XML character.</td>
<td>codepoints-to-string(0)</td>
</tr>
<tr>
<td>FOCH0002</td>
<td>An unsupported collation was specified in a function.</td>
<td>compare('a', 'a', 'unknown')</td>
</tr>
<tr>
<td>FOCH0003</td>
<td>An unsupported normalization form was specified in a function.</td>
<td>normalize-unicode('a', 'unknown')</td>
</tr>
<tr>
<td>FODC0001</td>
<td>The argument specified in fn:id() or fn:idref() must have a document node as root.</td>
<td>id(&quot;id0&quot;, &lt;xml/&gt;)</td>
</tr>
<tr>
<td>FODC0002</td>
<td>The specified document resource cannot be retrieved.</td>
<td>doc(&quot;unknown.xml&quot;)</td>
</tr>
<tr>
<td>FODC0004</td>
<td>The specified collection cannot be retrieved.</td>
<td>collection(&quot;unknown&quot;)</td>
</tr>
<tr>
<td>Error Code</td>
<td>Message</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>FODC0005</td>
<td>The specified URI to a document resource is invalid. doc(&quot;&lt;xml/&gt;&quot;)</td>
<td></td>
</tr>
<tr>
<td>FODC0006</td>
<td>The string passed to fn:parse-xml() is not well-formed. parse-xml(&quot;&lt;x/&gt;&quot;)</td>
<td></td>
</tr>
<tr>
<td>FODC0007</td>
<td>The base URI passed to fn:parse-xml() is invalid. parse-xml(&quot;&lt;x/&gt;&quot;, &quot;:&quot;)</td>
<td></td>
</tr>
<tr>
<td>FODF1280</td>
<td>The name of the decimal format passed to format-number(1, &quot;O&quot;, invalid) is invalid.</td>
<td></td>
</tr>
<tr>
<td>FODF1310</td>
<td>The picture string passed to fn:format-number(1, invalid) is invalid.</td>
<td></td>
</tr>
<tr>
<td>FODT0001</td>
<td>An arithmetic duration operation causes an over- or underflow. xs:date('2000-01-01') + xs:duration('P99999Y')</td>
<td></td>
</tr>
<tr>
<td>FODT0002</td>
<td>A duration declaration or operation causes an over- or implicit-timezone() div 0 underflow.</td>
<td></td>
</tr>
<tr>
<td>FODT0003</td>
<td>An invalid timezone was specified. adjust-time-to-timezone(xs:time(&quot;01:01:01&quot;), xs:dayTimeDuration(&quot;PT20H&quot;))</td>
<td></td>
</tr>
<tr>
<td>FOER0000</td>
<td>Error triggered by the fn:error() function. error()</td>
<td></td>
</tr>
<tr>
<td>FOFD1340</td>
<td>The picture string passed to fn:format-date(), fn:format-time() or fn:format-datetime() is invalid. format-date(current-date(), &quot;) format-time(current-time(), &quot;[Y2&quot;] format-datetime(current-date(), &quot;[Y2&quot;]</td>
<td></td>
</tr>
<tr>
<td>FONS0004</td>
<td>A function has a QName as argument that specifies an unbound prefix. resolve-QName(&quot;x:e&quot;, &lt;e/&gt;)</td>
<td></td>
</tr>
<tr>
<td>FORG0001</td>
<td>A value cannot be cast to the required target type. xs:integer(&quot;A&quot;) 1 + &lt;x&gt;a&lt;/x&gt;</td>
<td></td>
</tr>
<tr>
<td>FORG0002</td>
<td>The URI passed to fn:resolve-URI() is invalid. resolve-URI(&quot;:&quot;))</td>
<td></td>
</tr>
<tr>
<td>FORG0003</td>
<td>fn:zero-or-one() was called with more than one item. zero-or-one((1, 2))</td>
<td></td>
</tr>
<tr>
<td>FORG0004</td>
<td>fn:one-or-more() was called with zero items. one-or-more(()))</td>
<td></td>
</tr>
<tr>
<td>FORG0005</td>
<td>fn:exactly-one() was called with zero or more than one item. exactly-one((1, 2))</td>
<td></td>
</tr>
<tr>
<td>FORG0006</td>
<td>A wrong argument type was specified in a function call. sum((1, &quot;string&quot;))</td>
<td></td>
</tr>
<tr>
<td>FORG0008</td>
<td>The arguments passed to fn:dateTime() have different timezones. dateTime(xs:date(&quot;2001-01-01+01:01&quot;), current-time())</td>
<td></td>
</tr>
<tr>
<td>FORX0001</td>
<td>A function specifies an invalid regular expression flag. matches('input', 'query', 'invalid')</td>
<td></td>
</tr>
<tr>
<td>FORX0002</td>
<td>A function specifies an invalid regular expression. matches('input', &quot;&quot;]&quot;)</td>
<td></td>
</tr>
<tr>
<td>FORX0003</td>
<td>A regular expression matches an empty string. tokenize('input', &quot;.?&quot;)</td>
<td></td>
</tr>
<tr>
<td>FORX0004</td>
<td>The replacement string of a regular expression is invalid. replace(&quot;input&quot;, &quot;match&quot;, &quot;)</td>
<td></td>
</tr>
<tr>
<td>FOTY0012</td>
<td>An item has no typed value. count#1</td>
<td></td>
</tr>
<tr>
<td>FOTY0013</td>
<td>Functions items cannot be atomized, have no defined equality, and have no string representation. data(false#0)</td>
<td></td>
</tr>
<tr>
<td>FOTY0014</td>
<td>Function items have no string representation. string(map())</td>
<td></td>
</tr>
<tr>
<td>FOTY0015</td>
<td>Function items cannot be compared. deep-equal(false#0, true#0)</td>
<td></td>
</tr>
</tbody>
</table>
XQuery Errors

**Serialization Errors**

Error Codes: SEPM, SERE, SESU

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>SESU0007</td>
<td>The specified encoding is not supported.</td>
<td>declare option output:encoding &quot;xyz&quot;; 1</td>
</tr>
<tr>
<td>SEPM0009</td>
<td><code>omit-xml-declaration</code> is set to yes, and <code>standalone</code> has a value other than <code>omit</code>.</td>
<td></td>
</tr>
<tr>
<td>SEPM0010</td>
<td><code>method</code> is set to <code>xml</code>, <code>undeclare-prefixes</code> is set to yes, and <code>version</code> is set to 1.0.</td>
<td></td>
</tr>
<tr>
<td>SERE0014</td>
<td><code>method</code> is set to <code>html</code>, and an invalid HTML character is found.</td>
<td></td>
</tr>
<tr>
<td>SERE0015</td>
<td><code>method</code> is set to <code>html</code>, and a closing bracket (&gt;) appears inside a processing instruction.</td>
<td></td>
</tr>
<tr>
<td>SEPM0016</td>
<td>A specified parameter is unknown or has an invalid value.</td>
<td>declare option output:indent &quot;nope&quot;; 1</td>
</tr>
</tbody>
</table>

**Update Errors**

Error Codes: FOUP, XUDY, XUST, XUTY

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOUP0001</td>
<td>The first argument of <code>fn:put()</code> must be a document node or element.</td>
<td><code>fn:put(text { 1 }, 'file.txt')</code></td>
</tr>
<tr>
<td>FOUP0002</td>
<td>The second argument of <code>fn:put()</code> is not a valid URI.</td>
<td><code>fn:put(&lt;a/&gt;, '//')</code></td>
</tr>
<tr>
<td>XUDY0009</td>
<td>The target node of a replace expression needs a parent in order to be replaced.</td>
<td>replace node <code>&lt;target/&gt;</code> with <code>&lt;new/&gt;</code></td>
</tr>
<tr>
<td>XUDY0014</td>
<td>The expression updated by the <code>modify</code> clause was not created by the <code>copy</code> clause.</td>
<td>let $a := doc('a') return copy $b := $a modify delete node $a/* return $b</td>
</tr>
<tr>
<td>XUDY0015</td>
<td>In a rename expression, a target is renamed more than once.</td>
<td>let $a := &lt;xml/&gt; return (rename node $a as 'a', rename node $a as 'b')</td>
</tr>
<tr>
<td>XUDY0016</td>
<td>In a replace expression, a target is replaced more than once.</td>
<td>let $a := &lt;x&gt;&lt;x&gt;/node() return (replace node $a with &lt;a/&gt;, replace node $a with &lt;b/&gt;)</td>
</tr>
<tr>
<td>XUDY0017</td>
<td>In a replace value of expression, a target is replaced more than once.</td>
<td>let $a := &lt;x/&gt; return (replace value of node $a with 'a', replace value of node $a with 'a')</td>
</tr>
<tr>
<td>XUDY0021</td>
<td>The resulting update expression contains duplicate attributes.</td>
<td>copy $c := &lt;x a='a'/&gt; modify insert node attribute a (&quot;&quot;&quot;) into $c return $c</td>
</tr>
</tbody>
</table>

154
XQuery Errors

XUDY0023 The resulting update expression conflicts with existing namespaces.

```
rename node <a:ns xmlns:a='uri'/>
  as QName('URI', 'a:ns')
```

XUDY0024 New namespaces conflict with each other.

```
copy $n := <x/>
modify (insert node attribute { QName('uri1', 'a') } { '' } into $n, insert node attribute { QName('uri2', 'a') } { '' } into $n) return $n
```

XUDY0027 Target of an update expression is an empty sequence.

```
insert node <x/> into ()
```

XUDY0029 The target of an update expression has no parent node.

```
insert node <new/> before <target/>
```

XUDY0030 Attributes cannot be inserted before or after the child of a document node.

```
insert node <e a='a'/>/@a after document { <e/> }/*
```

XUDY0031 Multiple calls to fn:put() address the same URI.

```
for $i in 1 to 3 return put(<a/>, 'file.txt')
```

XUST0001 No updating expression is allowed here.

```
delete node /*, "finished."
```

XUST0002 An updating expression is expected in the modify clause or an updating function.

```
copy $a := <x/>
modify 1
return $a
```

XUST0003 The revalidation mode was declared more than once.

```
declare revalidation ...
```

XUST0026 The query contains a revalidate expression (revalidation is not supported by BaseX).

```
declare revalidation ...
```

XUST0028 no return type may be specified in an updating function.

```
declare updating function
local::x() as item() { () ; () }
```

XUTY0004 New attributes to be inserted must directly follow the root node.

```
insert node (<a/>, attribute a {''}) into <a/>
```

XUTY0005 A single element or document node is expected as target of an insert expression.

```
insert node <new/> into attribute a { '' }
```

XUTY0006 A single element, text, comment or processing instruction is expected as target of an insert attribute a { '' } before/after expression.

```
insert node <new/> after attribute a { '' }
```

XUTY0007 Only nodes can be deleted.

```
delete node "string"
```

XUTY0008 A single element, text, attribute, comment or processing instruction is expected as target of a replace > ) with <b/> expression.

```
replace node document { <a/> } with <b/>
```

XUTY0010 In a replace expression, in which no attributes are targeted, the replacing nodes must not be attributes as with attribute size { 1 } well.

```
replace node <a><b/></a>/b
```

XUTY0011 In the replace expression, in which attributes are targeted, the replacing nodes must be attributes as well. with <a/>

```
replace node <e a='''/>/@a
```

XUTY0012 In a rename expression, the target nodes must be an element, attribute or processing instruction.

```
rename node text { 1 } as <x/>
```

XUTY0013 An expression in the copy clause must return a single node.

```
copy $c := (<a/>, <b/>) modify () return $c
```

XUTY0022 An attribute must not be inserted into a document node.

```
insert node <e a='''/>/@a into document { 'a' }
```
Full-Text Errors

Error Codes: FTDY, FTST

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>FTDY0016</td>
<td>The specified weight value is out of range.</td>
<td>'a' contains text 'a' weight { 1001 }</td>
</tr>
<tr>
<td>FTDY0017</td>
<td>The not in operator contains a <em>string exclude</em>.</td>
<td>'a' contains text 'a' not in (ftnot 'a')</td>
</tr>
<tr>
<td>FTDY0020</td>
<td>The search term uses an invalid wildcard syntax.</td>
<td>'a' contains text '.{}' using wildcards</td>
</tr>
<tr>
<td>FTST0007</td>
<td>The full-text expression contains an ignore option (the ignore option is not supported by BaseX).</td>
<td>'a' contains text 'a' without content 'x'</td>
</tr>
<tr>
<td>FTST0008</td>
<td>The specified stop word file could not be opened or processed.</td>
<td>'a' contains text 'a' using stop words at 'unknown.txt'</td>
</tr>
<tr>
<td>FTST0009</td>
<td>The specified language is not supported.</td>
<td>'a' contains text 'a' using language 'aaa'</td>
</tr>
<tr>
<td>FTST0018</td>
<td>The specified thesaurus file could not be opened or processed.</td>
<td>'a' contains text 'a' using thesaurus at 'aaa'</td>
</tr>
<tr>
<td>FTST0019</td>
<td>A match option was specified more than once.</td>
<td>'a' contains text 'a' using stemming using stemming</td>
</tr>
</tbody>
</table>

BaseX Errors

Error Codes: BASX

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>BASX0000</td>
<td>Generic error, which is used for exceptions in context-aware Java bindings.</td>
<td>file:delete('file.txt'): Create rights needed.</td>
</tr>
<tr>
<td>BASX0001</td>
<td>The current user has insufficient permissions to execute an expression.</td>
<td>declare option db:xyz &quot;no&quot;; l</td>
</tr>
<tr>
<td>BASX0002</td>
<td>The specified database option is unknown.</td>
<td>%restxq:GET('x')</td>
</tr>
<tr>
<td>BASX0003</td>
<td>Errors related to RESTXQ.</td>
<td></td>
</tr>
</tbody>
</table>

Additional, module-specific error codes are listed in the descriptions of the query modules.
Part VI. XQuery Modules
Chapter 33. Admin Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions for performing admin-centric operations such as managing database users and log data.

Conventions

All functions in this module are assigned to the http://basex.org/modules/admin namespace, which is statically bound to the admin prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

Functions

admin:sessions

Signatures

admin:sessions() as element(session)*

Summary

Returns an element sequence with all currently opened sessions, including the user name, address (IP:port) and an optionally opened database. The output of this function and the SHOW SESSIONS command is similar.

Examples

• admin:sessions() may e.g. return
  `<session user="admin" address="127.0.0.1:6286" database="factbook"/>

admin:logs

Signatures

admin:logs() as element(file)*
admin:logs($date as xs:string) as element(entry)*
admin:logs($date as xs:string, $merge as xs:boolean) as element(entry)*

Summary

Returns Logging data compiled by the database or HTTP server:

• If no argument is specified, a list of all log files will be returned, including the file size and date.

• If a $date is specified, the contents of a single log file will be returned.

• If $merge is set to true, related log entries will be merged. Please note that the merge might not be 100% successful, as log entries may be ambiguous.

Examples

• admin:logs() may return `<file size="834367"/>2015-01-23</file>` if a single log file exists.

• admin:logs() ! admin:logs(.) lists the contents of all log files.

admin:write-log

Signatures

admin:write-log($text as xs:string) as empty-sequence()
admin:write-log($text as xs:string, $type as xs:string) as empty-sequence()

Summary

Writes a string to the database logs, along with current user data (timestamp, user name). An optional log $type can be specified, which must consist of letters in upper case. If omitted, the log type is INFO. If the function is called from a database client, the IP will be logged. Otherwise, the string SERVER will be logged.

Errors

BXAD0003: Log type must consist of uppercase letters.
admin:delete-logs

Signatures
admin:delete-logs($date as xs:string) as empty-sequence()

Summary
Deletes the log entries from the specified $date

Errors
BXAD0001: Today's log file cannot be deleted.
BXAD0002: An error occurred while deleting a log file.

Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXAD0001</td>
<td>Today's log file cannot be deleted.</td>
</tr>
<tr>
<td>BXAD0002</td>
<td>An error occurred while deleting a log file.</td>
</tr>
<tr>
<td>BXAD0003</td>
<td>Log type must consist of uppercase letters.</td>
</tr>
</tbody>
</table>

Changelog

Version 8.3
- Updated: admin:write-log: optional log type added

Version 8.2
- Added: admin:delete-logs

Version 8.0
- Added: admin:write-log
- Deleted: admin:users (renamed to user:list-details)

Version 7.8.2
- Updated: admin:users: md5-encoded password added to output.
- Updated: admin:logs: represent name of log files as string value; $merge argument added.

The Module was introduced with Version 7.5.
Chapter 34. Archive Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions to handle archives (including ePub, Open Office, JAR, and many other formats). New ZIP and GZIP archives can be created, existing archives can be updated, and the archive entries can be listed and extracted. The archive:extract-binary function includes an example for writing the contents of an archive to disk.

Conventions

All functions in this module are assigned to the http://basex.org/modules/archive namespace, which is statically bound to the archive prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

Functions

archive:create

Signatures
archive:create($entries as item(), $contents as item()*) as xs:base64Binary archive:create($entries as item(), $contents as item()*, $options as map(xs:string, xs:string)) as xs:base64Binary

Summary
Creates a new archive from the specified entries and contents. The $entries argument contains meta information required to create new entries. All items may either be of type xs:string, representing the entry name, or element (archive:entry), containing the name as text node and additional, optional attributes:

- last-modified: timestamp, specified as xs:dateTime (default: current time)
- compression-level: 0-9, 0 = uncompressed (default: 8)
- encoding: for textual entries (default: UTF-8)

An example:

<archive:entry last-modified='2011-11-11T11:11:11'
  compression-level='8'
  encoding='US-ASCII'>hello.txt</archive:entry>

The actual $contents must be xs:string or xs:base64Binary items. The $options parameter contains archiving options:

- format: allowed values are zip and gzip. zip is the default.
- algorithm: allowed values are deflate and stored (for the zip format). deflate is the default.

Errors
ARCH0001: the number of entries and contents differs.ARCH0002: the specified option or its value is invalid or not supported.ARCH0003: entry descriptors contain invalid entry names, timestamps or compression levels.ARCH0004: the specified encoding is invalid or not supported, or the string conversion failed. Invalid XML characters will be ignored if the CHECKSTRINGS option is turned off.ARCH0005: the chosen archive format only allows single entries.ARCH9999: archive creation failed for some other reason.

Examples
The following one-liner creates an archive archive.zip with one file file.txt:
archive:create(<archive:entry>file.txt</archive:entry>, 'Hello World')

The following function creates an archive mp3.zip, which contains all MP3 files of a local directory:

```xml
let $path  := 'audio/
let $files := file:list($path, true(), '*.mp3')
let $zip   := archive:create(
    $files ! element archive:entry { . },
    $files ! file:read-binary($path || .))
return file:write-binary('mp3.zip', $zip)
```

archive:create-from

<table>
<thead>
<tr>
<th>Signatures</th>
<th>archive:create-from($path as xs:string) as xs:base64Binary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>archive:create-from($path as xs:string, $options as map(xs:string, xs:string)) as xs:base64Binary</td>
</tr>
<tr>
<td></td>
<td>archive:create-from($path as xs:string, $options as map(xs:string, xs:string), $entries as item()*) as xs:base64Binary</td>
</tr>
</tbody>
</table>

| Summary | This convenience function creates an archive from all files in the specified directory $path. The $options parameter contains archiving options, and the files to be archived can be limited via $entries. The format of the two last arguments is the same as for archive:create. |
| Errors | no-dir: the specified path does not point to a directory. is-dir: one of the specified entries points to a directory. not-found: a specified entry does not exist. ARCH9999: archive creation failed for some other reason. |
| Examples | This example writes the files of a user's home directory to archive.zip: |

```xml
let $zip := archive:create-from('/home/user/')
return file:write-binary('archive.zip', $zip)
```

archive:entries

| Signatures | archive:entries($archive as xs:base64Binary) as element(archive:entry)* |
| Summary | Returns the entry descriptors of the specified $archive. A descriptor contains the following attributes, provided that they are available in the archive format: |
|           | • size: original file size |
|           | • last-modified: timestamp, formatted as xs:dateTime |
|           | • compressed-size: compressed file size |
| Errors | ARCH9999: archive creation failed for some other reason. |
| Examples | Sums up the file sizes of all entries of a JAR file: |

```xml
sum(archive:entries(file:read-binary('zip.zip'))/@size)
```
### archive:options

**Signatures**

<table>
<thead>
<tr>
<th>Signature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>archive:options($archive as xs:base64Binary) as element(archive:options)</code></td>
<td>Returns the options of the specified $archive in the format specified by <code>archive:create</code>.</td>
</tr>
</tbody>
</table>

**Errors**

ARCH0002: The packing format is not supported.
ARCH9999: archive creation failed for some other reason.

**Examples**

A standard ZIP archive will return the following options:

```xml
<archive:options xmlns:archive="http://basex.org/modules/archive">
  <archive:format value="zip"/>
  <archive:algorithm value="deflate"/>
</archive:options>
```

### archive:extract-text

**Signatures**

<table>
<thead>
<tr>
<th>Signature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>archive:extract-text($archive as xs:base64Binary) as xs:string*</code></td>
<td>Extracts entries of the specified $archive and returns them as texts. The returned entries can be limited via $entries. The format of the argument is the same as for <code>archive:create</code> (attributes will be ignored). The encoding of the input files can be specified via $encoding.</td>
</tr>
</tbody>
</table>

**Errors**

ARCH0004: the specified encoding is invalid or not supported, or the string conversion failed. Invalid XML characters will be ignored if the CHECKSTRINGS option is turned off.
ARCH9999: archive creation failed for some other reason.

**Examples**

The following expression extracts all .txt files from an archive:

```xquery
let $archive := file:read-binary("documents.zip")
for $entry in archive:entries($archive)[ends-with(., '.txt')]
return archive:extract-text($archive, $entry)
```

### archive:extract-binary

**Signatures**

<table>
<thead>
<tr>
<th>Signature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>archive:extract-binary($archive as xs:base64Binary)*</code></td>
<td>Extracts entries of the specified $archive and returns them as binaries. The returned entries can be limited via $entries. The format of the argument is the same as for <code>archive:create</code> (attributes will be ignored).</td>
</tr>
</tbody>
</table>

**Errors**

ARCH9999: archive creation failed for some other reason.

**Examples**

This example unzips all files of an archive to the current directory:

```xquery
let $archive := file:read-binary('archive.zip')
let $entries := archive:entries($archive)
let $contents := archive:extract-binary($archive)
return for-each-pair($entries, $contents, function($entry, $content) {
  file:create-dir(replace($entry, "[^/]+$", "")
  file:write-binary($entry, $content)
})
```
Archive Module

archive:extract-to

Signatures
archive:extract-to($path as xs:string, $archive as xs:base64Binary)
   as empty-sequence() archive:extract-to($path as xs:string, $archive as xs:base64Binary, $entries as item()*)
   as empty-sequence()

Summary
This convenience function writes files of an $archive directly to the specified directory $path. The archive entries to be written can be restricted via $entries. The format of the argument is the same as for archive:create (attributes will be ignored).

Errors
ARCH9999: archive creation failed for some other reason.

Examples
The following expression unzips all files of an archive to the current directory:

```
archive:extract-to('.', file:read-binary('archive.zip'))
```

archive:update

Signatures
archive:update($archive as xs:base64Binary, $entries as item()*,$contents as item()) as xs:base64Binary

Summary
Creates an updated version of the specified $archive with new or replaced entries. The format of $entries and $contents is the same as for archive:create.

Errors
ARCH0001: the number of entries and contents differs. ARCH0003: entry descriptors contain invalid entry names, timestamps, compression levels or encodings. ARCH0004: the specified encoding is invalid or not supported, or the string conversion failed. Invalid XML characters will be ignored if the CHECKSTRINGS option is turned off. ARCH0005: the entries of the given archive cannot be modified. ARCH9999: archive creation failed for some other reason.

Examples
This example replaces texts in a Word document:

```
declare variable $input := "HelloWorld.docx";
declare variable $output := "HelloUniverse.docx";
declare variable $doc := "word/document.xml";

let $archive := file:read-binary($input)
let $entry :=
   copy $c := fn:parse-xml(archive:extract-text($archive, $doc))
   modify replace value of node $c//*[text() = "HELLO WORLD!"] with
   "HELLO UNIVERSE!"
   return fn:serialize($c)
let $updated := archive:update($archive, $doc, $entry)
return file:write-binary($output, $updated)
```

archive:delete

Signatures
archive:delete($archive as xs:base64Binary, $entries as item()* as xs:base64Binary

Summary
Deletes entries from an $archive. The format of $entries is the same as for archive:create.

Errors
ARCH0005: the entries of the given archive cannot be modified. ARCH9999: archive creation failed for some other reason.

Examples
This example deletes all HTML files in an archive and creates a new file:

```
let $zip := file:read-binary('old.zip')
let $entries := archive:entries($zip)[matches(., '\.x?html?$', 'i')]
return file:write-binary('new.zip', archive:delete($zip, $entries))
```
Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARCH0001</td>
<td>The number of specified entries and contents differs.</td>
</tr>
<tr>
<td>ARCH0002</td>
<td>The packing format or the specified option is invalid or not supported.</td>
</tr>
<tr>
<td>ARCH0003</td>
<td>Entry descriptors contain invalid entry names, timestamps or compression</td>
</tr>
<tr>
<td></td>
<td>levels.</td>
</tr>
<tr>
<td>ARCH0004</td>
<td>The specified encoding is invalid or not supported, or the string</td>
</tr>
<tr>
<td></td>
<td>conversion failed. Invalid XML characters will be ignored if the</td>
</tr>
<tr>
<td></td>
<td>CHECKSTRINGS option is turned off.</td>
</tr>
<tr>
<td>ARCH0005</td>
<td>The entries of the given archive cannot be modified.</td>
</tr>
<tr>
<td>ARCH0006</td>
<td>The chosen archive format only allows single entries.</td>
</tr>
<tr>
<td>ARCH9999</td>
<td>Archive processing failed for some other reason.</td>
</tr>
</tbody>
</table>

Changelog

Version 8.3

- Added: archive:create-from, archive:extract-to (replaces archive:write)

Version 7.7

- Added: archive:write

The module was introduced with Version 7.3.
Chapter 35. Array Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions for manipulating arrays, which will officially be introduced with XQuery 3.1. Please note that the functions are subject to change until the specification has reached its final stage.

Conventions

All functions in this module are assigned to the http://www.w3.org/2005/xpath-functions/array namespace, which is statically bound to the array prefix.

Functions

array:size

Signatures
array:size($input as array(*)) as xs:integer

Summary
Returns the number of members in $array. Note that because an array is an item, the fn:count function when applied to an array always returns 1.

Examples
• array:size(array ( 1 to 10 )) returns 10.
• array:size([1 to 10]) returns 1, because the array contains a single sequence with 10 integers.

array:get

Signatures
array:get($array as array(*), $position as xs:integer) as item()*

Summary
Returns the $array member at the specified $position.

Errors
FOAY0001: $position is not in the range 1 to array:size($array) inclusive.

Examples
• array:get(array ( reverse(1 to 5) ), 5) returns the value 1.

array:append

Signatures
array:append($array as array(*), $member as item()) as array(*)

Summary
Returns a copy of $array with a new $member attached.

Examples
• array:append([], 'member1') returns the array ["member1"].

array:subarray

Signatures
array:subarray($array as array(*), $position as xs:integer) as array(*)
array:subarray($array as array(*), $position as xs:integer, $length as xs:integer) as array(*)

Summary
Constructs a new array with with $length members of $array beginning from the specified $position. The two-argument version of the function returns the same result as the three-argument version when called with $length equal to the value of array:size($array) - $position + 1.

Errors
FOAY0001: $position is less than one, or if $position + $length is greater than array:size($array) + 1.FOAY0002: $length is less than zero.

Examples
• array:append(['member1', 'member2']) returns the array ["member1", "member2"].
array:remove

**Signatures**

```
array:remove($array as array(*), $position as xs:integer) as array(*)
```

**Summary**

Returns a copy of $array without the member at the specified $position.

**Errors**

FOAY0001: $position is not in the range 1 to array:size($array) inclusive.

**Examples**

• array:append(["a"], 1) returns the array [].

array:insert-before

**Signatures**

```
array:insert-before($array as array(*), $position as xs:integer, $member as item()*) as array(*)
```

**Summary**

Returns a copy of $array with one new $member at the specified $position. Setting $position to the value array:size($array) + 1 yields the same result as array:append($array, $insert).

**Errors**

FOAY0001: $position is not in the range 1 to array:size($array) + 1 inclusive.

**Examples**

• array:insert-before(["a"], 1, "b") returns the array ["b", "a"].

array:head

**Signatures**

```
array:head($array as array(*)) as item()*
```

**Summary**

Returns the first member of $array. This function is equivalent to the expression $array(1).

**Errors**

FOAY0001: The array is empty.

**Examples**

• array:head(["a", "b"]) returns "a".

  • array:head([ ["a", "b"], ["c", "d"] ]) returns the array ["a", "b"].

array:tail

**Signatures**

```
array:tail($array as array(*)) as array(*)
```

**Summary**

Returns a new array with all members except the first from $array. This function is equivalent to the expression array:remove($array, 1).

**Errors**

FOAY0001: The array is empty.

**Examples**

• array:insert-before(["a"], 1, "b") returns the array ["b", "a"].

array:reverse

**Signatures**

```
array:reverse($array as array(*)) as array(*)
```

**Summary**

Returns a new array with all members of $array in reverse order.

**Examples**

• array:reverse(array { 1 to 3 }) returns the array [3, 2, 1].

array:join

**Signatures**

```
array:join($arrays as array(*)*) as array(*)
```

**Summary**

Concatenates the contents of several $arrays into a single array.

**Examples**

• array:join() returns the array [].

  • array:join((1 to 3) ! array { . }) returns the array [1, 2, 3].
### array:flatten

**Signatures**

`array:flatten($items as item()*) as item(*)`

**Summary**

Recursively flattens all arrays that occur in the supplied `$items`.

**Examples**

- `array:flatten(["a","b")`) returns the sequence "a", "b".
- `array:flatten([1,[2,3],4]])` returns the sequence 1, 2, 3, 4.

### array:for-each

**Signatures**

`array:for-each($array as array(*), $function as function(item()*) as item()*) as array(*)`

**Summary**

Returns a new array, in which each member is computed by applying `$function` to the corresponding member of `$array`.

**Examples**

The following query returns the array [2, 3, 4, 5, 6]:

```
array:for-each(
    array { 1 to 5 },
    function($i) { $i + 1}
)
```

### array:filter

**Signatures**

`array:filter($array as array(*), $function as function(item()*) as xs:boolean) as array(*)`

**Summary**

Returns a new array with those members of `$array` for which `$function` returns `true`.

**Examples**

The following query returns the array [0, 1, 3]:

```
array:filter(
    array { 0, 1, -2, 3, -4 },
    function($i) { $i > 0 }
)
```

### array:fold-left

**Signatures**

`array:fold-left($array as array(*), $zero as item()*, $function as function(item()*, item()*) as item()*) as item(*)`

**Summary**

Evaluates the supplied `$function` cumulatively on successive members of the supplied `$array` from left to right and using `$zero` as first argument.

**Examples**

The following query returns 55 (the sum of the integers 1 to 10):

```
array:fold-left(
    array { 1 to 10 },
    0,
    function($a, $b) { $a + $b }
)
```

### array:fold-right

**Signatures**

`array:fold-right($array as array(*), $zero as item()*, $function as function(item()*, item()*) as item()*) as item(*)`

**Summary**

Evaluates the supplied `$function` cumulatively on successive members of the supplied `$array` from right to left and using `$zero` as first argument.
Examples

The following query is equivalent to the expression array:reverse(array { 1 to 5 }):

```xml
array {
    array:fold-right(
        array { 1 to 5 },
        (),
        function($a, $b) { $b, $a }
    )
}
```

array:for-each-pair

Signatures

array:for-each-pair($array1 as array(*), $array2 as array(*), $function as function(item(*) as item(*)) as array(*))

Summary

Returns a new array obtained by evaluating the supplied $function for each pair of members at the same position in $array1 and $array2.

Examples

The following query returns the array [5, 7, 9]:

```xml
array:for-each-pair(
    array { 1 to 3 },
    array { 4 to 6 },
    function($a + $b) { $a + $b }
)
```

array:sort

Signatures

array:sort($array as array(*)) as array(*)

array:sort($array as array(*), $key as function(item(*) as xs:anyAtomicType*) as array(*))

Summary

Returns a new array with sorted $array members. If a sort $key function is given, it will be applied on all array members. The items of the resulting values will be sorted using the semantics of the lt expression.

Examples

• array:sort(array { reverse(1 to 3) }) returns [1, 2, 3]
• array:sort([3, -2, 1], abs#1) returns [1, -2, 3]
• array:sort([1,2,3], function($x) { -$x }) returns [3, 2, 1]
• array:sort((1, 'a')) returns an error (strings and integers cannot be compared)

Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOAY0001</td>
<td>The specified index extends beyonds the bounds of an array.</td>
</tr>
<tr>
<td>FOAY0002</td>
<td>The specified length is less than zero.</td>
</tr>
</tbody>
</table>

Changelog

Version 8.4

• Removed: array:serialize (use fn:serialize instead)

Introduced with Version 8.0.
Chapter 36. Binary Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions to process binary data, including extracting subparts, searching, basic binary operations and conversion between binary and structured forms.

This module is based on the EXPath Binary Module.

Conventions

All functions and errors in this module are assigned to the http://expath.org/ns/binary namespace, which is statically bound to the bin prefix.

Constants and Conversions

**bin:hex**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>bin:hex($in as xs:string?) as xs:base64Binary?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the binary form of the set of octets written as a sequence of (ASCII) hex digits ([0-9A-Fa-f]). $in will be effectively zero-padded from the left to generate an integral number of octets, i.e. an even number of hexadecimal digits. If $in is an empty string, then the result will be an xs:base64Binary with no embedded data. Byte order in the result follows (per-octet) character order in the string. If the value of $in is the empty sequence, the function returns an empty sequence.</td>
</tr>
<tr>
<td>Errors</td>
<td>non-numeric-character: the input cannot be parsed as a hexadecimal number.</td>
</tr>
<tr>
<td>Examples</td>
<td>bin:hex('11223F4E') yields ESI/Tg==. xs:hexBinary(bin:hex('FF')) yields FF.</td>
</tr>
</tbody>
</table>

**bin:bin**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>bin:bin($in as xs:string?) as xs:base64Binary?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the binary form of the set of octets written as a sequence of (8-wise) (ASCII) binary digits ([01]). $in will be effectively zero-padded from the left to generate an integral number of octets. If $in is an empty string, then the result will be an xs:base64Binary with no embedded data. Byte order in the result follows (per-octet) character order in the string. If the value of $in is the empty sequence, the function returns an empty sequence.</td>
</tr>
<tr>
<td>Errors</td>
<td>non-numeric-character: the input cannot be parsed as a binary number.</td>
</tr>
<tr>
<td>Examples</td>
<td>bin:bin('1101000111010101') yields 0dU=. xs:hexBinary(bin:bin('1000111010101')) yields 11D5.</td>
</tr>
</tbody>
</table>

**bin:octal**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>bin:octal($in as xs:string?) as xs:base64Binary?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the binary form of the set of octets written as a sequence of (ASCII) octal digits ([0-7]). $in will be effectively zero-padded from the left to generate an integral number of octets. If $in is an empty string, then the result will be an xs:base64Binary with no embedded data. Byte order in the result follows (per-octet) character order in the string. If the value of $in is the empty sequence, the function returns an empty sequence.</td>
</tr>
<tr>
<td>Errors</td>
<td>non-numeric-character: the input cannot be parsed as an octal number.</td>
</tr>
<tr>
<td>Examples</td>
<td>xs:hexBinary(bin:octal('11223047')) yields 252627.</td>
</tr>
</tbody>
</table>
## Binary Module

### bin:to-octets

**Signatures**

\[
\text{bin:to-octets}(\text{\$in as xs:base64Binary}) \rightarrow \text{xs:integer}^* 
\]

**Summary**

Returns binary data as a sequence of octets. If \$in is a zero length binary data then the empty sequence is returned. Octets are returned as integers from 0 to 255.

### bin:from-octets

**Signatures**

\[
\text{bin:from-octets}(\text{\$in as xs:integer}^*) \rightarrow \text{xs:base64Binary}
\]

**Summary**

Converts a sequence of octets into binary data. Octets are integers from 0 to 255. If the value of \$in is the empty sequence, the function returns zero-sized binary data.

**Errors**

- octet-out-of-range: one of the octets lies outside the range 0 - 255.

### Basic Operations

#### bin:length

**Signatures**

\[
\text{bin:length}(\text{\$in as xs:base64Binary}) \rightarrow \text{xs:integer}
\]

**Summary**

Returns the size of binary data in octets.

#### bin:part

**Signatures**

\[
\text{bin:part}(\text{\$in as xs:base64Binary?}, \text{\$offset as xs:integer}) \rightarrow \text{xs:base64Binary}
\]

**Summary**

Returns a section of binary data starting at the \$offset octet. If \$size is specified, the size of the returned binary data is \$size octets. If \$size is absent, all remaining data from \$offset is returned. The \$offset is zero based. If the value of \$in is the empty sequence, the function returns an empty sequence.

**Errors**

- negative-size: the specified size is negative.
- index-out-of-range: the specified offset + size is out of range.

**Examples**

Test whether binary data starts with binary content consistent with a PDF file:

\[
\text{bin:part}(\text{\$data, 0, 4) eq bin:hex("25504446")}. 
\]

#### bin:join

**Signatures**

\[
\text{bin:join}(\text{\$in as xs:base64Binary}) \rightarrow \text{xs:base64Binary}
\]

**Summary**

Returns an \text{xs:base64Binary} created by concatenating the items in the sequence \$in, in order. If the value of \$in is the empty sequence, the function returns a binary item containing no data bytes.

#### bin:insert-before

**Signatures**

\[
\text{bin:insert-before}(\text{\$in as xs:base64Binary?}, \text{\$offset as xs:integer}, \text{\$extra as xs:base64Binary?}) \rightarrow \text{xs:base64Binary}
\]

**Summary**

Returns binary data consisting sequentially of the data from \$in up to and including the \$offset - 1 octet, followed by all the data from \$extra, and then the remaining data from \$in. The \$offset is zero based. If the value of \$in is the empty sequence, the function returns an empty sequence.

**Errors**

- index-out-of-range: the specified offset is out of range.
**Binary Module**

**bin:pad-left**

**Signatures**

```
bin:pad-left($in as xs:base64Binary?, $size as xs:integer) as xs:base64Binary?
bin:pad-left($in as xs:base64Binary?, $size as xs:integer, $octet as xs:integer) as xs:base64Binary?
```

**Summary**

Returns an `xs:base64Binary` created by padding the input with `$size` octets in front of the input. If `$octet` is specified, the padding octets each have that value, otherwise they are zero. If the value of `$in` is the empty sequence, the function returns an empty sequence.

**Errors**

- `negative-size`: the specified size is negative.
- `octet-out-of-range`: the specified octet lies outside the range 0-255.

**bin:pad-right**

**Signatures**

```
bin:pad-right($in as xs:base64Binary?, $size as xs:integer) as xs:base64Binary?
bin:pad-right($in as xs:base64Binary?, $size as xs:integer, $octet as xs:integer) as xs:base64Binary?
```

**Summary**

Returns an `xs:base64Binary` created by padding the input with `$size` octets after the input. If `$octet` is specified, the padding octets each have that value, otherwise they are zero. If the value of `$in` is the empty sequence, the function returns an empty sequence.

**Errors**

- `negative-size`: the specified size is negative.
- `octet-out-of-range`: the specified octet lies outside the range 0-255.

**bin:find**

**Signatures**

```
bin:find($in as xs:base64Binary?, $offset as xs:integer, $search as xs:base64Binary) as xs:integer?
```

**Summary**

Returns the first location of the binary search sequence in the input, or if not found, the empty sequence. The `$offset` and the returned location are zero based. If the value of `$in` is the empty sequence, the function returns an empty sequence.

**Errors**

- `index-out-of-range`: the specified offset + size is out of range.

**Text Decoding and Encoding**

**bin:decode-string**

**Signatures**

```
bin:decode-string($in as xs:base64Binary?, $encoding as xs:string) as xs:string?
bin:decode-string($in as xs:base64Binary?, $encoding as xs:string, $offset as xs:integer) as xs:string?
bin:decode-string($in as xs:base64Binary?, $encoding as xs:string, $offset as xs:integer, $size as xs:integer) as xs:string?
```

**Summary**

Decodes binary data as a string in a given `$encoding`. If `$offset` and `$size` are provided, the `$size` octets from `$offset` are decoded. If `$offset` alone is provided, octets from `$offset` to the end are decoded. If the value of `$in` is the empty sequence, the function returns an empty sequence.

**Errors**

- `negative-size`: the specified size is negative.
- `index-out-of-range`: the specified offset + size is out of range.
- `unknown-encoding`: the specified encoding is unknown.
- `conversion-error`: an error or malformed input occurred during decoding the string.

**Examples**

Tests whether the binary data starts with binary content consistent with a PDF file:

```
bin:decode-string($data, 'UTF-8', 0, 4) eq '%PDF'.
```

**bin:encode-string**

**Signatures**

```
bin:encode-string($in as xs:string?, $encoding as xs:string) as xs:base64Binary?
```

---

171
### Summary
Encodes a string into binary data using a given $encoding. If the value of $in is the empty sequence, the function returns an empty sequence.

<table>
<thead>
<tr>
<th>Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>unknown-encoding: the specified encoding is unknown.</td>
</tr>
<tr>
<td>conversion-error: an error or malformed input occurred during encoding the string.</td>
</tr>
</tbody>
</table>

### Packing and Unpacking of Numeric Values

The functions have an optional parameter $octet-order whose string value controls the order: Least-significant-first order is indicated by any of the values least-significant-first, little-endian, or LE. Most-significant-first order is indicated by any of the values most-significant-first, big-endian, or BE.

#### bin:pack-double

**Signatures**

<table>
<thead>
<tr>
<th>bin:pack-double($in as xs:double) as xs:base64Binary</th>
</tr>
</thead>
<tbody>
<tr>
<td>bin:pack-double($in as xs:double, $octet-order as xs:string) as xs:base64Binary</td>
</tr>
</tbody>
</table>

**Summary**

Returns the 8-octet binary representation of a double value. Most-significant-octet-first number representation is assumed unless the $octet-order parameter is specified.

**Errors**

| unknown-significance-order: the specified octet order is unknown. |

#### bin:pack-float

**Signatures**

<table>
<thead>
<tr>
<th>bin:pack-float($in as xs:float) as xs:base64Binary</th>
</tr>
</thead>
<tbody>
<tr>
<td>bin:pack-float($in as xs:float, $octet-order as xs:string) as xs:base64Binary</td>
</tr>
</tbody>
</table>

**Summary**

Returns the 4-octet binary representation of a float value. Most-significant-octet-first number representation is assumed unless the $octet-order parameter is specified.

**Errors**

| unknown-significance-order: the specified octet order is unknown. |

#### bin:pack-integer

**Signatures**

<table>
<thead>
<tr>
<th>bin:pack-integer($in as xs:integer, $size as xs:integer) as xs:base64Binary</th>
</tr>
</thead>
<tbody>
<tr>
<td>bin:pack-integer($in as xs:integer, $size as xs:integer, $octet-order as xs:string) as xs:base64Binary</td>
</tr>
</tbody>
</table>

**Summary**

Returns the two's-complement binary representation of an integer value treated as $size octets long. Any 'excess' high-order bits are discarded. Most-significant-octet-first number representation is assumed unless the $octet-order parameter is specified. Specifying a $size of zero yields an empty binary data.

**Errors**

| unknown-significance-order: the specified octet order is unknown. |
| negative-size: the specified size is negative. |

#### bin:unpack-double

**Signatures**

<table>
<thead>
<tr>
<th>bin:unpack-double($in as xs:base64Binary, $offset as xs:integer) as xs:double</th>
</tr>
</thead>
<tbody>
<tr>
<td>bin:unpack-double($in as xs:base64Binary, $offset as xs:integer, $octet-order as xs:string) as xs:double</td>
</tr>
</tbody>
</table>

**Summary**

Extracts the double value stored at the particular offset in binary data. Most-significant-octet-first number representation is assumed unless the $octet-order parameter is specified. The $offset is zero based.

**Errors**

| index-out-of-range: the specified offset is out of range. |
| unknown-significance-order: the specified octet order is unknown. |
**bin:unpack-float**

**Signatures**

\[
\text{bin:unpack-float}($\text{in} \text{ as xs:base64Binary}, \text{ $\text{offset} \text{ as xs:integer}}) \text{ as xs:float}
\]

\[
\text{bin:unpack-float}($\text{in} \text{ as xs:base64Binary}, \text{ $\text{offset} \text{ as xs:integer}}, \text{ $\text{octet-order} \text{ as xs:string}}) \text{ as xs:float}
\]

**Summary**

Extracts the float value stored at the particular offset in binary data. Most-significant-octet-first number representation is assumed unless the $octet-order parameter is specified. The $offset is zero based.

**Errors**

index-out-of-range: the specified offset + size is out of range
unknown-significance-order: the specified octet order is unknown.

**bin:unpack-integer**

**Signatures**

\[
\text{bin:unpack-integer}($\text{in} \text{ as xs:base64Binary}, \text{ $\text{offset} \text{ as xs:integer}}, \text{ $\text{size} \text{ as xs:integer}}) \text{ as xs:integer}
\]

\[
\text{bin:unpack-integer}($\text{in} \text{ as xs:base64Binary}, \text{ $\text{offset} \text{ as xs:integer}}, \text{ $\text{size} \text{ as xs:integer}}, \text{ $\text{octet-order} \text{ as xs:string}}) \text{ as xs:integer}
\]

**Summary**

Returns a signed integer value represented by the $size octets starting from $offset in the input binary representation. Necessary sign extension is performed (i.e. the result is negative if the high order bit is ‘1’). Most-significant-octet-first number representation is assumed unless the $octet-order parameter is specified. The $offset is zero based. Specifying a $size of zero yields the integer 0.

**Errors**

negative-size: the specified size is negative
index-out-of-range: the specified offset + size is out of range
unknown-significance-order: the specified octet order is unknown.

**bin:unpack-unsigned-integer**

**Signatures**

\[
\text{bin:unpack-unsigned-integer}($\text{in} \text{ as xs:base64Binary}, \text{ $\text{offset} \text{ as xs:integer}}, \text{ $\text{size} \text{ as xs:integer}}) \text{ as xs:integer}
\]

\[
\text{bin:unpack-unsigned-integer}($\text{in} \text{ as xs:base64Binary}, \text{ $\text{offset} \text{ as xs:integer}}, \text{ $\text{size} \text{ as xs:integer}}, \text{ $\text{octet-order} \text{ as xs:string}}) \text{ as xs:integer}
\]

**Summary**

Returns an unsigned integer value represented by the $size octets starting from $offset in the input binary representation. Most-significant-octet-first number representation is assumed unless the $octet-order parameter is specified. The $offset is zero based. Specifying a $size of zero yields the integer 0.

**Errors**

negative-size: the specified size is negative
index-out-of-range: the specified offset + size is out of range
unknown-significance-order: the specified octet order is unknown.

**Bitwise Operations**

**bin:or**

**Signatures**

\[
\text{bin:or}($\text{a} \text{ as xs:base64Binary?}, \text{ $\text{b} \text{ as xs:base64Binary?}}) \text{ as xs:base64Binary}
\]

**Summary**

Returns the "bitwise or" of two binary arguments. If either argument is the empty sequence, an empty sequence is returned.

**Errors**

differing-length-arguments: the input arguments are of differing length.

**bin:xor**

**Signatures**

\[
\text{bin:xor}($\text{a} \text{ as xs:base64Binary?}, \text{ $\text{b} \text{ as xs:base64Binary?}}) \text{ as xs:base64Binary}
\]
Binary Module

Summary
Returns the "bitwise xor" of two binary arguments. If either argument is the empty sequence, an empty sequence is returned.

Errors
differing-length-arguments: the input arguments are of differing length.

bin:and

Signatures
bin:and($a as xs:base64Binary?, $b as xs:base64Binary?) as xs:base64Binary?

Summary
Returns the "bitwise and" of two binary arguments. If either argument is the empty sequence, an empty sequence is returned.

Errors
differing-length-arguments: the input arguments are of differing length.

bin:not

Signatures
bin:not($in as xs:base64Binary?) as xs:base64Binary?

Summary
Returns the "bitwise not" of a binary argument. If the argument is the empty sequence, an empty sequence is returned.

bin:shift

Signatures
bin:shift($in as xs:base64Binary?, $by as xs:integer) as xs:base64Binary?

Summary
Shifts bits in binary data. If $by is zero, the result is identical to $in. If $by is positive then bits are shifted to the left. Otherwise, bits are shifted to the right. If the absolute value of $by is greater than the bit-length of $in then an all-zeros result is returned. The result always has the same size as $in. The shifting is logical: zeros are placed into discarded bits. If the value of $in is the empty sequence, the function returns an empty sequence.

Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>differing-length-arguments</td>
<td>The arguments to a bitwise operation have different lengths.</td>
</tr>
<tr>
<td>index-out-of-range</td>
<td>An offset value is out of range.</td>
</tr>
<tr>
<td>negative-size</td>
<td>A size value is negative.</td>
</tr>
<tr>
<td>octet-out-of-range</td>
<td>An octet value lies outside the range 0-255.</td>
</tr>
<tr>
<td>non-numeric-character</td>
<td>Binary data cannot be parsed as number.</td>
</tr>
<tr>
<td>unknown-encoding</td>
<td>An encoding is not supported.</td>
</tr>
<tr>
<td>conversion-error</td>
<td>An error or malformed input during converting a string.</td>
</tr>
<tr>
<td>unknown-significance-order</td>
<td>An octet-order value is unknown.</td>
</tr>
</tbody>
</table>

Changelog

Introduced with Version 7.8.
Chapter 37. Client Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions to access remote BaseX server instances from XQuery. With this module, you can on the one hand execute database commands and on the other hand evaluate queries, the results of which are returned as XDM sequences.

Conventions

All functions in this module are assigned to the http://basex.org/modules/client namespace, which is statically bound to the client prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

Functions

client:connect

Signatures

client:connect($host as xs:string, $port as xs:integer, $user as xs:string, $password as xs:string) as xs:anyURI

Summary

This function establishes a connection to a remote BaseX server, creates a new client session, and returns a session id. The parameter $host is the name of the database server, $port specifies the server port, and $user and $password represent the login data.

Errors

BXCL0001: an error occurs while creating the session (possible reasons: server not available, access denied).

client:execute

Signatures

client:execute($id as xs:anyURI, $command as xs:string) as xs:string

Summary

This function executes a command and returns the result as string. The parameter $id contains the session id returned by client:connect. The $command argument represents a single command, which will be executed by the server.

Errors

BXCL0003: an I/O error occurs while transferring data from or to the server. BXCL0004: an error occurs while executing a command.

Examples

The following query creates a new database TEST on a remote BaseX server:

client:connect('basex.server.org', 8080, 'admin', 'admin') !
client:execute(., 'create database TEST')

client:info

Signatures

client:info($id as xs:anyURI) as xs:string

Summary

This function returns an information string, created by the last call of client:execute. $id specifies the session id.

client:query

Signatures

client:query($id as xs:anyURI, $query as xs:string) as item()*

client:query($id as xs:anyURI, $query as xs:string, $bindings as map(*)) as item()*

Summary

Evaluates a query and returns the result as sequence. The parameter $id contains the session id returned by client:connect, and $query represents the query string, which will be evaluated by the
server. Variables and the context item can be declared via $bindings. The specified keys must be QNames or strings:

- If a key is a QName, it will be directly adopted as variable name.
- If a key is a string, it may be prefixed with a dollar sign. A namespace can be specified using the Clark Notation. If the specified string is empty, the value will be bound to the context item.

### Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXCL0003</td>
<td>an I/O error occurs while transferring data from or to the server.</td>
</tr>
<tr>
<td>BXCL0005</td>
<td>an error occurs while evaluating a query, and if the original error cannot be extracted from the returned error string.</td>
</tr>
</tbody>
</table>

### Examples

The following query sends a query on a local server instance, binds the integer 123 to the variable $n and returns 246:

```xml
let $c := client:connect('localhost', 1984, 'admin', 'admin')
return client:query($c, "declare variable $n external; $n * 2", map ('n': 123 ))
```

The following query performs a query on a first server, the results of which are passed on to a second server:

```xml
let $c1 := client:connect('basex1.server.org', 8080, 'jack', 'C0S19tt2X')
let $c2 := client:connect('basex2.server.org', 8080, 'john', '465wFHe26')
for $it in client:query($c1, '1 to 10')
return client:query($c2, $it || '* 2')
```

### client:close

**Signatures**

client:close($id as xs:anyURI) as empty-sequence()

**Summary**

This function closes a client session. $id specifies the session id. At the end of query execution, open sessions will be automatically closed.

**Errors**

BXCL0003: an I/O error occurs while transferring data from or to the server.

### Changelog

**Version 8.0**

- Updated: Bound values may now contain no or more than one item in client:query.

**Version 7.5**

- Added: client:info
The module was introduced with Version 7.3.
Chapter 38. Conversion Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions to convert data between different formats.

Conventions

All functions in this module are assigned to the http://basex.org/modules/convert namespace, which is statically bound to the convert prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

Strings

convert:binary-to-string

Signatures

convert:binary-to-string($bytes as xs:anyAtomicType) as xs:string
convert:binary-to-string($bytes as xs:anyAtomicType, $encoding as xs:string) as xs:string

Summary

Converts the specified binary data (xs:base64Binary, xs:hexBinary) to a string. The UTF-8 default encoding can be overwritten with the optional $encoding argument.

Errors

BXCO0001: The input is an invalid XML string, or the wrong encoding has been specified. Invalid XML characters will be ignored if the CHECKSTRINGS option is turned off.BXCO0002: The specified encoding is invalid or not supported.

Examples

• convert:binary-to-string(xs:hexBinary('48656c6c6f576f726c64')) returns the string HelloWorld.

convert:string-to-base64

Signatures

convert:string-to-base64($input as xs:string) as xs:base64Binary
convert:string-to-base64($input as xs:string, $encoding as xs:string) as xs:base64Binary

Summary

Converts the specified string to a xs:base64Binary item. If the default encoding is chosen, conversion will be cheap, as both xs:string and xs:base64Binary items are internally represented as byte arrays. The UTF-8 default encoding can be overwritten with the optional $encoding argument.

Errors

BXCO0001: The input cannot be represented in the specified encoding.BXCO0002: The specified encoding is invalid or not supported.

Examples

• convert:string-to-base64('HelloWorld') returns the xs:base64binary item SGVsbG9Xb3JsZA==.

convert:string-to-hex

Signatures

convert:string-to-hex($input as xs:string) as xs:hexBinary
convert:string-to-hex($input as xs:string, $encoding as xs:string) as xs:hexBinary

Summary

Converts the specified string to a xs:hexBinary item. If the default encoding is chosen, conversion will be cheap, as both xs:string and xs:hexBinary items are internally represented as byte arrays. The UTF-8 default encoding can be overwritten with the optional $encoding argument.

Errors

BXCO0001: The input cannot be represented in the specified encoding.BXCO0002: The specified encoding is invalid or not supported.
Conversion Module

Examples

• `convert:string-to-hex('HelloWorld')` returns the Base64 item `48656C6C6F576F726C64`.

Binary Data

convert:bytes-to-base64

Signatures

`convert:bytes-to-base64($input as xs:byte*) as xs:base64Binary`

Summary

Converts the specified byte sequence to a `xs:base64Binary` item. Conversion is cheap, as `xs:base64Binary` items are internally represented as byte arrays.

Errors

- `BXCO0001`: The input cannot be represented in the specified encoding.
- `BXCO0002`: The specified encoding is invalid or not supported.

Examples

• `convert:string-to-base64('HelloWorld')` returns the `xs:base64binary` item `SGVsbG9Xb3JsZA==`.

convert:bytes-to-hex

Signatures

`convert:bytes-to-hex($input as xs:byte*) as xs:hexBinary`

Summary

Converts the specified byte sequence to a `xs:hexBinary` item. Conversion is cheap, as `xs:hexBinary` items are internally represented as byte arrays.

convert:binary-to-bytes

Signatures

`convert:binary-to-bytes($bin as xs:anyAtomicType) as xs:byte*`

Summary

Returns the specified binary data (`xs:base64Binary`, `xs:hexBinary`) as a sequence of bytes.

Examples

• `convert:binary-to-`bytes(xs:base64Binary('QmFzZVggaXMgY29vbA==')) returns the sequence `(66, 97, 115, 101, 88, 32, 105, 115, 32, 99, 111, 111, 108)`.

• `convert:binary-to-bytes(xs:hexBinary("4261736558"))` returns the sequence `(66 97 115 101 88)`.

Numbers

convert:integer-to-base

Signatures

`convert:integer-to-base($num as xs:integer, $base as xs:integer) as xs:string`

Summary

Converts `$num` to base `$base`, interpreting it as a 64-bit unsigned integer. The first `$base` elements of the sequence `'0',..,'9','a',..,'z` are used as digits. Valid bases are 2, ..., 36.

Errors

- `BXCO0004`: The specified base is not in the range 2-36.

Examples

• `convert:integer-to-base(-1, 16)` returns the hexadecimal string `'fffffff11111111'`.

• `convert:integer-to-base(22, 5)` returns `'42'`.

convert:integer-from-base

Signatures

`convert:integer-from-base($str as xs:string, $base as xs:integer) as xs:integer`

Summary

Decodes an `xs:integer` from `$str`, assuming that it’s encoded in base `$base`. The first `$base` elements of the sequence `'0',..,'9','a',..,'z` are allowed as digits, case doesn’t matter.
Valid bases are 2 - 36. If $\text{str}$ contains more than 64 bits of information, the result is truncated arbitrarily.

<table>
<thead>
<tr>
<th>Errors</th>
<th>BXCO0004: The specified base is not in the range 2-36. BXCO0005: The specified digit is not valid for the given range.</th>
</tr>
</thead>
</table>

| Examples | • `convert:integer-from-base('ffffffffffffffff', 16)` returns -1.  
• `convert:integer-from-base('CAFEBABE', 16)` returns 3405691582.  
• `convert:integer-from-base('42', 5)` returns 22.  
• `convert:integer-from-base(convert:integer-to-base(123, 7), 7)` returns 123. |

### Dates and Durations

#### `convert:integer-to-dateTime`

<table>
<thead>
<tr>
<th>Signatures</th>
<th><code>convert:integer-to-dateTime($ms as xs:integer) as xs:dateTime</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Converts the specified number of milliseconds since 1 Jan 1970 to an item of type xs:dateTime.</td>
</tr>
</tbody>
</table>
| Examples   | • `convert:integer-to-dateTime(0)` returns 1970-01-01T00:00:00Z.  

#### `convert:dateTime-to-integer`

<table>
<thead>
<tr>
<th>Signatures</th>
<th><code>convert:dateTime-to-integer($dateTime as xs:dateTime) as xs:integer</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Converts the specified item of type xs:dateTime to the number of milliseconds since 1 Jan 1970.</td>
</tr>
<tr>
<td>Examples</td>
<td>• <code>convert:dateTime-to-integer(xs:dateTime('1970-01-01T00:00:00Z'))</code> returns 0.</td>
</tr>
</tbody>
</table>

#### `convert:integer-to-dayTime`

<table>
<thead>
<tr>
<th>Signatures</th>
<th><code>convert:integer-to-dayTime($ms as xs:integer) as xs:dayTimeDuration</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Converts the specified number of milliseconds to an item of type xs:dayTimeDuration.</td>
</tr>
<tr>
<td>Examples</td>
<td>• <code>convert:integer-to-dayTime(1234)</code> returns PT1.234S.</td>
</tr>
</tbody>
</table>

#### `convert:dayTime-to-integer`

<table>
<thead>
<tr>
<th>Signatures</th>
<th><code>convert:dayTime-to-integer($dayTime as xs:dayTimeDuration) as xs:integer</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Converts the specified item of type xs:dayTimeDuration to milliseconds represented by an integer.</td>
</tr>
<tr>
<td>Examples</td>
<td>• <code>convert:dayTime-to-integer(xs:dayTimeDuration('PT1S'))</code> returns 1000.</td>
</tr>
</tbody>
</table>

### Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXCO0001</td>
<td>The input is an invalid XML string, or the wrong encoding has been specified.</td>
</tr>
<tr>
<td>BXCO0002</td>
<td>The specified encoding is invalid or not supported.</td>
</tr>
<tr>
<td>BXCO0003</td>
<td>The specified base is not in the range 2-36.</td>
</tr>
</tbody>
</table>
BXCO0004 The specified encoding is invalid or not supported.
BXCO0005 The specified digit is not valid for the given range.

**Changelog**

Version 7.5

- **Added:** `convert:integer-to-dateTime`, `convert:dateTime-to-integer`, `convert:integer-to-dayTime`, `convert:dayTime-to-integer`

The module was introduced with Version 7.3. Some of the functions have been adopted from the obsolete Utility Module.
Chapter 39. Cryptographic Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions to perform cryptographic operations in XQuery. The cryptographic module is based on an early draft of the EXPath Cryptographic Module and provides the following functionality: creation of message authentication codes (HMAC), encryption and decryption, and creation and validation of XML Digital Signatures.

Conventions

All functions in this module are assigned to the http://expath.org/ns/crypto namespace, which is statically bound to the crypto prefix. All errors are assigned to the http://expath.org/ns/error namespace, which is statically bound to the experr prefix.

Message Authentication

crypto:hmac

| Signatures | crypto:hmac($message as xs:string, $key as xs:string, $algorithm as xs:string) as xs:string crypto:hmac($message as xs:string, $key as xs:string, $algorithm as xs:string, $encoding as xs:string) as xs:string |
| Summary | Creates a message authentication code via a cryptographic hash function and a secret $key. $encoding must either be hex, base64 or the empty string and specifies the encoding of the returned authentication code. Default is base64. $algorithm describes the hash algorithm which is used for encryption. Currently supported are md5, sha1, sha256, sha384, sha512. Default is md5. |
| Errors | CX0013: the specified hashing algorithm is not supported. CX0014: the specified encoding method is not supported. CX0019: the specified secret key is invalid. |
| Example | Returns the message authentication code (MAC) for a given string. Query: |
| | crypto:hmac('message','secretkey','md5','base64') |
| Result | 34D1E3818B347252A75A4F6D747B21C2 |

Encryption & Decryption

The encryption and decryption functions underlie several limitations:

- Cryptographic algorithms are currently limited to symmetric algorithms only. This means that the same secret key is used for encryption and decryption.
- Available algorithms are DES and AES.
- Padding is fixed to PKCS5Padding.
- The result of an encryption using the same message, algorithm and key looks different each time it is executed. This is due to a random initialization vector (IV) which is appended to the message and simply increases security.
- As the IV has to be passed along with the encrypted message somehow, data which has been encrypted by the crypto:encrypt function in BaseX can only be decrypted by calling the crypto:decrypt function.
**crypto:encrypt**

**Signatures**
crypto:encrypt($input as xs:string, $encryption as xs:string, $key as xs:string, $algorithm as xs:string) as xs:string

**Summary**
Encrypts the given input string. $encryption must be symmetric, as asymmetric encryption is not supported so far. Default is symmetric. $key is the secret key which is used for both encryption and decryption of input data. Its length is fixed and depends on the chosen algorithm: 8 bytes for DES, 16 bytes for AES. $algorithm must either be DES or AES. Other algorithms are not supported so far, but, of course, can be added on demand. **Default is DES.**

**Errors**
- CX0016: padding problems arise.
- CX0017: padding is incorrect.
- CX0018: the encryption type is not supported.
- CX0019: the secret key is invalid.
- CX0020: the block size is incorrect.
- CX0021: the specified encryption algorithm is not supported.

**Example**
Encrypts input data. Query:

```
crypto:encrypt('message', 'symmetric','keykeyke','DES')
```

**crypto:decrypt**

**Signatures**
crypto:decrypt($input as xs:string, $type as xs:string, $key as xs:string, $algorithm as xs:string) as xs:string

**Summary**
Decrypts the encrypted $input. $type must be symmetric. An option for asymmetric encryption will most likely be added with another version of BaseX. Default is symmetric. $key is the secret key which is used for both encryption and decryption of input data. Its length is fixed and depends on the chosen algorithm: 8 bytes for DES, 16 bytes for AES. $algorithm must either be DES or AES. Other algorithms are not supported so far, but, of course, can be added on demand. **Default is DES.**

**Errors**
- CX0016: padding problems arise.
- CX0017: padding is incorrect.
- CX0018: the encryption type is not supported.
- CX0019: the secret key is invalid.
- CX0020: the block size is incorrect.
- CX0021: the specified encryption algorithm is not supported.

**Example**
Decrypts input data and returns the original string. Query:

```
let $encrypted := crypto:encrypt('message', 'symmetric','keykeyke','DES')
return crypto:decrypt($encrypted, 'symmetric','keykeyke','DES')
```

**Result:**
```
message
```

**XML Signatures**

XML Signatures are used to sign data. In our case, the data which is signed is an XQuery node. The following example shows the basic structure of an XML signature.

**XML Signature**

```xml
<Signature>
  <SignedInfo>
    <CanonicalizationMethod/>
    <SignatureMethod/>
    <Reference>
      <Transforms/>
      <DigestMethod/>
      <DigestValue/>
    </Reference>
  </SignedInfo>
</Signature>
```
• **SignedInfo** contains or references the signed data and lists algorithm information

• **Reference** references the signed node

• **Transforms** contains transformations (i.e. XPath expressions) that are applied to the input node in order to sign a subset

• **DigestValue** holds digest value of the transformed references

• **SignatureValue** contains the Base64 encoded value of the encrypted digest of the SignedInfo element

• **KeyInfo** provides information on the key that is used to validate the signature

• **Object** contains the node which is signed if the signature is of type `enveloping`

**Signature Types**

Depending on the signature type, the signature element is either placed as a child of the signed node (enveloped type), or directly contains the signed node (enveloping type). Detached signatures are so far not supported.

**Digital Certificate**

The `generate-signature` function allows to pass a digital certificate. This certificate holds parameters that allow to access key information stored in a Java key store which is then used to sign the input document. Passing a digital certificate simply helps re-using the same key pair to sign and validate data. The digital certificate is passed as a node and has the following form:

```xml
<digital-certificate>
  <keystore-type>JKS</keystore-type>
  <keystore-password>...</keystore-password>
  <key-alias>...</key-alias>
  <private-key-password>...</private-key-password>
  <keystore-uri>...</keystore-uri>
</digital-certificate>
```

**crypto:generate-signature**

| Signatures | crypto:generate-signature($input as node(), $canonicalization as xs:string, $digest as xs:string, $signature as(xs:string, $prefix as(xs:string, $type as(xs:string)) as node() crypto:generate-signature($input as node(), $canonicalization as xs:string, $digest as xs:string, $signature as(xs:string, $prefix as(xs:string, $type as(xs:string, $xpath as(xs:string, $certificate as(xs:string, $certicate as node()) as node() crypto:generate-signature($input as node(), $canonicalization as(xs:string, $digest as(xs:string, $signature as(xs:string, $prefix as(xs:string, $type as(xs:string, $ext as item()) as node()) as node)

| Summary | $canonicalization must either be inclusive-with-comments, inclusive-excluding-comments or exclusive. Default is inclusive-with-comments. $digest must be one of the following: SHA1, SHA256 or SHA512. Default is SHA1. $signature must either be RSA_SHA1 or DSA_SHA1. Default is RSA_SHA1. $prefix may be empty and prefixes the Signature element accordingly. $type is the signature type. It must
either be enveloped or enveloping (detached signatures are not supported so far). **Default is enveloped**. $xpath$ is an arbitrary XPath expression which specifies a subset of the document that is to be signed. $certificate$ is the digital certificate used to sign the input document. $ext$ may either be an $xpath$ expression or a $certificate$.

**Errors**

CX0001: the canonicalization algorithm is not supported.
CX0002: the digest algorithm is not supported.
CX0003: the signature algorithm is not supported.
CX0004: the $xpath$-expression is invalid.
CX0005: the root name of $digital-certificate$ is not 'digital-certificate'.
CX0007: the key store is null.
CX0012: the key cannot be found in the specified key store.
CX0023: the certificate alias is invalid.
CX0024: an invalid algorithm is specified.
CX0025: an exception occurs while the signing the document.
CX0026: an exception occurs during key store initialization.
CX0027: an IO exception occurs.
CX0028: the signature type is not supported.

**Example**

Generates an XML Signature. Query:

```
crypto:generate-signature(<a/>, '', '', '', '', '')
```

**Result:**

```
<a>
  <Signature xmlns="http://www.w3.org/2000/09/xmldsig#"/>
  <SignedInfo>
    <CanonicalizationMethod Algorithm="http://www.w3.org/TR/2001/REC-xml-c14n-20010315#WithComments"/>
    <SignatureMethod Algorithm="http://www.w3.org/2000/09/xmldsig#rsa-sha1"/>
    <Reference URI=""/>
    <Transforms>
      <Transform Algorithm="http://www.w3.org/2000/09/xmldsig#enveloped-signature"/>
    </Transforms>
    <DigestMethod Algorithm="http://www.w3.org/2000/09/xmldsig#sha1"/>
    <DigestValue>9hvH4qztn1YgYfJDRLnEMPJdoaY="</DigestValue>
  </SignedInfo>
  <SignatureValue>Pn/Jr44WBcdARff2UVYEiwyYW1563XdqgU87nusAIaHgzd+U3SrjVJhPFLDe0DjfRxVtYzLFaZnTYE P3deofMyA="</SignatureValue>
  <KeyInfo>
    <KeyValue>
      <RSAKeyValue>
        <Modulus>rtvpFSbCIE2BJeP1VYLIRjX10R7ESr2+D+JOVKn7AM7VzbcbRDPegRbjskEz1HWC/N067tjB3qH 4/4PPT9bGQ="</Modulus>
        <Exponent>AQAB</Exponent>
      </RSAKeyValue>
    </KeyValue>
  </KeyInfo>
</Signature>
</a>
```

**crypto:validate-signature**

**Signatures**

```
crypto:validate-signature($input-doc as node()) as xs:boolean
```

**Summary**

Checks if the given node contains a Signature element and whether the signature is valid. In this case true is returned. If the signature is invalid the function returns false.

**Errors**

CX0015: the signature element cannot be found.
CX9994: an unspecified problem occurs during validation.
CX9996: an IO exception occurs during validation.

**Example**

Validates an XML Signature. Query:
let $sig := crypto:generate-signature(<a/>, '', '', '', '', '')
return crypto:validate-signature($sig)

Result:
true

Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CX0001</td>
<td>The canonicalization algorithm is not supported.</td>
</tr>
<tr>
<td>CX0002</td>
<td>The digest algorithm is not supported.</td>
</tr>
<tr>
<td>CX0003</td>
<td>The signature algorithm is not supported.</td>
</tr>
<tr>
<td>CX0004</td>
<td>The XPath expression is invalid.</td>
</tr>
<tr>
<td>CX0005</td>
<td>The root element of argument $digital-certificate must have the name 'digital-certificate'.</td>
</tr>
<tr>
<td>CX0006</td>
<td>The child element of argument $digital-certificate having position $position must have the name $child-element-name.</td>
</tr>
<tr>
<td>CX0007</td>
<td>The keystore is null.</td>
</tr>
<tr>
<td>CX0008</td>
<td>I/O error while reading keystore.</td>
</tr>
<tr>
<td>CX0009</td>
<td>Permission denied to read keystore.</td>
</tr>
<tr>
<td>CX0010</td>
<td>The keystore URL is invalid.</td>
</tr>
<tr>
<td>CX0011</td>
<td>The keystore type is not supported.</td>
</tr>
<tr>
<td>CX0012</td>
<td>Cannot find key for alias in given keystore.</td>
</tr>
<tr>
<td>CX0013</td>
<td>The hashing algorithm is not supported.</td>
</tr>
<tr>
<td>CX0014</td>
<td>The encoding method is not supported.</td>
</tr>
<tr>
<td>CX0015</td>
<td>Cannot find Signature element.</td>
</tr>
<tr>
<td>CX0016</td>
<td>No such padding.</td>
</tr>
<tr>
<td>CX0017</td>
<td>Incorrect padding.</td>
</tr>
<tr>
<td>CX0018</td>
<td>The encryption type is not supported.</td>
</tr>
<tr>
<td>CX0019</td>
<td>The secret key is invalid.</td>
</tr>
<tr>
<td>CX0020</td>
<td>Illegal block size.</td>
</tr>
<tr>
<td>CX0021</td>
<td>The algorithm is not supported.</td>
</tr>
<tr>
<td>CX0023</td>
<td>An invalid certificate alias is specified. Added to the official specification.</td>
</tr>
<tr>
<td>CX0024</td>
<td>The algorithm is invalid. Added to the official specification.</td>
</tr>
<tr>
<td>CX0025</td>
<td>Signature cannot be processed. Added to the official specification.</td>
</tr>
<tr>
<td>CX0026</td>
<td>Keystore cannot be processed. Added to the official specification.</td>
</tr>
<tr>
<td>CX0027</td>
<td>An I/O Exception occurred. Added to the official specification.</td>
</tr>
<tr>
<td>CX0028</td>
<td>The specified signature type is not supported. Added to the official specification.</td>
</tr>
</tbody>
</table>

Changelog

The Module was introduced with Version 7.0.
Chapter 40. CSV Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains a single function to parse CSV input. CSV (comma-separated values) is a popular representation for tabular data, exported e.g. from Excel.

Conventions

All functions in this module are assigned to the http://basex.org/modules/csv namespace, which is statically bound to the csv prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

Conversion

XML: Direct, Attributes

CSV is converted to XML as follows:

- The resulting XML document has a <csv> root element.
- Rows are represented via <record> elements.
- Fields are represented via <entry> elements. The value of a field is represented as text node.
- If the header option is set to true, the first text line is parsed as table header, and the <entry> elements are replaced with the field names:
  - Empty names are represented by a single underscore (_), and characters that are not valid in element names are replaced with underscores or (when invalid as first character of an element name) prefixed with an underscore.
  - If the lax option is set to false, invalid characters will be rewritten to an underscore and the character’s four-digit Unicode, and underscores will be represented as two underscores (____). The resulting element names may be less readable, but can always be converted back to the original field names.
- If format is set to attributes, field names will be stored in name attributes.

Map

If format is set to map, the CSV data will be converted to an XQuery map:

- All records are enumerated with positive integers.
- By default, all entries of a records are represented in a sequence.
- If the header option is set to true, a map is created, which contains all field names and its values.

A little advice: in the Database Creation dialog of the GUI, if you select CSV Parsing and switch to the Parsing tab, you can see the effects of some of the conversion options.

Options

The following options are available:

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
<th>Allowed</th>
<th>Default</th>
</tr>
</thead>
<tbody>
<tr>
<td>separator</td>
<td>Defines the character which separates the entries of a record in a single line.</td>
<td>comma, semicolon, colon, tab, space or a single character</td>
<td>comma</td>
</tr>
</tbody>
</table>
header: Indicates if the first line of the parsed or serialized CSV data is a table header.  

<table>
<thead>
<tr>
<th>Format</th>
<th>Specifies the format of the XML data. The format is only relevant if the header option is activated:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• With direct conversion, field names are represented as element names</td>
<td></td>
</tr>
<tr>
<td>• With attributes conversion, field names are stored in name attributes</td>
<td></td>
</tr>
<tr>
<td>• With map conversion, the input is converted to an XQuery map</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lax</th>
<th>Specifies if a lax approach is used to convert QNames to JSON names.</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Quotes</th>
<th>Specifies if quotes should be parsed in the input and generated in the output.</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Backslashes</th>
<th>Specifies if characters are escaped by backslashes. Otherwise, a double quote is encoded by a second consecutive quote.</th>
</tr>
</thead>
</table>

## Functions

### csv:parse

```
csv:parse($input as xs:string) as document-node(element(csv))
csv:parse($input as xs:string, $options as map(xs:string, item())) as item()
```

**Summary**

Converts the CSV data specified by `$input` to an XML document or a map. The `$options` argument can be used to control the way the input is converted.

**Errors**

BXCS0001: the input cannot be parsed.

### csv:serialize

```
csv:serialize($input as node()) as xs:string
```

**Summary**

Serializes the node specified by `$input` as CSV data, and returns the result as `xs:string`. Items can also be serialized as JSON if the Serialization Parameter method is set to `csv`. The `$options` argument can be used to control the way the input is serialized.

**Errors**

BXCS0002: the input cannot be serialized.

## Examples

**Example 1:** Converts CSV data to XML, interpreting the first row as table header:

**Input** addressbook.csv:

<table>
<thead>
<tr>
<th>Name, First Name, Address, City</th>
</tr>
</thead>
<tbody>
<tr>
<td>Huber, Sepp, Hauptstraße 13, 93547 Hintertupfing</td>
</tr>
</tbody>
</table>

**Query:**

```
let $text := file:read-text('addressbook.csv')
```

---

188
return csv:parse($text, map { 'header': true() })

**Result:**

```
<csv>
  <record>
    <Name>Huber</Name>
    <First_Name>Sepp</First_Name>
    <Address>Hauptstraße 13</Address>
    <City>93547 Hintertupfing</City>
  </record>
</csv>
```

**Example 2:** Converts some CSV data to XML and back, and checks if the input and output are equal. The expected result is `true`:

**Query:**

```
let $text := file:read-text('some-data.csv')
let $options := map { 'lax': false() }
let $xml := csv:parse($text, $options)
let $csv := csv:serialize($xml, $options)
return $text eq $csv
```

**Example 3:** Converts CSV data to an XQuery map item and serializes its contents:

**Query:**

```
let $text := "Name;City" || out:nl() || "John;Newton" || out:nl() || "Jack;Oldtown"
let $options :=
  <csv:options>
    <csv:separator value=';'/>
    <csv:format value='map'/>
    <csv:header value='yes'/>
  </csv:options>
return csv:parse($text, $options)
```

**Result:**

```
{
  1: {
    "City": "Newton",
    "Name": "John"
  },
  2: {
    "City": "Oldtown",
    "Name": "Jack"
  }
}
```

**Errors**

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXCS00001</td>
<td>The input cannot be parsed.</td>
</tr>
<tr>
<td>BXCS00002</td>
<td>The node cannot be serialized.</td>
</tr>
</tbody>
</table>

**Changelog**

Version 8.0
• Added: backslashes option

Version 7.8

• Updated: csv:parse now returns a document node instead of an element, or an XQuery map if format is set to map.

• Added: format and lax options

The module was introduced with Version 7.7.2.
Chapter 41. Database Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions for processing databases from within XQuery. Existing databases can be opened and listed, its contents can be directly accessed, documents can be added to and removed, etc.

Conventions

All functions in this module are assigned to the http://basex.org/modules/db namespace, which is statically bound to the db prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

Database Nodes

Database nodes are XML nodes which are either stored in a persistent database or part of a so-called database fragment. All XML fragments can be converted to database fragments by e.g. applying the transform expression on an XML fragment:

```
copy $c := element hello { 'world' } modify () return $c
```

General Functions

**db:system**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>db:system() as element(system)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns information on the database system, such as the database path and current database settings. The output is similar to the INFO command.</td>
</tr>
</tbody>
</table>

**db:info**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>db:info($db as xs:string) as element(database)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns meta information on the database $db. The output is similar to the INFO DB command.</td>
</tr>
<tr>
<td>Errors</td>
<td>BXDB0002: the addressed database does not exist or could not be opened.</td>
</tr>
</tbody>
</table>

**db:list**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>db:list() as xs:string* db:list($db as xs:string) as xs:string* db:list($db as xs:string, $path as xs:string) as xs:string*</th>
</tr>
</thead>
</table>
| Summary          | Returns a string sequence with the names of all databases:  
|                  | • If a database $db is specified, all documents and raw files of the specified database are returned.  
|                  | • The list of resources can be further restricted by the $path argument. |
| Errors           | BXDB0002: the addressed database does not exist or could not be opened. |
| Examples         | • db:list("docs") returns the names of all documents from the database named docs. |

**db:list-details**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>db:list-details() as element(database)* db:list-details($db as xs:string) as element(resource)* db:list-details($db as xs:string, $path as xs:string) as element(resource)*</th>
</tr>
</thead>
</table>

### Database Module

**Summary**
- If no argument is specified, a sequence of elements is returned. A single element contains the name of a database, the number of stored resources, the date of modification, and the database path.
- If $db$ is specified, a sequence of elements is returned, comprising information on all resources of the addressed database. An element contains the name of the resource, the content type, the modified date, the raw flag and the size of a resource.
- For binary resources, the size is the number of bytes; for document resources, the size is the number of nodes.
- Returned databases resources can be further restricted by the $path$ argument.

**Errors**
- BXDB0002: the addressed database does not exist or could not be opened.

**Examples**
- `db:list-details("docs")` returns the names plus additional data of all documents from the database named docs.

---

#### db:backups

**Signatures**
- `db:backups() as element(backup)*`
- `db:backups($db as xs:string) as element(backup)*`

**Summary**
Returns an element sequence containing all available database backups. If a database $db$ is specified, the sequence will be restricted to the backups matching this database.

**Examples**
- `db:backups("factbook")` returns all backups that have been made from the factbook database.

---

#### Read Operations

#### db:open

**Signatures**
- `db:open($db as xs:string) as document-node()*`
- `db:open($db as xs:string, $path as xs:string) as document-node()*`

**Summary**
Opens the database $db and returns all document nodes. The document nodes to be returned can be filtered with the $path argument.

**Errors**
- BXDB0002: the addressed database does not exist or could not be opened.

**Examples**
- `db:open("docs")` returns all documents from the database named docs.
- `db:open("db", "one")` returns all documents from the database named db located in the path one.
- `for $i in 1 to 3 return db:open("db" || $i)//item` returns all item elements from the databases db1, db2 and db3.

---

#### db:open-pre

**Signatures**
- `db:open-pre($db as xs:string, $pre as xs:integer) as node()`

**Summary**
Opens the database $db and returns the node with the specified $pre value. The PRE value provides very fast access to an existing database node, but it will change whenever a node with a smaller $pre value is added to or deleted from a database.

**Errors**
- BXDB0002: the addressed database does not exist or could not be opened.
- BXDB0009: the specified $pre value does not exist in the database.

**Examples**
- `db:open-pre("docs", 0)` returns the first database node from the database named docs.

---

#### db:open-id

**Signatures**
- `db:open-id($db as xs:string, $id as xs:integer) as node()`
Summary
Opens the database $db and returns the node with the specified $id value. Each database node has a persistent ID value. Access to the node id can be sped up by turning on the UPDINDEX option.

Errors
BXDB0002: the addressed database does not exist or could not be opened. BXDB0009: the specified id value does not exist in the database.

db:node-pre

Signatures
db:node-pre($nodes as node()) as xs:integer*

Summary
Returns the pre values of the nodes supplied by $nodes, which must all be database nodes. The PRE value provides very fast access to an existing database node, but it will change whenever a node with a smaller pre values is added to or deleted from a database.

Errors
BXDB0001: $nodes contains a node which is not stored in a database.

Examples
• db:node-pre(doc("input")) returns 0 if the database input contains a single document.

db:node-id

Signatures
db:node-id($nodes as node()) as xs:integer*

Summary
Returns the id values of the nodes supplied by $nodes, which must all be database nodes. Each database node has a persistent ID value. Access to the node id can be sped up by turning on the UPDINDEX option.

Errors
BXDB0001: $nodes contains a node which is not stored in a database.

db:retrieve

Signatures
db:retrieve($db as xs:string, $path as xs:string) as xs:base64Binary
db:retrieve($db as xs:string, $path as xs:string, $params as item()) as empty-sequence()

Summary
Returns a binary resource addressed by the database $db and $path as streamable xs:base64Binary.

Errors
BXDB0002: the addressed database does not exist or could not be opened. BXDB0003: the database is not persistent (stored on disk). FODC0002: the addressed resource cannot be retrieved. FODC0007: the specified path is invalid.

Examples
• db:retrieve("DB", "music/01.mp3") returns the specified audio file as raw data.
• stream:materialize(db:retrieve("DB", "music/01.mp3")) materializes the streamable result in main-memory before returning it.

db:export

Signatures
db:export($db as xs:string, $path as xs:string) as empty-sequence()
db:export($db as xs:string, $path as xs:string, $params as item()) as empty-sequence()

Summary
Exports the specified database $db to the specified file $path. Existing files will be overwritten. The $params argument contains serialization parameters (see Serialization for more details), which can either be specified

• as children of an <output:serialization-parameters/> element, as defined for the fn:serialize() function; e.g.:

```xml
<output:serialization-parameters>
  <output:method value='xml'/>
  <output:cdata-section-elements value="div"/>
</output:serialization-parameters>
```
### Value Indexes

#### `db:text`

**Signatures**

```
db:text($db as xs:string, $string as item()) as text()*
```

**Summary**

Returns all text nodes of the database `$db` that have `$string` as their string value and that are stored in the text index.

**Errors**

- BXDB0002: the addressed database does not exist or could not be opened.
- BXDB0004: the index is not available.

**Examples**

- `db:text("DB", "QUERY")/..` returns the parents of all text nodes of the database `DB` that match the string `QUERY`.

#### `db:text-range`

**Signatures**

```
db:text-range($db as xs:string, $min as xs:string, $max as xs:string) as text()*
```

**Summary**

Returns all text nodes of the database `$db` that are located in between the `$min` and `$max` strings and that are stored in the text index.

**Errors**

- BXDB0002: the addressed database does not exist or could not be opened.
- BXDB0004: the index is not available.

**Examples**

- `db:text-range("DB", "2000", "2001")` returns all text nodes of the database `DB` that are found in between `2000` and `2001`.

#### `db:attribute`

**Signatures**

```
db:attribute($db as xs:string, $string as item()) as attribute()*
db:attribute($db as xs:string, $string as item(), $name as xs:string) as attribute()*
```

**Summary**

Returns all attribute nodes of the database `$db` that have `$string` as string value and that are stored in the attribute index. If `$name` is specified, the resulting attribute nodes are filtered by their attribute name.
### Database Module

<table>
<thead>
<tr>
<th>Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXDB0002: the addressed database does not exist or could not be opened. BXDB0004: the index is not available.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>• <code>db:attribute(&quot;DB&quot;, &quot;QUERY&quot;, &quot;id&quot;)//..</code> returns the parents of all <code>id</code> attribute nodes of the database <code>DB</code> that have <code>QUERY</code> as string value.</td>
</tr>
</tbody>
</table>

#### `db:attribute-range`

<table>
<thead>
<tr>
<th>Signatures</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>db:attribute-range($db as xs:string, $min as xs:string, $max as xs:string) as attribute()</code>* <code>db:attribute-range($db as xs:string, $min as xs:string, $max as xs:string, $name as xs:string) as attribute()</code>*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Returns all attributes of the database <code>$db</code>, the string values of which are larger than or equal to <code>$min</code> and smaller than or equal to <code>$max</code> and that are stored in the attribute index.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXDB0002: the addressed database does not exist or could not be opened. BXDB0004: the index is not available.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>• <code>db:attribute-range(&quot;DB&quot;, &quot;id456&quot;, &quot;id473&quot;, 'id')</code> returns all <code>@id</code> attributes of the database <code>DB</code> that have a string value in between <code>id456</code> and <code>id473</code>.</td>
</tr>
</tbody>
</table>

#### `db:token`

<table>
<thead>
<tr>
<th>Signatures</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>db:token($db as xs:string, $token as item()) as attribute()</code>* <code>db:token($db as xs:string, $token as item(), $name as xs:string) as attribute()</code>*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Returns all attribute nodes of the database <code>$db</code>, the value of which contains the specified <code>$token</code>. If <code>$name</code> is specified, the resulting attribute nodes are filtered by their attribute name.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXDB0002: the addressed database does not exist or could not be opened. BXDB0004: the index is not available.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>• <code>db:token(&quot;DB&quot;, &quot;row&quot;, &quot;class&quot;)/parent::div</code> returns all <code>div</code> nodes of database <code>DB</code> with a <code>class</code> attribute that contains the token <code>row</code>.</td>
</tr>
</tbody>
</table>

### Updates

**Important note:** All functions in this section are updating functions: they will not be immediately executed, but queued on the Pending Update List, which will be processed after the actual query has been evaluated. This means that the order in which the functions are specified in the query does usually not reflect the order in which the code will be evaluated.

#### `db:create`

<table>
<thead>
<tr>
<th>Signatures</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>db:create($db as xs:string) as empty-sequence()</code> <code>db:create($db as xs:string, $inputs as item()) as empty-sequence()</code> <code>db:create($db as xs:string, $inputs as item(), $paths as xs:string*) as empty-sequence()</code> <code>db:create($db as xs:string, $inputs as item(), $paths as xs:string*, $options as map(*)) as empty-sequence()</code></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creates a new database with name <code>$db</code> and adds initial documents specified via <code>$inputs</code> to the specified <code>$paths</code>:</td>
</tr>
<tr>
<td>• <code>$inputs</code> may be strings or nodes:</td>
</tr>
<tr>
<td>• nodes may be of any type except for attributes</td>
</tr>
<tr>
<td>• strings may be file/directory references, or XML strings (which are detected by the leading <code>&lt;</code> character)</td>
</tr>
<tr>
<td>• a path must be specified if the input is not a file or directory reference</td>
</tr>
</tbody>
</table>
The $options argument can be used to change the indexing behavior:

- the allowed options are ADDCACHE and the indexing, full-text, parsing and XML parsing options, all in lower case.

- the options are only applied to string input (XML strings or file/directory references).

- An existing database will be overwritten. Next, $db:create will be placed last on the Pending Update List. As a consequence, a newly created database cannot be addressed in the same query.

Errors

| FODC0002: an input points to an unknown resource. | FOUP0001: an attribute was specified as input. |
| BXDB0007: a database is opened by another process. |
| BXDB0011: the specified name is not a valid database name. |
| BXDB0012: two $db:create statements with the same database name were specified. |
| BXDB0013: the number of specified inputs and paths differs. |

Examples

- $db:create("DB") creates the empty database DB.
- $db:create("DB", "/home/dir/doc.xml") creates the database DB and adds the document /home/dir/doc.xml as initial content.
- $db:create("DB", <a/>, "doc.xml") creates the database DB and adds the document with content <a/> under the name doc.xml.
- $db:create("DB", "/home/dir/", "docs/dir") creates the database DB and adds the documents in /home/dir to the database under the path docs/dir.
- $db:create("DB", file:list('.'), (), map { 'ftindex': true() }) adds all files of the current working directory to a new database, preserving relative filesystem paths and creating a full-text index.

### $db:drop

**Signatures**

$db:drop($db as xs:string) as empty-sequence()

**Summary**

Drops the database $db and all connected resources.

**Errors**

| BXDB0002: the addressed database does not exist or could not be opened. |
| BXDB0007: a database is opened by another process. |

**Examples**

- $db:drop("DB") drops the database DB.

### $db:add

**Signatures**

$db:add($db as xs:string, $input as item()) as empty-sequence()

$db:add($db as xs:string, $input as item(), $path as xs:string) as empty-sequence()

$db:add($db as xs:string, $input as item(), $path as xs:string, $options as map(*)) as empty-sequence()

**Summary**

Adds documents specified by $input to the database $db with the specified $path:

- A document with the same path may occur more than once in a database. If you want to enforce single instances, use $db:replace instead.

- For more information on the arguments, see $db:create (note that the specified options cannot be indexing and full-text options).

**Errors**

| BXDB0002: the addressed database does not exist or could not be opened. |
| FODC0002: the input points to an unknown resource. |
| FOUP0001: an attribute was specified as input. |

**Examples**

- $db:add("DB", "/home/dir/doc.xml") adds the file /home/dir/doc.xml to the database DB.
- $db:add("DB", <a/>, "doc.xml") adds a document node to the database DB under the name doc.xml.

---

196
### db:add

**Signatures**

```xml
db:add($db as xs:string, $path as xs:string) as empty-sequence()
```

**Summary**

Adds all documents in the path specified by `$path` to the database specified by `$db`.

**Errors**

- BXDB0002: the addressed database does not exist or could not be opened.
- BXDB0008: the specified path is invalid.

**Examples**

- `db:add("DB", "/home/dir", "docs/dir")` adds all documents in `/home/dir` to the database `DB` under the path `docs/dir`.

### db:delete

**Signatures**

```xml
db:delete($db as xs:string, $path as xs:string) as empty-sequence()
```

**Summary**

Deletes resource(s), specified by `$path`, from the database `$db`.

**Errors**

- BXDB0002: the addressed database does not exist or could not be opened.
- BXDB0008: the specified path is invalid.

**Examples**

- `db:delete("DB", "docs/dir/doc.xml")` deletes the resource `docs/dir/doc.xml` from `DB`.
- `db:delete("DB", "docs/dir")` deletes all resources from `DB` in the specified path `docs/dir`.

### db:copy

**Signatures**

```xml
db:copy($db as xs:string, $newname as xs:string) as empty-sequence()
```

**Summary**

Creates a copy of the database specified by `$db` to `$newname`.

**Errors**

- BXDB0002: the addressed database does not exist or could not be opened.
- BXDB0011: invalid database name.
- BXDB0016: name of source and target database is equal.

### db:alter

**Signatures**

```xml
db:alter($db as xs:string, $newname as xs:string) as empty-sequence()
```

**Summary**

Renames the database specified by `$db` to `$newname`.

**Errors**

- BXDB0002: the addressed database does not exist or could not be opened.
- BXDB0011: invalid database name.
- BXDB0016: name of source and target database is equal.

### db:create-backup

**Signatures**

```xml
db:create-backup($db as xs:string) as empty-sequence()
```

**Summary**

Creates a backup of the database `$db`.

**Errors**

- BXDB0002: the addressed database does not exist or could not be opened.
- BXDB0011: invalid database name.

**Examples**

- `db:create-backup("DB")` creates a backup of the database `DB`.

### db:drop-backup

**Signatures**

```xml
db:drop-backup($name as xs:string) as empty-sequence()
```

**Summary**

Drops all backups of the database with the specified `$name`. If the given `$name` points to a specific backup file, only this specific backup file is deleted.

**Errors**

- BXDB0002: No backup file found.
- BXDB0011: invalid database name.

**Examples**

- `db:drop-backup("DB")` drops all backups of the database `DB`.
- `db:drop-backup("DB-2014-03-13-17-36-44")` drops the specific backup file `DB-2014-03-13-17-36-44.zip` of the database `DB`.

### db:restore

**Signatures**

```xml
db:restore($name as xs:string) as empty-sequence()
```

**Summary**

Restores the database from the specified backup file `$name`.

**Errors**

- BXDB0002: No backup file found.
- BXDB0011: invalid database name.

**Examples**

- `db:restore("DB")` restores the database `DB` from the backup file `DB.zip`.
- `db:restore("DB-2014-03-13-17-36-44")` restores the specific backup file `DB-2014-03-13-17-36-44.zip` of the database `DB`.
<table>
<thead>
<tr>
<th>Summary</th>
<th>Restores the database with the specified $name. The $name may include the timestamp of the backup file.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Errors</td>
<td>BXDB0011: invalid database name. BXDB0015: No backup found.</td>
</tr>
<tr>
<td>Examples</td>
<td>• db:restore(&quot;DB&quot;) restores the database DB.</td>
</tr>
<tr>
<td></td>
<td>• db:restore(&quot;DB-2014-03-13-18-05-45&quot;) restores the database DB from the backup file with the given timestamp.</td>
</tr>
</tbody>
</table>

**db:optimize**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>db:optimize($db as xs:string) as empty-sequence() db:optimize($db as xs:string, $all as xs:boolean) as empty-sequence() db:optimize($db as xs:string, $all as xs:boolean, $options as map(*)) as empty-sequence()</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Optimizes the meta data and indexes of the database $db. If $all is true, the complete database will be rebuilt. The $options argument can be used to control indexing. The syntax is identical to the db:create function: Allowed options are all indexing and full-text options. UPDINDEX is only supported if $all is true.</td>
</tr>
<tr>
<td>Errors</td>
<td>BXDB0002: the addressed database does not exist or could not be opened. FOUP0002: an error occurred while optimizing the database.</td>
</tr>
<tr>
<td>Examples</td>
<td>• db:optimize(&quot;DB&quot;) optimizes the database structures of the database DB.</td>
</tr>
<tr>
<td></td>
<td>• db:optimize(&quot;DB&quot;, true(), map { 'ftindex': true() }) optimizes all database structures of the database DB and creates a full-text index.</td>
</tr>
</tbody>
</table>

**db:rename**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>db:rename($db as xs:string, $source as xs:string, $target as xs:string) as empty-sequence()</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Moves all resources(s) of database $db, which are found in the supplied $source path, to the supplied $target path. The paths may point to single resources or directories. No updates will take place if a non-existing source path is supplied.</td>
</tr>
<tr>
<td>Errors</td>
<td>BXDB0002: the addressed database does not exist or could not be opened. BXDB0008: the specified source or target path, or one of its descendants, is invalid.</td>
</tr>
<tr>
<td>Examples</td>
<td>• db:rename(&quot;DB&quot;, &quot;docs/dir/doc.xml&quot;, &quot;docs/dir/newdoc.xml&quot;) renames the resource docs/dir/doc.xml to docs/dir/newdoc.xml in the database DB.</td>
</tr>
<tr>
<td></td>
<td>• db:rename(&quot;DB&quot;, &quot;docs/dir&quot;, &quot;docs/newdir&quot;) moves all resources in the database DB from docs/dir to {Code</td>
</tr>
</tbody>
</table>

**db:replace**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>db:replace($db as xs:string, $path as xs:string, $input as item()) as empty-sequence() db:replace($db as xs:string, $path as xs:string, $input as item(), $options as map(*)) as empty-sequence()</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Replaces a resource, specified by $path, in the database $db with the contents of $input, or adds it as a new resource: See db:add for more details on the arguments (for historical reasons, the order of the 2nd and 3rd argument differs!).</td>
</tr>
<tr>
<td>Errors</td>
<td>BXDB0002: the addressed database does not exist or could not be opened. BXDB0014: the path points to a directory. FOUC0002: the input points to an unknown resource. FOUP0001: an attribute was specified as input.</td>
</tr>
</tbody>
</table>
Examples

- `db:replace("DB", "docs/dir/doc.xml", "/home/dir/doc.xml")` replaces the content of the document `docs/dir/doc.xml` in the database `DB` with the content of the file `/home/dir/doc.xml`.

- `db:replace("DB", "docs/dir/doc.xml", "<a/>")` replaces the content of the document `docs/dir/doc.xml` in the database `DB` with `<a/>`.

- `db:replace("DB", "docs/dir/doc.xml", document { <a/> })` replaces the content of the document `docs/dir/doc.xml` in the database `DB` with the specified document node.

The following query can be used to import files from a directory to a database:

```xml
let $source := '/home/john/xml/source'
for $file in file:list($source, true())
let $path := $source || $file
where not(file:is-dir($path))
return db:replace('db', $file, doc($path))
```

\[ \text{db:store} \]

**Signatures**

```xml
db:store($db as xs:string, $path as xs:string, $input as item()) as empty-sequence()
```

**Summary**

Stores a binary resource specified by `$input` in the database `$db` and the location specified by `$path`.

**Errors**

- BXDB0002: the addressed database does not exist or could not be opened.
- BXDB0003: the database is not persistent (stored on disk).
- FODC0007: the specified path is invalid.
- FOUP0002: the resource cannot be stored at the specified location.

**Examples**

- `db:store("DB", "video/sample.mov", file:read-binary('video.mov'))` stores the addressed video file at the specified location.

\[ \text{db:output} \]

**Signatures**

```xml
db:output($result as item()*) as empty-sequence()
```

**Summary**

This function can be used to both perform updates and return results in a single query. The argument of the function will be evaluated, and the resulting items will be cached and returned after the updates on the pending update list have been processed. As nodes may be updated, they will be copied before being cached. The function can only be used together with updating expressions; if the function is called within a transform expression, its results will be discarded.

**Examples**

- `db:output("Prices have been deleted."), delete node //price` deletes all price elements in a database and returns an info message.

\[ \text{db:output-cache} \]

**Signatures**

```xml
db:output-cache() as item()*
```

**Summary**

Returns the items that have been cached by `db:output`. It can be used to check which items will eventually be returned as result of an updating function. This function is non-deterministic: It will return different results before and after items have been cached. It is e.g. useful when writing unit tests.

\[ \text{db:flush} \]

**Signatures**

```xml
db:flush($db as xs:string) as empty-sequence()
```

**Summary**

Explicitly flushes the buffers of the database `$db`. This command is only useful if AUTOFLUSH has been set to false.
Errors | BXDB0002: the addressed database does not exist or could not be opened.

**Helper Functions**

**db:name**

**Signatures**

\[
db:name($node as node()) as xs:string
\]

**Summary**

Returns the name of the database in which the specified database node $node is stored.

**Errors**

BXDB0001: $nodes contains a node which is not stored in a database.

**db:path**

**Signatures**

\[
db:path($node as node()) as xs:string
\]

**Summary**

Returns the path of the database document in which the specified database node $node is stored.

**Errors**

BXDB0001: $nodes contains a node which is not stored in a database.

**db:exists**

**Signatures**

\[
db:exists($db as xs:string) as xs:boolean
\]

**Examples**

- \[db:exists("DB")\] returns true if the database DB exists.
- \[db:exists("DB", "resource")\] returns true if resource is an XML document or a raw file.

**db:is-raw**

**Signatures**

\[
db:is-raw($db as xs:string, $path as xs:string) as xs:boolean
\]

**Errors**

BXDB0002: the addressed database does not exist or could not be opened.

**Examples**

- \[db:is-raw("DB", "music/01.mp3")\] returns true.

**db:is-xml**

**Signatures**

\[
db:is-xml($db as xs:string, $path as xs:string) as xs:boolean
\]

**Summary**

Checks if the specified resource in the database $db and the path $path exists, and if it is an XML document.

**Examples**

- \[db:is-xml("DB", "dir/doc.xml")\] returns true.

**db:content-type**

**Signatures**

\[
db:content-type($db as xs:string, $path as xs:string) as xs:string
\]

**Summary**

Retrieves the content type of a resource in the database $db and the path $path. The file extension is used to recognize the content-type of a resource stored in the database. Content-type application/xml will be returned for any XML document stored in the database, regardless of its file name extension.
Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXDB0001</td>
<td>The referenced XML node is no database node, i.e. it is neither stored in a database nor represented as database fragment.</td>
</tr>
<tr>
<td>BXDB0002</td>
<td>The addressed database does not exist or could not be opened.</td>
</tr>
<tr>
<td>BXDB0003</td>
<td>The addressed database is not persistent (stored on disk).</td>
</tr>
<tr>
<td>BXDB0004</td>
<td>The database lacks an index structure required by the called function.</td>
</tr>
<tr>
<td>BXDB0005</td>
<td>A query is expected to exclusively return database nodes of a single database.</td>
</tr>
<tr>
<td>BXDB0006</td>
<td>A database path addressed with doc() contains more than one document.</td>
</tr>
<tr>
<td>BXDB0007</td>
<td>A database cannot be updated because it is opened by another process.</td>
</tr>
<tr>
<td>BXDB0008</td>
<td>The specified database path is invalid.</td>
</tr>
<tr>
<td>BXDB0009</td>
<td>The addressed database id or pre value is out of range.</td>
</tr>
<tr>
<td>BXDB0011</td>
<td>The name of the specified database is invalid.</td>
</tr>
<tr>
<td>BXDB0012</td>
<td>A database can only be created once.</td>
</tr>
<tr>
<td>BXDB0013</td>
<td>The number of specified inputs and paths differs.</td>
</tr>
<tr>
<td>BXDB0014</td>
<td>Path points to a directory.</td>
</tr>
</tbody>
</table>

Examples

- `db:content-type("DB", "docs/doc01.pdf")` returns `application/pdf`.
- `db:content-type("DB", "docs/doc01.xml")` returns `application/xml`.
- `db:content-type("DB", "docs/doc01")` returns `application/xml` if `db:is-xml("DB", "docs/doc01")` returns `true`.

Changelog

Version 8.4

- Updated: `db:create, db:add, db:replace`: support for ADDCACHE option.
- Added: `db:token`

Version 8.3

- Updated: `db:list-details`: attributes with name of database and date of backup added to results.
- Updated: `db:backups` now include attributes with name of database and date of backup.
- Updated: Value Indexes: raise error if no index exists.

Version 8.2

- Added: `db:output-cache`
- Removed: `db:event`

Version 7.9

- Updated: parsing options added to `db:create, db:add` and `db:replace`.
- Updated: allow UPDINDEX if `$all` is `true`. 
Version 7.8.2

- Added: `db:alter, db:copy, db:create-backup, db:drop-backup, db:restore`

Version 7.8

- Removed: `db:fulltext` (use `ft:search` instead)

Version 7.7

- Added: `db:export, db:name, db:path`
- Updated: `$options` argument added to `db:create` and `db:optimize`.
- Updated: the functions no longer accept `Database Nodes` as reference. Instead, the name of a database must now be specified.

Version 7.6

- Updated: `db:create`: allow more than one input and path.

Version 7.5

- Updated: `db:add`: input nodes will be automatically converted to document nodes
- Added: `db:backups`
- Added: `db:create`
- Added: `db:drop`

Version 7.3

- Added: `db:flush`

Version 7.2.1

- Added: `db:text-range, db:attribute-range, db:output`

Version 7.1

- Added: `db:list-details, db:content-type`
- Updated: `db:info, db:system, db:retrieve`

Version 7.0

- Added: `db:retrieve, db:store, db:exists, db:is-raw, db:is-xml`
- Updated: `db:list, db:open, db:add`
Chapter 42. Fetch Module

Read this entry online in the BaseX Wiki.

This XQuery Module provides simple functions to fetch the content of resources identified by URIs. Resources can be stored locally or remotely and e.g. use the file:// or http:// scheme. If more control over HTTP requests is required, the HTTP Module can be used. With the HTML Module, retrieved HTML documents can be converted to XML.

Conventions

All functions in this module are assigned to the http://basex.org/modules/fetch namespace, which is statically bound to the fetch prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

Functions

fetch:binary

Signatures

<table>
<thead>
<tr>
<th>Signature</th>
</tr>
</thead>
<tbody>
<tr>
<td>fetch:binary($uri as xs:string) as xs:base64Binary</td>
</tr>
</tbody>
</table>

Summary

Fetches the resource referred to by the given URI and returns it as streamable xs:base64Binary.

Errors

BXFE0001: the URI could not be resolved, or the resource could not be retrieved.

Examples


fetch:text

Signatures

<table>
<thead>
<tr>
<th>Signature</th>
</tr>
</thead>
<tbody>
<tr>
<td>fetch:text($uri as xs:string) as xs:string fetch:text($uri as xs:string, $encoding as xs:string) as xs:string</td>
</tr>
</tbody>
</table>

Summary

Fetches the resource referred to by the given URI and returns it as streamable xs:string.

Errors

BXFE0001: the URI could not be resolved, or the resource could not be retrieved. Invalid XML characters will be ignored if the CHECKSTRINGS option is turned off. BXFE0002: the specified $encoding is not supported, or unknown.

Examples


fetch:xml

Signatures

<table>
<thead>
<tr>
<th>Signature</th>
</tr>
</thead>
<tbody>
<tr>
<td>fetch:xml($uri as xs:string) as document-node() fetch:xml($uri as xs:string, $options as map(*)) as document-node()</td>
</tr>
</tbody>
</table>

Summary

Fetches the resource referred to by the given $uri and returns it as an XML document. In contrast to fn:doc, each function call returns a different document node. As a consequence, document instances created by this function will not be kept in memory until the end of query evaluation. The $options argument can be used to change the parsing behavior. Allowed options are all parsing and XML parsing options in lower case.

Errors

BXFE0001: the URI could not be resolved, or the resource could not be retrieved.
Examples

The following expression returns an XML representation of the English Wikipedia main HTML page and chops all whitespace nodes:

```xml
fetch:xml("http://en.wikipedia.org", map { 'chop': true() })
```

**fetch:content-type**

**Signatures**

`fetch:content-type($uri as xs:string) as xs:string`

**Summary**

Returns the content-type (also called mime-type) of the resource specified by `$uri`:

- If a remote resource is addressed, the request header will be evaluated.
- If the addressed resource is locally stored, the content-type will be guessed based on the file extension.

**Errors**

`BXFE0001`: the URI could not be resolved, or the resource could not be retrieved.

**Examples**


**Errors**

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXFE0001</td>
<td>The URI could not be resolved, or the resource could not be retrieved.</td>
</tr>
<tr>
<td>BXFE0002</td>
<td>The specified encoding is not supported, or unknown.</td>
</tr>
</tbody>
</table>

**Changelog**

Version 8.0

- Added: `fetch:xml`

The module was introduced with Version 7.6.
Chapter 43. File Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions related to file system operations, such as listing, reading, or writing files.

This module is based on the EXPath File Module.

Conventions

All functions and errors in this module are assigned to the http://expath.org/ns/file namespace, which is statically bound to the file prefix.

For serialization parameters, the http://www.w3.org/2010/xslt-xquery-serialization namespace is used, which is statically bound to the output prefix.

Returned strings that refer to existing directories are suffixed with a directory separator. The error invalid-path is raised if a path is invalid.

Read Operations

file:list

<table>
<thead>
<tr>
<th>Signatures</th>
<th>file:list($dir as xs:string) as xs:string* file:list($dir as xs:string, $recursive as xs:boolean) as xs:string* file:list($dir as xs:string, $recursive as xs:boolean, $pattern as xs:string) as xs:string*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Lists all files and directories found in the specified $dir. The returned paths are relative to the provided path. The optional parameter $recursive specifies whether sub-directories will be traversed, too. The optional parameter $pattern defines a file name pattern in the Glob Syntax. If present, only those files and directories are returned that correspond to the pattern. Several patterns can be separated with a comma (,).</td>
</tr>
<tr>
<td>Errors</td>
<td>not-found: the specified file does not exist. no-dir: the specified path does not point to a directory. io-error: the operation fails for some other reason.</td>
</tr>
</tbody>
</table>

file:children

<table>
<thead>
<tr>
<th>Signatures</th>
<th>file:children($dir as xs:string) as xs:string*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the full paths to all files and directories found in the specified $dir. The inverse function is file:parent. The related function file:list returns relative file paths.</td>
</tr>
<tr>
<td>Errors</td>
<td>not-found: the specified file does not exist. no-dir: the specified path does not point to a directory. io-error: the operation fails for some other reason.</td>
</tr>
</tbody>
</table>

file:read-binary

<table>
<thead>
<tr>
<th>Signatures</th>
<th>file:read-binary($path as xs:string) as xs:base64Binary file:read-binary($path as xs:string, $offset as xs:integer) as xs:base64Binary file:read-binary($path as xs:string, $offset as xs:integer, $length as xs:integer) as xs:base64Binary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Reads the binary content of the file specified by $path and returns it as streamable xs:base64Binary. The optional parameters $offset and $length can be used to read chunks of a file.</td>
</tr>
<tr>
<td>Errors</td>
<td>not-found: the specified file does not exist. is-dir: the specified path is a directory. out-of-range: the offset or length is negative, or the chosen values would exceed the file bounds. io-error: the operation fails for some other reason.</td>
</tr>
</tbody>
</table>
Examples

- `stream:materialize(file:read-binary("config.data"))` returns a materialized representation of the streamable result.

**file:read-text**

**Signatures**

```
file:read-text($path as xs:string) as xs:string
file:read-text($path as xs:string, $encoding as xs:string) as xs:string
```

**Summary**

Reads the textual contents of the file specified by `$path` and returns it as streamable `xs:string`. The optional parameter `$encoding` defines the encoding of the file.

**Errors**

- `not-found`: the specified file does not exist.
- `is-dir`: the specified path is a directory.
- `unknown-encoding`: the specified encoding is not supported, or unknown.
- `io-error`: the operation fails for some other reason.

Invalid XML characters will be ignored if the `CHECKSTRINGS` option is turned off.

**Examples**

- `stream:materialize(file:read-text("config.txt"))` returns a materialized representation of the streamable result.

**file:read-text-lines**

**Signatures**

```
file:read-text-lines($path as xs:string) as xs:string
file:read-text-lines($path as xs:string, $encoding as xs:string) as xs:string*
```

**Summary**

Reads the textual contents of the file specified by `$path` and returns it as a sequence of `xs:string` items. The optional parameter `$encoding` defines the encoding of the file.

**Errors**

- `not-found`: the specified file does not exist.
- `is-dir`: the specified path is a directory.
- `unknown-encoding`: the specified encoding is not supported, or unknown.
- `io-error`: the operation fails for some other reason.

**Write Operations**

**file:create-dir**

**Signatures**

```
file:create-dir($dir as xs:string) as empty-sequence()
```

**Summary**

Creates the directory specified by `$dir`, including all non-existing parent directories.

**Errors**

- `exists`: a file with the same path already exists.
- `io-error`: the operation fails for some other reason.

**file:create-temp-dir**

**Signatures**

```
file:create-temp-dir($prefix as xs:string, $suffix as xs:string) as xs:string
file:create-temp-dir($prefix as xs:string, $suffix as xs:string, $dir as xs:string) as xs:string
```

**Summary**

Creates a new temporary directory that did not exist before this function was called, and returns its full file path. The directory name begins and ends with the specified `$prefix` and `$suffix`. If no directory is specified via `$dir`, the directory will be placed in the system’s default temporary directory. The operation will create all non-existing parent directories.

**Errors**

- `no-dir`: the specified directory points to a file.
- `io-error`: the directory could not be created.

**file:create-temp-file**

**Signatures**

```
file:create-temp-file($prefix as xs:string, $suffix as xs:string) as xs:string
file:create-temp-file($prefix as xs:string, $suffix as xs:string, $dir as xs:string) as xs:string
```

**Summary**

Creates a new temporary file that did not exist before this function was called, and returns its full file path. The file name begins and ends with the specified `$prefix` and `$suffix`. If no directory is
specified via $dir, the file will be placed in the system’s default temporary directory. The operation will create all non-existing parent directories.

<table>
<thead>
<tr>
<th>Errors</th>
<th>no-dir: the specified directory points to a file. io-error: the directory could not be created.</th>
</tr>
</thead>
</table>

**file:delete**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>file:delete($path as xs:string) as empty-sequence() file:delete($path as xs:string, $recursive as xs:boolean) as empty-sequence()</th>
</tr>
</thead>
</table>

**Summary**

Recursively deletes a file or directory specified by $path. The optional parameter $recursive specifies whether sub-directories will be deleted, too.

<table>
<thead>
<tr>
<th>Errors</th>
<th>not-found: the specified path does not exist. io-error: the operation fails for some other reason.</th>
</tr>
</thead>
</table>

**file:write**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>file:write($path as xs:string, $items as item()* as empty-sequence()) file:write($path as xs:string, $items as item()*, $params as item()) as empty-sequence()</th>
</tr>
</thead>
</table>

**Summary**

Writes a serialized sequence of items to the specified file. If the file already exists, it will be overwritten. The $params argument contains serialization parameters (see Serialization for more details), which can either be specified

- as children of an `<output:serialization-parameters/>` element, as defined for the fn:serialize() function; e.g.:

```
<output:serialization-parameters>
  <output:method value='xml'/>
  <output:cdata-section-elements value="div"/>
  ...
</output:serialization-parameters>
```

- as map, which contains all key/value pairs:

```
map { "method": "xml", "cdata-section-elements": "div", ... }
```

<table>
<thead>
<tr>
<th>Errors</th>
<th>no-dir: the parent of specified path is no directory. is-dir: the specified path is a directory. io-error: the operation fails for some other reason.</th>
</tr>
</thead>
</table>

**file:write-binary**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>file:write-binary($path as xs:string, $value as xs:anyAtomicType) as empty-sequence() file:write-binary($path as xs:string, $value as xs:integer) as empty-sequence()</th>
</tr>
</thead>
</table>

**Summary**

Writes a binary item (xs:base64Binary, xs:hexBinary) to the specified file. If the file already exists, it will be overwritten. If $offset is specified, data will be written at this file position. An existing file may be resized by that operation.

<table>
<thead>
<tr>
<th>Errors</th>
<th>no-dir: the parent of specified path is no directory. is-dir: the specified path is a directory. out-of-range: the offset is negative, or it exceeds the current file size. io-error: the operation fails for some other reason.</th>
</tr>
</thead>
</table>

**file:write-text**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>file:write-text($path as xs:string, $value as xs:string) as empty-sequence() file:write-text($path as xs:string, $value as xs:string, $encoding as xs:string) as empty-sequence()</th>
</tr>
</thead>
</table>
### Summary
Writes a string to the specified file. If the file already exists, it will be overwritten. The optional parameter $encoding defines the output encoding (default: UTF-8).

### Errors
- **no-dir**: the parent of specified path is no directory.
- **is-dir**: the specified path is a directory.
- **unknown-encoding**: the specified encoding is not supported, or unknown.
- **io-error**: the operation fails for some other reason.

---

#### file:write-text-lines

**Signatures**

```xml
file:write-text-lines($path as xs:string, $values as xs:string*) as empty-sequence()
file:write-text-lines($path as xs:string, $values as xs:string*, $encoding as xs:string) as empty-sequence()
```

**Summary**
Writes a sequence of strings to the specified file, each followed by the system specific newline character. If the file already exists, it will be overwritten. The optional parameter $encoding defines the output encoding (default: UTF-8).

**Errors**
- **no-dir**: the parent of specified path is no directory.
- **is-dir**: the specified path is a directory.
- **unknown-encoding**: the specified encoding is not supported, or unknown.
- **io-error**: the operation fails for some other reason.

---

#### file:append

**Signatures**

```xml
file:append($path as xs:string, $items as item()* as empty-sequence())
file:append($path as xs:string, $items as item()*, $params as item()) as empty-sequence()
```

**Summary**
Appends a serialized sequence of items to the specified file. If the file does not exist, a new file is created.

**Errors**
- **no-dir**: the parent of specified path is no directory.
- **is-dir**: the specified path is a directory.
- **io-error**: the operation fails for some other reason.

---

#### file:append-binary

**Signatures**

```xml
file:append-binary($path as xs:string, $value as xs:anyAtomicType) as empty-sequence()
```

**Summary**
Appends a binary item (xs:base64Binary, xs:hexBinary) to the specified file. If the file does not exist, a new one is created.

**Errors**
- **no-dir**: the parent of specified path is no directory.
- **is-dir**: the specified path is a directory.
- **io-error**: the operation fails for some other reason.

---

#### file:append-text

**Signatures**

```xml
file:append-text($path as xs:string, $value as xs:string) as empty-sequence()
file:append-text($path as xs:string, $value as xs:string, $encoding as xs:string) as empty-sequence()
```

**Summary**
Appends a string to a file specified by $path. If the specified file does not exist, a new file is created. The optional parameter $encoding defines the output encoding (default: UTF-8).

**Errors**
- **no-dir**: the parent of specified path is no directory.
- **is-dir**: the specified path is a directory.
- **unknown-encoding**: the specified encoding is not supported, or unknown.
- **io-error**: the operation fails for some other reason.

---

#### file:append-text-lines

**Signatures**

```xml
file:append-text-lines($path as xs:string, $values as xs:string*) as empty-sequence()
file:append-text-lines($path as xs:string, $values as xs:string*, $encoding as xs:string) as empty-sequence()
```
Summary
Appends a sequence of strings to the specified file, each followed by the system specific newline character. If the specified file does not exists, a new file is created. The optional parameter $encoding defines the output encoding (default: UTF-8).

Errors
no-dir: the parent of specified path is no directory.

is-dir: the specified path is a directory.

unknown-encoding: the specified encoding is not supported, or unknown.

io-error: the operation fails for some other reason.

copy

Signatures
file:copy($source as xs:string, $target as xs:string) as empty-sequence()

Summary
Copies a file or directory specified by $source to the file or directory specified by $target. If the target file already exists, it will be overwritten. No operation will be performed if the source and target path are equal.

Errors
not-found: the specified source does not exist.

exists: the specified source is a directory and the target is a file.

no-dir: the parent of the specified target is no directory.

io-error: the operation fails for some other reason.

move

Signatures
file:move($source as xs:string, $target as xs:string) as empty-sequence()

Summary
Moves or renames the file or directory specified by $source to the path specified by $target. If the target file already exists, it will be overwritten. No operation will be performed if the source and target path are equal.

Errors
not-found: the specified source does not exist.

exists: the specified source is a directory and the target is a file.

no-dir: the parent of the specified target is no directory.

io-error: the operation fails for some other reason.

File Properties

exists

Signatures
file:exists($path as xs:string) as xs:boolean

Summary
Returns an xs:boolean indicating whether a file or directory specified by $path exists in the file system.

is-dir

Signatures
file:is-dir($path as xs:string) as xs:boolean

Summary
Returns an xs:boolean indicating whether the argument $path points to an existing directory.

is-absolute

Signatures
file:is-absolute($path as xs:string) as xs:boolean

Summary
Returns an xs:boolean indicating whether the argument $path is absolute. The behavior of this function depends on the operating system: On Windows, an absolute path starts with the drive letter and a colon, whereas on Linux it starts with a slash.

is-file

Signatures
file:is-file($path as xs:string) as xs:boolean
Summary | Returns an xs:boolean indicating whether the argument $path points to an existing file.

file:last-modified

Signatures | file:last-modified($path as xs:string) as xs:dateTime
Summary | Retrieves the timestamp of the last modification of the file or directory specified by $path.
Errors | not-found: the specified path does not exist.

file:size

Signatures | file:size($file as xs:string) as xs:integer
Summary | Returns the size, in bytes, of the file specified by $file, or 0 for directories.
Errors | not-found: the specified file does not exist.

Path Functions

file:name

Signatures | file:name($path as xs:string) as xs:string
Summary | Returns the name of a file or directory specified by $path. An empty string is returned if the path points to the root directory.

file:parent

Signatures | file:parent($path as xs:string) as xs:string?
Summary | Returns the absolute path to the parent directory of a file or directory specified by $path. An empty sequence is returned if the path points to a root directory. The inverse function is file:children.
Examples | • file:parent(static-base-uri()) returns the directory of the current XQuery module.

file:path-to-native

Signatures | file:path-to-native($path as xs:string) as xs:string
Summary | Transforms the $path argument to its native representation on the operating system.
Errors | not-found: the specified file does not exist.

file:resolve-path

Signatures | file:resolve-path($path as xs:string) as xs:string file:resolve-path($path as xs:string, $base as xs:string) as xs:string
Summary | Transforms the $path argument to an absolute operating system path. If the path is relative, and if an absolute $base path is specified, it will be resolved against this path.
Errors | is-relative: the specified base path is relative.
Examples | The following examples apply to Windows:
• file:resolve-path('file.txt', 'C:/Temp/') returns C:/Temp/file.txt.
• file:resolve-path('file.txt', 'C:/Temp') returns C:/file.txt.
• file:resolve-path('file.txt', 'Temp') raises an error.
file:path-to-uri

Signatures  | file:path-to-uri($path as xs:string) as xs:string
Summary     | Transforms the path specified by $path into a URI with the file:// scheme.

System Properties

file:dir-separator

Signatures  | file:dir-separator() as xs:string
Summary     | Returns the directory separator used by the operating system, such as / or \.  

file:path-separator

Signatures  | file:path-separator() as xs:string
Summary     | Returns the path separator used by the operating system, such as ; or :.

file:line-separator

Signatures  | file:line-separator() as xs:string
Summary     | Returns the line separator used by the operating system, such as 
, or 
.

file:temp-dir

Signatures  | file:temp-dir() as xs:string
Summary     | Returns the system’s default temporary-file directory.

file:current-dir

Signatures  | file:current-dir() as xs:string
Summary     | Returns the current working directory. - This function returns the same result as the function call file:resolve-path()<null/> </td>

file:base-dir

Signatures  | file:base-dir() as xs:string?
Summary     | Returns the parent directory of the static base URI. If the Base URI property is undefined, the empty sequence is returned. - If a static base URI exists, and if points to a local file path, this function returns the same result as the expression file:parent(static-base-uri()).

Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>not-found</td>
<td>A specified path does not exist.</td>
</tr>
<tr>
<td>invalid-path</td>
<td>A specified path is invalid.</td>
</tr>
<tr>
<td>exists</td>
<td>A file with the same path already exists.</td>
</tr>
<tr>
<td>no-dir</td>
<td>The specified path does not point to a directory.</td>
</tr>
<tr>
<td>is-dir</td>
<td>The specified path is a directory.</td>
</tr>
<tr>
<td>is-relative</td>
<td>The specified path is relative (and must be absolute).</td>
</tr>
</tbody>
</table>
unknown-encoding | The specified encoding is not supported, or unknown.
---|---
out-of-range | The specified offset or length is negative, or the chosen values would exceed the file bounds.
io-error | The operation fails for some other reason specific to the operating system.

## Changelog

### Version 8.2
- Added: `file:is-absolute`
- Updated: `file:resolve-path`: base argument added

### Version 8.0

### Version 7.8
- Added: `file:parent`, `file:name`
- Updated: error codes; `file:read-binary`, `file:write-binary`: $offset and $length arguments added.
- Deleted: `file:base-name`, `file:dir-name`

### Version 7.7
- Updated: all returned strings that refer to existing directories will be suffixed with a directory separator.

### Version 7.3
- Added: `file:append-text`, `file:write-text`, `file:append-text-lines`, `file:write-text-lines`, `file:line-separator`
- Updated: `file:write-binary`, `file:append-binary`: output limited to a single value

### Version 7.2.1
- Updated: `file:delete`: $recursive parameter added to prevent sub-directories from being accidentally deleted.
- Fixed: `file:list` now returns relative instead of absolute paths.
Chapter 44. Full-Text Module

Read this entry online in the BaseX Wiki.

This XQuery Module extends the W3C Full Text Recommendation with some useful functions: The index can be directly accessed, fulltext results can be marked with additional elements, or the relevant parts can be extracted. Moreover, the score value, which is generated by the contains text expression, can be explicitly requested from items.

Conventions

All functions in this module are assigned to the http://basex.org/modules/ft namespace, which is statically bound to the ft prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

Functions

ft:search

<table>
<thead>
<tr>
<th>Signatures</th>
</tr>
</thead>
<tbody>
<tr>
<td>ft:search($db as xs:string, $terms as item()* as text()*)</td>
</tr>
<tr>
<td>ft:search($db as xs:string, $terms as item()<em>, $options as map(xs:string, item())) as text()</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Summary</th>
</tr>
</thead>
</table>
| Returns all text nodes from the full-text index of the database $db that contain the specified $terms. The options used for tokenizing the input and building the full-text will also be applied to the search terms. As an example, if the index terms have been stemmed, the search string will be stemmed as well. The $options argument can be used to control full-text processing. The following options are supported (the introduction on Full-Text processing gives you equivalent expressions in the XQuery Full-Text notation):

- mode: determines the mode how tokens are searched. Allowed values are any, any word, all, all words, and phrase. any is the default search mode.

- fuzzy: turns fuzzy querying on or off. Allowed values are true and false. By default, fuzzy querying is turned off.

- wildcards: turns wildcard querying on or off. Allowed values are true and false. By default, wildcard querying is turned off.

- ordered: requires that all tokens occur in the order in which they are specified. Allowed values are true and false. The default is false.

- content: specifies that the matched tokens need to occur at the beginning or end of a searched string, or need to cover the entire string. Allowed values are start, end, and entire. By default, the option is turned off.

- scope: defines the scope in which tokens must be located. The option has following sub options:

  - same: can be set to true or false. It specifies if tokens need to occur in the same or different units.

  - unit: can be sentence or paragraph. It specifies the unit for finding tokens.

  - window: sets up a window in which all tokens must be located. By default, the option is turned off. It has following sub options:

    - size: specifies the size of the window in terms of units.

    - unit: can be sentences, sentences or paragraphs. The default is words.
- **distance**: specifies the distance in which tokens must occur. By default, the option is turned off. It has following sub options:
  - **min**: specifies the minimum distance in terms of *units*. The default is 0.
  - **max**: specifies the maximum distance in terms of *units*. The default is #.
  - **unit**: can be words, sentences or paragraphs. The default is words.

**Errors**

- BXDB0002: The addressed database does not exist or could not be opened.
- BXDB0004: the index is not available.
- BXFT0001: the fuzzy and wildcard option cannot be both specified.

**Examples**

- `ft:search("DB", "QUERY")`: Return all text nodes of the database DB that contain the term QUERY.
- Return all text nodes of the database DB that contain the numbers 2010 and 2020:
  ```xml
  ft:search("DB", ("2010", "2020"), map { 'mode': 'all' })
  ```
- Return text nodes that contain the terms A and B in a distance of at most 5 words:
  ```xml
  ft:search("db", ("A", "B"), map {
    "mode": "all words",
    "distance": map {
      "max": "5",
      "unit": "words"
    }
  })
  ```
- Iterate over three databases and return all elements containing terms similar to Hello World in the text nodes:
  ```xml
  let $terms := "Hello Worlds"
  let $fuzzy := true()
  for $db in 1 to 3
    let $dbname := 'DB' || $db
    return ft:search($dbname, $terms, map { 'fuzzy': $fuzzy })/..
  ```

**ft:contains**

**Signatures**

- `ft:contains($input as item()*, $terms as item()*) as xs:boolean`
- `ft:contains($input as item()*, $terms as item()*, $options as map(xs:xstring, item())) as xs:boolean`

**Summary**

Checks if the specified `$input` items contain the specified `$terms`. The function does the same as the Full-Text expression `contains text`, but options can be specified more dynamically. The `$options` are the same as for `ft:search`, and the following ones in addition:

- **case**: determines how character case is processed. Allowed values are insensitive, sensitive, upper and lower. By default, search is case insensitive.
- **diacritics**: determines how diacritical characters are processed. Allowed values are insensitive and sensitive. By default, search is diacritical insensitive.
- **stemming**: determines if tokens are stemmed. Allowed values are true and false. By default, stemming is turned off.
- **language**: determines the language. This option is relevant for stemming tokens. All language codes are supported. The default language is en.

**Errors**

- BXFT0001: the fuzzy and wildcard option cannot be both specified.

**Examples**

- Checks if jack or john occurs in the input string John Doe:
ft:contains("John Doe", ("jack", "john"), map { "mode": "any" })

- Calls the function with stemming turned on and off:

(true(), false()) ! ft:contains("Häuser", "Haus", map { 'stemming': ., 'language':'de' })

**ft:mark**

**Signatures**

<table>
<thead>
<tr>
<th>Signature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ft:mark($nodes as node()<em>) as node()</em> ft:mark($nodes as node()<em>, $name as xs:string) as node()</em></td>
<td>Puts a marker element around the resulting $nodes of a full-text index request. The default name of the marker element is mark. An alternative name can be chosen via the optional $name argument. Please note that:</td>
</tr>
</tbody>
</table>

- the full-text expression that computes the token positions must be specified as argument of the ft:mark() function, as all position information is lost in subsequent processing steps. You may need to specify more than one full-text expression if you want to use the function in a FLWOR expression, as shown in Example 2.

- the XML node to be transformed must be an internal "database" node. The transform expression can be used to apply the method to a main-memory fragment, as shown in Example 3.

**Examples**

**Example 1:** The following query returns `<XML><mark>hello</mark> world</XML>`, if one text node of the database DB has the value "hello world":

ft:mark(db:open('DB')//*[text() contains text 'hello'] )

**Example 2:** The following expression loops through the first ten full-text results and marks the results in a second expression:

let $start := 1
let $end   := 10
let $term  := 'welcome'
for $ft in (db:open('DB')//*[text() contains text { $term }])[position() = $start to $end]
return element hit {
  ft:mark($ft[text() contains text { $term }])
}

**Example 3:** The following expression returns `<p><b>word</b></p>`:

copy $p := <p>word</p>
modify ()
return ft:mark($p[text() contains text 'word'], 'b')

**ft:extract**

**Signatures**

<table>
<thead>
<tr>
<th>Signature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ft:extract($nodes as node()<em>) as node()</em> ft:extract($nodes as node()<em>, $name as xs:string) as node()</em> ft:extract($nodes as node()<em>, $name as xs:string, $length as xs:integer) as node()</em></td>
<td>Extracts and returns relevant parts of full-text results. It puts a marker element around the resulting $nodes of a full-text index request and chops irrelevant sections of the result. The default element name of the marker element is mark. An alternative element name can be chosen via the optional $name argument. The default length of the returned text is 150 characters. An alternative length can be specified via the optional $length argument. Note that the effective text length may differ</td>
</tr>
</tbody>
</table>

**Summary**

Extracts and returns relevant parts of full-text results. It puts a marker element around the resulting $nodes of a full-text index request and chops irrelevant sections of the result. The default element name of the marker element is mark. An alternative element name can be chosen via the optional $name argument. The default length of the returned text is 150 characters. An alternative length can be specified via the optional $length argument. Note that the effective text length may differ...
from the specified text due to formatting and readability issues. For more details on this function, please have a look at ft:mark.

### Examples

- The following query may return `<XML>...<b>hello</b>...<XML>` if a text node of the database DB contains the string "hello world":

  ```xml
  ft:extract(db:open('DB')//*[text() contains text 'hello'], 'b', 1)
  ```

#### ft:count

**Signatures**

```xml
ft:count($nodes as node()) as xs:integer
```

**Summary**

Returns the number of occurrences of the search terms specified in a full-text expression.

**Examples**

- `ft:count//*[text() contains text 'QUERY']` returns the `xs:integer` value 2 if a document contains two occurrences of the string "QUERY".

#### ft:score

**Signatures**

```xml
ft:score($item as item()) as xs:double*
```

**Summary**

Returns the score values (0.0 - 1.0) that have been attached to the specified items. 0 is returned a value if no score was attached.

**Examples**

- `ft:score('a' contains text 'a')` returns the `xs:double` value 1.

#### ft:tokens

**Signatures**

```xml
ft:tokens($db as xs:string) as element(value)*
```

**Summary**

Returns all full-text tokens stored in the index of the database $db, along with their numbers of occurrences. If $prefix is specified, the returned nodes will be refined to the strings starting with that prefix. The prefix will be tokenized according to the full-text used for creating the index.

**Errors**

BXDB0002: The addressed database does not exist or could not be opened.

BXDB0004: the full-text index is not available.

**Examples**

Returns the number of occurrences for a single, specific index entry:

```xml
let $term := ft:tokenize($term)
return number(ft:tokens('db', $term)[. = $term]/@count)
```

#### ft:tokenize

**Signatures**

```xml
ft:tokenize($input as xs:string) as xs:string* ft:tokenize($input as xs:string, $options as map(xs:xstring, item())) as xs:string*
```

**Summary**

Tokenizes the given $input string, using the current default full-text options or the $options specified as second argument. The following options are available:

- **case**: determines how character case is processed. Allowed values are insensitive, sensitive, upper and lower. By default, search is case insensitive.

- **diacritics**: determines how diacritical characters are processed. Allowed values are insensitive and sensitive. By default, search is diacritical insensitive.

- **stemming**: determines is tokens are stemmed. Allowed values are true and false. By default, stemming is turned off.

- **language**: determines the language. This option is relevant for stemming tokens. All language codes are supported. The default language is en.
The $options argument can be used to control full-text processing.

**Examples**
- `ft:tokenize("No Doubt")` returns the two strings no and doubt.
- `ft:tokenize("École", map { 'diacritics': 'sensitive' })` returns the string école.
- `declare ft-option using stemming; ft:tokenize("GIFTS")` returns a single string gift.

### ft:normalize

**Signatures**
```
ft:normalize($input as xs:string) as xs:string*
```
```
ft:normalize($input as xs:string, $options as map(xs:xstring, item())) as xs:string*
```

**Summary**
Normalizes the given $input string, using the current default full-text options or the $options specified as second argument. The function expects the same arguments as `ft:tokenize`.

**Examples**
- `ft:tokenize("Häuser am Meer", map { 'case': 'sensitive' })` returns the string Hauser am Meer.

### Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXFT0001</td>
<td>Both wildcards and fuzzy search have been specified as search options.</td>
</tr>
</tbody>
</table>

### Changelog

#### Version 8.0
- Added: `ft:contains`, `ft:normalize`
- Updated: Options added to `ft:tokenize`

#### Version 7.8
- Added: `ft:contains`
- Updated: Options added to `ft:search`

#### Version 7.7
- Updated: the functions no longer accept Database Nodes as reference. Instead, the name of a database must now be specified.

#### Version 7.2
- Updated: `ft:search` (second argument generalized, third parameter added)

#### Version 7.1
- Added: `ft:tokens`, `ft:tokenize`
Chapter 45. Geo Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions that may be applied to geometry data conforming to the Open Geospatial Consortium (OGC) Simple Feature (SF) data model. It is based on the EXPath Geo Module and uses the JTS library.

Geometries introduced in GML 2 are: Point, LineString, LinearRing, Polygon, MultiPoint, MultiLineString, MultiPolygon, and MultiGeometry. All nodes queried by BaseX should be a valid geometry. The only geometry type which is not supported by BaseX right now is MultiGeometry.

Conventions

• This module is included in the complete distributions of BaseX (zip, exe, war).

• All functions are assigned to the http://expath.org/ns/geo namespace, which must be dynamically imported:

```xml
import module namespace geo = "http://expath.org/ns/geo";
...
```

• In this documentation, the namespace is bound to the geo prefix.

• All errors are assigned to the http://expath.org/ns/error namespace, which is statically bound to the experr prefix.

General Functions

**geo:dimension**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>geo:dimension($geometry as element(*)) as xs:integer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the dimension of the given geometry $geometry.</td>
</tr>
<tr>
<td>Errors</td>
<td>GEO0001: the given element is not recognized as a valid geometry. GEO0002: the given element cannot be read by reader for some reason.</td>
</tr>
</tbody>
</table>
| Example            | import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
geo:dimension(<gml:Point><gml:coordinates>1,2</gml:coordinates></gml:Point>) |

**geo:geometry-type**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>geo:geometry-type($geometry as element(*)) as xs:QName</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the name of the geometry type of given geometry $geometry, if the geometry is not recognized with an error massage.</td>
</tr>
<tr>
<td>Errors</td>
<td>GEO0001: the given element is not recognized as a valid geometry. GEO0002: the given element cannot be read by reader for some reason.</td>
</tr>
</tbody>
</table>
| Example            | import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
geo:geometry-type(<gml:Point><gml:coordinates>1,2</gml:coordinates></gml:Point>) |
### Geo:srid

**Signatures**

```
geo:srid($geometry as element(*)) as xs:integer
```

**Summary**

Returns the ID of the Spatial Reference System used by the given geometry $geometry. Spatial Reference System information is supported in the simple way defined in the SFS. A Spatial Reference System ID (SRID) is present in each Geometry object. Geometry provides basic accessor operations for this field, but no others. The SRID is represented as an integer (based on the OpenGIS Simple Features Specifications For SQL). Here is a difference between the EXPath Geo Module and the implementation in BaseX, since the specification return the URI.

**Errors**

GEO0001: the given element is not recognized as a valid geometry. GEO0002: the given element cannot be read by reader for some reason.

### Geo:envelope

**Signatures**

```
geo:envelope($geometry as element(*)) as element(*)
```

**Summary**

Returns the gml:Envelope of the given geometry $geometry. The envelope is the minimum bounding box of this geometry. If this Geometry is the empty geometry, returns an empty Point. If the Geometry is a point, returns a non-empty Point. Otherwise, returns a Polygon whose points are (minx, miny), (maxx, miny), (maxx, maxy), (minx, maxy), (minx, miny).

**Errors**

GEO0001: the given element is not recognized as a valid geometry. GEO0002: the given element cannot be read by reader for some reason.

**Example**

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
geo:envelope(<gml:LinearRing>
  <gml:coordinates>1,1 20,1 20,20 1,20 1,1</gml:coordinates>
</gml:LinearRing>)
```

### Geo:as-text

**Signatures**

```
geo:as-text($geometry as element(*)) as xs:string
```

**Summary**

Returns the WKT (Well-known Text) representation of the given geometry $geometry. The envelope is the minimum bounding box of this geometry.

**Errors**

GEO0001: the given element is not recognized as a valid geometry. GEO0002: the given element cannot be read by reader for some reason.

**Example**

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
geo:as-text(<gml:Point><gml:coordinates>1,2</gml:coordinates></gml:Point>)
```

### Geo:as-binary

**Signatures**

```
geo:as-binary($geometry as element(*)) as xs:base64Binary
```

**Summary**

Returns the WKB (Well-known Binary) representation of the given geometry $geometry.

**Errors**

GEO0001: the given element is not recognized as a valid geometry. GEO0002: the given element cannot be read by reader for some reason.

**Example**

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
geo:as-binary(<gml:Point><gml:coordinates>1,2</gml:coordinates></gml:Point>)
```
**Geo Module**

<table>
<thead>
<tr>
<th>Function</th>
<th>Signatures</th>
<th>Summary</th>
<th>Errors</th>
<th>Example</th>
</tr>
</thead>
</table>
| geo:is-simple | `geo:is-simple($geometry as element(*)) as xs:boolean` | Returns whether the given geometry is simple `$geometry` and does not have no anomalous geometric points (i.e. the geometry does not self-intersect and does not pass through the same point more than once (may be a ring)). | GEO0001: the given element is not recognized as a valid geometry. GEO0002: the given element cannot be read by reader for some reason. | ```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
geo:is-simple(
    <gml:MultiLineString>
        <gml:LineString>
            <gml:coordinates>1,1 0,0 2,1</gml:coordinates>
        </gml:LineString>
        <gml:LineString>
            <gml:coordinates>2,1 3,3 4,4</gml:coordinates>
        </gml:LineString>
    </gml:MultiLineString>
)
``` |
| geo:boundary | `geo:boundary($geometry as element(*)) as element(*)?` | Returns the boundary of the given geometry `$geometry`, in GML 2. The return value is a sequence of either gml:Point or gml:LinearRing elements as a GeometryCollection object. For a Point or MultiPoint, the boundary is the empty geometry, nothing is returned. | GEO0001: the given element is not recognized as a valid geometry. GEO0002: the given element cannot be read by reader for some reason. GEO0005: the output object cannot be written as an element by writer for some reason. | ```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
geo:boundary(
    <gml:LineString>
        <gml:coordinates>1,1 0,0 2,1</gml:coordinates>
    </gml:LineString>
)
``` |
| geo:num-geometries | `geo:num-geometries($geometry as element(*)) as xs:integer` | Returns the number of geometry in a geometry-collection `$geometry`, in GML. For the geometries which are not a collection, it returns the instant value 1. This function is implemented wider than the specification and accepts all types of geometries, while the specification limits it to the collection types (MultiPoint, MultiPolygon, ...). | GEO0001: the given element(s) is not recognized as a valid geometry (QName). GEO0002: the given element cannot be read by reader for some reason. | ```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
geo:num-geometries(
    <gml:MultiLineString>
)
``` |
### Geo:geometry-n

**Signatures**

`geo:geometry-n($geometry as element(*), $geoNumber as xs:integer) as element(*)`

**Summary**

Returns the Nth geometry in geometry-collection `$geometry`, in GML. For the geometries which are not a collection, it returns the geometry if `$geoNumber` is 1. This function is implemented wider than the specification and accepts all types of geometries, while the specification limits it to the collection types (MultiPoint, MultiPolygon, ...).

**Errors**

GEO0001: the given element(s) is not recognized as a valid geometry (QName). GEO0002: the given element cannot be read by reader for some reason. GEO0004: the input index of geometry is out of range. GEO0005: the output object cannot be written as an element by writer for some reason.

**Example**

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
let $line :=
  <gml:MultiLineString>
    <gml:LineString><gml:coordinates>1,1 0,0 2,1</gml:coordinates></gml:LineString>
    <gml:LineString><gml:coordinates>2,1 3,3 4,4</gml:coordinates></gml:LineString>
  </gml:MultiLineString>
return geo:geometry-n($line, 1)
```

### Geo:length

**Signatures**

`geo:length($geometry as element(*)) as xs:double`

**Summary**

Returns the length of the geometry `$geometry`. If the geometry is a point, zero value will be returned.

**Errors**

GEO0001: the given element(s) is not recognized as a valid geometry (QName). GEO0002: the given element cannot be read by reader for some reason.

**Example**

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
let $polygon :=
  <gml:Polygon>
    <gml:outerBoundaryIs>
      <gml:LinearRing>
        <gml:coordinates>1,1 2,1 2,2 1,2 1,1</gml:coordinates>
      </gml:LinearRing>
    </gml:outerBoundaryIs>
  </gml:Polygon>
return geo:length($polygon)
```

### Geo:num-points

**Signatures**

`geo:num-points($geometry as element(*)) as xs:integer`

**Summary**

Returns integer value of number of the points in the given geometry `$geometry`. It can be used not only for Lines, also any other geometry types, like MultiPolygon. For Point geometry it will
Geo Module

return 1. This is an implementation different from the EXPath geo specification, as it limits the input geometry type only to lines.

<table>
<thead>
<tr>
<th>Errors</th>
<th>GEO0001: the given element(s) is not recognized as a valid geometry (QName). GEO0002: the given element cannot be read by reader for some reason.</th>
</tr>
</thead>
</table>

| Example       | import module namespace geo = "http://expath.org/ns/geo";
|               | declare namespace gml='http://www.opengis.net/gml';
|               | geo:num-points(
|               |   <gml:LineString><gml:coordinates>2,1 3,3 4,4</gml:coordinates></gml:LineString>
|               | ) |

| geo:area      | **Signatures** geo:area($geometry as element(*)) as xs:double  
|               | **Summary** Returns the area of the given geometry $geometry. For points and line the return value will be zero.  
|               | **Errors** GEO0001: the given element(s) is not recognized as a valid geometry (QName). GEO0002: the given element cannot be read by reader for some reason.  
|               | **Example** import module namespace geo = "http://expath.org/ns/geo";
|               | declare namespace gml='http://www.opengis.net/gml';
|               | let $polygon :=
|               |   <gml:Polygon>
|               |     <gml:outerBoundaryIs>
|               |       <gml:LinearRing>
|               |         <gml:coordinates>1,1 2,1 2,2 1,2 1,1</gml:coordinates>
|               |       </gml:LinearRing>
|               |     </gml:outerBoundaryIs>
|               |   </gml:Polygon>
|               | return geo:area($polygon)  

| geo:centroid  | **Signatures** geo:centroid($geometry as element(*)) as element(*)  
|               | **Summary** Returns the mathematical centroid of the given geometry $geometry, as a gml:Point. Based on the definition, this point is not always on the surface of the geometry $geometry.  
|               | **Errors** GEO0001: the given element(s) is not recognized as a valid geometry (QName). GEO0002: the given element cannot be read by reader for some reason. GEO0005: the output object cannot be written as an element by writer for some reason.  
|               | **Example** import module namespace geo = "http://expath.org/ns/geo";
|               | declare namespace gml='http://www.opengis.net/gml';
|               | let $point :=
|               |   <gml:MultiPoint>
|               |     <gml:Point><gml:coordinates>1,1</gml:coordinates></gml:Point>
|               |     <gml:Point><gml:coordinates>10,10</gml:coordinates></gml:Point>
|               |     <gml:Point><gml:coordinates>2,2</gml:coordinates></gml:Point>
|               |   </gml:MultiPoint>
|               | return geo:centroid($point)  

| geo:point-on-surface | **Signatures** geo:point-on-surface($geometry as element(*)) as element(*)  
|                      | **Summary** Returns an interior point on the given geometry $geometry, as a gml:Point. It is guaranteed to be on surface. Otherwise, the point may lie on the boundary of the geometry.  

222
Errors

GEO0001: the given element(s) is not recognized as a valid geometry (QName).
GEO0002: the given element cannot be read by reader for some reason.
GEO0005: the output object cannot be written as an element by writer for some reason.

Example

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
geo:point-on-surface(
    <gml:Polygon>
        <gml:outerBoundaryIs>
            <gml:LinearRing>
                <gml:coordinates>1,1 2,1 2,2 1,2 1,1</gml:coordinates>
            </gml:LinearRing>
        </gml:outerBoundaryIs>
    </gml:Polygon>
)
```

Spatial Predicate Functions

geo:equals

Signatures

geo:equals($geometry1 as element(*), $geometry2 as element(*)) as xs:boolean

Summary

Returns whether geometry1 is spatially equal to geometry2.

Errors

GEO0001: the given element(s) is not recognized as a valid geometry (QName).
GEO0002: the given element cannot be read by reader for some reason.

Example

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
geo:equals(
    <gml:LineString><gml:coordinates>1,1 55,99 2,1</gml:coordinates></gml:LineString>,
    <gml:LineString><gml:coordinates>1,1 1,1 55,99 2,1</gml:coordinates></gml:LineString>
)
```

geo:disjoint

Signatures

geo:disjoint($geometry1 as element(*), $geometry2 as element(*)) as xs:boolean

Summary

Returns whether geometry1 is spatially disjoint from geometry2 (they have no point in common, they do not intersect each other, and the DE-9IM Intersection Matrix for the two geometries is FF*FF****).

Errors

GEO0001: the given element(s) is not recognized as a valid geometry (QName).
GEO0002: the given element cannot be read by reader for some reason.

Example

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
let $lines :=
    <gml:MultiLineString>
        <gml:LineString><gml:coordinates>1,1 0,0 2,1</gml:coordinates></gml:LineString>
        <gml:LineString><gml:coordinates>2,1 3,3 4,4</gml:coordinates></gml:LineString>
    </gml:MultiLineString>
let $line :=
    <gml:LineString><gml:coordinates>0,0 2,1 3,3</gml:coordinates></gml:LineString>
```
# Geo Module

| geo:disjoint($lines, $line) |

## geo:intersects

### Signatures
geo:intersects($geometry1 as element(*), $geometry2 as element(*)) as xs:boolean

### Summary
Returns whether geometry1 $geometry1 is spatially intersects $geometry2 $geometry2. This is true if disjoint function of the two geometries returns false.

### Errors
- GEO0001: the given element(s) is not recognized as a valid geometry (QName).
- GEO0002: the given element cannot be read by reader for some reason.

### Example
```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
let $lines :=
  <gml:MultiLineString>
    <gml:LineString><gml:coordinates>1,1 0,0 2,1</gml:coordinates></gml:LineString>
    <gml:LineString><gml:coordinates>2,1 3,3 4,4</gml:coordinates></gml:LineString>
  </gml:MultiLineString>
let $line :=
  <gml:LineString><gml:coordinates>0,0 2,1 3,3</gml:coordinates></gml:LineString>
return geo:intersects($lines, $line)
```

## geo:intersects

### Signatures
gml:intersects($geometry1 as element(*), $geometry2 as element(*)) as xs:boolean

### Summary
Returns whether geometry1 $geometry1 is spatially touches $geometry2 $geometry2.

### Errors
- GEO0001: the given element(s) is not recognized as a valid geometry (QName).
- GEO0002: the given element cannot be read by reader for some reason.

### Example
```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
let $line :=
  <gml:LinearRing><gml:coordinates>1,1 2,1 5,3 1,1</gml:coordinates></gml:LinearRing>
let $polygon :=
  <gml:Polygon>
    <gml:outerBoundaryIs>
      <gml:LinearRing><gml:coordinates>1,1 2,1 5,3 1,1</gml:coordinates></gml:LinearRing>
    </gml:outerBoundaryIs>
  </gml:Polygon>
return geo:intersects($line, $polygon)
```

## geo:crosses

### Signatures
gml:crosses($geometry1 as element(*), $geometry2 as element(*)) as xs:boolean

### Summary
Returns whether geometry1 $geometry1 is spatially crosses $geometry2 $geometry2. It means, if the geometries have some but not all interior points in common. Returns true if the DE-9IM intersection matrix for the two geometries is:

- $T*T******$ (for P/L, P/A, and L/A situations)
- $T*****T**$ (for L/P, A/P, and A/L situations)
- $0********$ (for L/L situations).

### Errors
- GEO0001: the given element(s) is not recognized as a valid geometry (QName).
- GEO0002: the given element cannot be read by reader for some reason.
## geo:within

**Signatures**

geo:within($geometry1 as element(*), $geometry2 as element(*)) as xs:boolean

**Summary**

Returns whether $geometry1 is spatially within $geometry2.

**Errors**

GEO0001: the given element(s) is not recognized as a valid geometry (QName). GEO0002: the given element cannot be read by reader for some reason.

**Example**

```
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
let $line :=
  <gml:LinearRing><gml:coordinates>1,1 2,1 5,3 1,1</gml:coordinates></gml:LinearRing>
let $polygon :=
  <gml:Polygon>
    <gml:outerBoundaryIs>
      <gml:LinearRing><gml:coordinates>1,1 2,1 5,3 1,1</gml:coordinates></gml:LinearRing>
    </gml:outerBoundaryIs>
  </gml:Polygon>
return geo:crosses($line, $polygon)
```

## geo:contains

**Signatures**

geo:contains($geometry1 as element(*), $geometry2 as element(*)) as xs:boolean

**Summary**

Returns whether $geometry1 spatially contains $geometry2. Returns true if within function of these two geometries also returns true.

**Errors**

GEO0001: the given element(s) is not recognized as a valid geometry (QName). GEO0002: the given element cannot be read by reader for some reason.

**Example**

```
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
let $point :=
  <gml:Point><gml:coordinates>1,1</gml:coordinates></gml:Point>
let $polygon :=
  <gml:Polygon>
    <gml:outerBoundaryIs>
      <gml:LinearRing><gml:coordinates>1,1 2,1 5,3 1,1</gml:coordinates></gml:LinearRing>
    </gml:outerBoundaryIs>
  </gml:Polygon>
return geo:contains($polygon, $point)
```
geo:overlaps

Signatures
geo:overlaps($geometry1 as element(*) , $geometry2 as element(*))
as xs:boolean

Summary
Returns whether $geometry1 overlaps $geometry2.

Errors
GEO0001: the given element(s) is not recognized as a valid geometry (QName).
GEO0002: the given element cannot be read by reader for some reason.

Example
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
let $polygon1 :=
<gml:Polygon>
  <gml:outerBoundaryIs>
    <gml:LinearRing>
      <gml:coordinates>1,1 20,1 20,20 30,20 30,30 1,30 1,1</gml:coordinates>
    </gml:LinearRing>
  </gml:outerBoundaryIs>
  <gml:innerBoundaryIs>
    <gml:LinearRing>
      <gml:coordinates>2,2 3,2 3,3 2,3 2,2</gml:coordinates>
    </gml:LinearRing>
  </gml:innerBoundaryIs>
  <gml:innerBoundaryIs>
    <gml:LinearRing>
      <gml:coordinates>10,10 19,10 19,19 10,19 10,10</gml:coordinates>
    </gml:LinearRing>
  </gml:innerBoundaryIs>
</gml:Polygon>
let $polygon2 :=
<gml:Polygon>
  <gml:outerBoundaryIs>
    <gml:LinearRing>
      <gml:coordinates>1,1 2,1 5,3 1,1</gml:coordinates>
    </gml:LinearRing>
  </gml:outerBoundaryIs>
</gml:Polygon>
return geo:overlaps($polygon1, $polygon2)

geo:relate

Signatures
geo:relate($geometry1 as element(*) , $geometry2 as element(*), $intersectionMatrix as xs:string) as xs:boolean

Summary
Returns whether relationships between the boundaries, interiors and exteriors of $geometry1 and $geometry2 match the pattern specified in $intersectionMatrix, which should have the length of 9 characters. The values in the DE-9IM can be T, F, *, 0, 1, 2. - T means the intersection gives a non-empty result. - F means the intersection gives an empty result. - * means any result. - 0, 1, 2 gives the expected dimension of the result (point, curve, surface)

Errors
GEO0001: the given element(s) is not recognized as a valid geometry (QName).
GEO0002: the given element cannot be read by reader for some reason.
GEO0006: the given matrix is invalid.

Example
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
let $point :=
<gml:Point><gml:coordinates>18,11</gml:coordinates></gml:Point>
let $polygon :=
<gml:Polygon>
  <gml:outerBoundaryIs>
Analysis Functions

geo:distance

Signatures geo:distance($geometry1 as element(*), $geometry2 as element(*)) as xs:double

Summary Returns the shortest distance, in the units of the spatial reference system of geometry1 $geometry1, between the geometries, where that distance is the distance between a point on each of the geometries.

Errors GEO0001: the given element(s) is not recognized as a valid geometry (QName). GEO0002: the given element cannot be read by reader for some reason.

Example
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
let $line :=
<gml:LinearRing>
  <gml:coordinates>10,400 20,200 30,100 20,100 10,400</gml:coordinates>
</gml:LinearRing>
let $polygon :=
<gml:Polygon>
  <gml:outerBoundaryIs>
    <gml:LinearRing>
      <gml:coordinates>10,10 20,10 30,40 20,40 10,10</gml:coordinates>
    </gml:LinearRing>
  </gml:outerBoundaryIs>
</gml:Polygon>
return geo:distance($line, $polygon)

geo:buffer

Signatures geo:buffer($geometry as element(*), $distance as xs:double) as element(*)

Summary Returns polygonal geometry representing the buffer by distance $distance of geometry $geometry a buffer area around this geometry having the given width, in the spatial reference system of geometry. The buffer of a Geometry is the Minkowski sum or difference of the geometry with a disc of radius abs(distance). The buffer is constructed using 8 segments per quadrant to represent curves.

Errors GEO0001: the given element(s) is not recognized as a valid geometry (QName). GEO0002: the given element cannot be read by reader for some reason. GEO0005: the output object cannot be written as an element by writer for some reason.

Example
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
let $polygon :=
<gml:Polygon>
  <gml:outerBoundaryIs>
    <gml:LinearRing>
      <gml:coordinates>10,10 20,10 30,40 20,40 10,10</gml:coordinates>
    </gml:LinearRing>
  </gml:outerBoundaryIs>
</gml:Polygon>
geo:convex-hull

Signatures | geo:convex-hull($geometry as element(*)) as element(*)
---|---
Summary | Returns the convex hull geometry of the given geometry $geometry in GML, or the empty sequence. Actually returns the object of smallest dimension possible.
Errors | GEO0001: the given element(s) is not recognized as a valid geometry (QName).GEO0002: the given element cannot be read by reader for some reason.GEO0005: the output object cannot be written as an element by writer for some reason.
Example | import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
geo:convex-hull(
    <gml:LinearRing>
    <gml:coordinates>10,400 20,200 30,100 20,100 10,400</gml:coordinates>
    </gml:LinearRing>
)
geo:difference

Signatures
geo:difference($geometry1 as element(*), $geometry2 as element(*)) as element(*)?

Summary
Returns the difference geometry of $geometry1 with $geometry2, in GML, or empty sequence if the difference is empty, as a set of point in $geometry1 and not included in $geometry2.

Errors
GEO0001: the given element(s) is not recognized as a valid geometry (QName).
GEO0002: the given element cannot be read by reader for some reason.
GEO0005: the output object cannot be written as an element by writer for some reason.

Example
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
let $point :=
  <gml:Point><gml:coordinates>1.00,1.00</gml:coordinates></gml:Point>
let $line :=
  <gml:LineString><gml:coordinates>2,1 3,3 4,4</gml:coordinates></gml:LineString>
return geo:difference($point, $line)

geo:sym-difference

Signatures
geo:sym-difference($geometry1 as element(*), $geometry2 as element(*)) as element(*)?

Summary
Returns the symmetric difference geometry of $geometry1 with $geometry2, in GML, or empty sequence if the difference is empty, as a set of point in one of the geometries and not included in the other.

Errors
GEO0001: the given element(s) is not recognized as a valid geometry (QName).
GEO0002: the given element cannot be read by reader for some reason.
GEO0005: the output object cannot be written as an element by writer for some reason.

Example
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
let $point :=
  <gml:Point><gml:coordinates>1.00,1.00</gml:coordinates></gml:Point>
let $line :=
  <gml:LineString><gml:coordinates>2,1 3,3 4,4</gml:coordinates></gml:LineString>
return geo:sym-difference($point, $line)

Functions Specific to Geometry Type

geo:x

Signatures
geo:x($point as element(*)) as xs:double

Summary
Returns the x coordinate of point $point. A point has to have an x coordinate.

Errors
GEO0001: the given element(s) is not recognized as a valid geometry (QName).
GEO0002: the given element cannot be read by reader for some reason.

Example
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
### geo:z

**Signatures**

```
geo:z($point as element(*)) as xs:double?
```

**Summary**

Returns the z coordinate of point $point. If the point does not have the y coordinate, 0 will be returned.

**Errors**

GEO0001: the given element(s) is not recognized as a valid geometry (QName). GEO0002: the given element cannot be read by reader for some reason.

**Example**

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
geo:z(<gml:Point><gml:coordinates>1.00,1.00,3.00</gml:coordinates></gml:Point>)
```

### geo:start-point

**Signatures**

```
geo:start-point($line as element(*)) as element(*)
```

**Summary**

Returns the starting point of the given line $line. $line has to be a single line, LineString or LinearRing.

**Errors**

GEO0001: the given element(s) is not recognized as a valid geometry (QName). GEO0002: the given element cannot be read by reader for some reason. GEO0003: the given element has to be a line. Other geometries are not accepted.

**Example**

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
geo:start-point(
    <gml:LineString><gml:coordinates>2,1 3,3 4,4</gml:coordinates></gml:LineString>
)
```

### geo:end-point

**Signatures**

```
geo:end-point($line as element(*)) as element(*)
```

**Summary**

Returns the ending point of the given line $line. $line has to be a single line, LineString or LinearRing.

**Errors**

GEO0001: the given element(s) is not recognized as a valid geometry (QName). GEO0002: the given element cannot be read by reader for some reason. GEO0003: the given element has to be a line. Other geometries are not accepted.

**Example**

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
geo:end-point(
    <gml:LineString><gml:coordinates>2,1 3,3 4,4</gml:coordinates></gml:LineString>
)
```
### geo:is-closed

**Signatures**

```
geo:is-closed($line as element(*)) as xs:boolean
```

**Summary**

Returns a boolean value that shows the line $line is a closed loop (start point and end point are the same) or not. $line has to be a line, as a geometry, LineString or LinearRing, and MultiLineString.

**Errors**

- GEO0001: the given element(s) is not recognized as a valid geometry (QName).
- GEO0002: the given element cannot be read by reader for some reason.
- GEO0003: the given element has to be a line. Other geometries are not accepted.

**Example**

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
geo:is-closed(
    <gml:LineString>
        <gml:coordinates>2,1 3,3 4,4</gml:coordinates>
    </gml:LineString>
)
```

### geo:is-ring

**Signatures**

```
geo:is-ring($line as element(*)) as xs:boolean
```

**Summary**

Returns a boolean value that shows the line $line is a ring (closed loop and single) or not. $line has to be a single line, as a geometry, LineString or LinearRing.

**Errors**

- GEO0001: the given element(s) is not recognized as a valid geometry (QName).
- GEO0002: the given element cannot be read by reader for some reason.
- GEO0003: the given element has to be a line. Other geometries are not accepted.

**Example**

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
geo:is-ring(
    <gml:LineString>
        <gml:coordinates>2,1 3,3 4,4</gml:coordinates>
    </gml:LineString>
)
```

### geo:point-n

**Signatures**

```
geo:point-n($line as element(*)) as element(*)
```

**Summary**

Returns the Nth point in the given line $geometry. $line has to be a single line, as a geometry, LineString or LinearRing.

**Errors**

- GEO0001: the given element(s) is not recognized as a valid geometry (QName).
- GEO0002: the given element cannot be read by reader for some reason.
- GEO0003: the given element has to be a line. Other geometries are not accepted.
- GEO0004: the the input index of geometry is out of range.
- GEO0005: the output object cannot be written as an element by writer for some reason.

**Example**

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
let $line :=
    <gml:LineString>
        <gml:coordinates>2,1 3,3 4,4</gml:coordinates>
    </gml:LineString>
```
### geo:exterior-ring

<table>
<thead>
<tr>
<th>Signatures</th>
<th>geo:exterior-ring($polygon as element(<em>)) as element(</em>)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the outer ring of the given polygon $geometry, as a gml:LineString.</td>
</tr>
<tr>
<td>Errors</td>
<td>GEO0001: the given element(s) is not recognized as a valid geometry (QName). GEO0002: the given element cannot be read by reader for some reason. GEO0003: the given element has to be a polygon. Other geometries are not accepted. GEO0005: the output object cannot be written as an element by writer for some reason.</td>
</tr>
</tbody>
</table>

### geo:num-interior-ring

<table>
<thead>
<tr>
<th>Signatures</th>
<th>geo:num-interior-ring($polygon as element(*)) as xs:integer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the number of interior rings in the given polygon $geometry.</td>
</tr>
<tr>
<td>Errors</td>
<td>GEO0001: the given element(s) is not recognized as a valid geometry (QName). GEO0002: the given element cannot be read by reader for some reason. GEO0003: the given element has to be a polygon. Other geometries are not accepted.</td>
</tr>
</tbody>
</table>

### geo:interior-ring-n

<table>
<thead>
<tr>
<th>Signatures</th>
<th>geo:interior-ring-n($polygon as element(<em>)) as element(</em>)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the outer ring of the given polygon $geometry, as a gml:LineString.</td>
</tr>
</tbody>
</table>
Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GEO0001</td>
<td>Unrecognized Geo type.</td>
</tr>
<tr>
<td>GEO0002</td>
<td>The input GML node cannot be read by GMLreader.</td>
</tr>
<tr>
<td>GEO0003</td>
<td>Input geometry is not an appropriate geometry for this function.</td>
</tr>
<tr>
<td>GEO0004</td>
<td>The input index is out of range.</td>
</tr>
<tr>
<td>GEO0005</td>
<td>The result geometry can not be written by GMLwriter.</td>
</tr>
<tr>
<td>GEO0006</td>
<td>The input matrix is invalid.</td>
</tr>
</tbody>
</table>

Changelog

The module was introduced with Version 7.6.
Chapter 46. Hashing Module

This XQuery Module provides functions that perform different hash operations.

Conventions

All functions in this module are assigned to the http://basex.org/modules/hash namespace, which is statically bound to the hash prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

Functions

hash:md5

Signatures

hash:md5($value as xs:anyAtomicType) as xs:base64Binary

Summary

Computes the MD5 hash of the given $value, which may be of type xs:string, xs:base64Binary, or xs:hexBinary.

Examples

• xs:hexBinary(hash:md5("BaseX")) returns 0D65185C9E296311C0A2200179E479A2.

• hash:md5(xs:base64Binary("")) returns 1B2M2Y8AsgTpgAmY7PhCfg=.

hash:sha1

Signatures

hash:sha1($value as xs:anyAtomicType) as xs:base64Binary

Summary

Computes the SHA-1 hash of the given $value, which may be of type xs:string, xs:base64Binary, or xs:hexBinary.

Examples

• xs:hexBinary(hash:sha1("BaseX")) returns 3AD5958F0F27D5AFFDCA2957560F121D0597A4ED.

• hash:sha1(xs:base64Binary("")) returns 2jmj7l5rSw0yVb/vlWAYkK/YBwk=.

hash:sha256

Signatures

hash:sha256($value as xs:anyAtomicType) as xs:base64Binary

Summary

Computes the SHA-256 hash of the given $value, which may be of type xs:string, xs:base64Binary, or xs:hexBinary.

Examples

• xs:hexBinary(hash:sha256("BaseX")) returns 15D570763DEB75D728BB69643392873B835C7C94A2F1E881909DA47662821A3.

• hash:sha256(xs:base64Binary("")) returns 47DEQpj8HBSa+/TImW+5JceuQeRkm5NmpJWZG3hSuFU=.

hash:hash

Signatures

hash:hash($value as xs:anyAtomicType, $algorithm as xs:string) as xs:base64Binary

Summary

Computes the hash of the given $value, using the specified $algorithm. The specified values may be of type xs:string, xs:base64Binary, or xs:hexBinary. The following three algorithms are supported: MD5, SHA-1, and SHA-256.

Errors

HASH0001: the specified hashing algorithm is unknown.
Hashing Module

Examples

- `xs:hexBinary(hash:hash("", "MD5"))` returns `D41D8CD98F00B204E9800998ECF8427E`.
- `hash:hash("", "")` raises an error.

Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HASH0001</td>
<td>The specified hash algorithm is unknown.</td>
</tr>
</tbody>
</table>

Changelog

The module was introduced with Version 7.3.
Chapter 47. Higher-Order Functions Module

Read this entry online in the BaseX Wiki.

This XQuery Module adds some useful higher-order functions, additional to the Higher-Order Functions provided by the official specification.

Conventions

All functions in this module are assigned to the http://basex.org/modules/hof namespace, which is statically bound to the hof prefix.

Functions

**hof:id**

**Signatures**

| hof:id($expr as item()*) as item()*

**Summary**

Returns its argument unchanged. This function isn't useful on its own, but can be used as argument to other higher-order functions.

**Examples**

- hof:id(1 to 5) returns 1 2 3 4 5
- With higher-order functions:

  ```xquery```
  ```
  let $sort := sort(?, hof:id#1)
  let $reverse-sort := sort(?, function($x) { -$x })
  return (  
    $sort((1, 5, 3, 2, 4)),
    '|' ,
    $reverse-sort((1, 5, 3, 2, 4))
  )
  returns: 1 2 3 4 5 | 5 4 3 2 1
  ```

**hof:const**

**Signatures**

| hof:const($expr as item()*, $ignored as item()*) as item()*

**Summary**

Returns its first argument unchanged and ignores the second. This function isn't useful on its own, but can be used as argument to other higher-order functions, e.g. when a function combining two values is expected and one only wants to retain the left one.

**Examples**

- hof:const(42, 1337) returns 42.
- With higher-order functions:

  ```xquery```
  ```
  let $zip-sum := function($f, $seq1, $seq2) {
    sum(for-each-pair($seq1, $seq2, $f))
  }
  let $sum-all := $zip-sum(function($a, $b) { $a + $b }, ?, ?)
  let $sum-left := $zip-sum(hof:const#2, ?, ?)
  return (
    $sum-all((1, 1, 1, 1, 1), 1 to 5),
    $sum-left((1, 1, 1, 1, 1), 1 to 5)
  )
  ```
• Another use-case: When inserting a key into a map, $f$ decides how to combine the new value with a possibly existing old one. \texttt{hof:const} here means ignoring the old value, so that's normal insertion.

let $\text{insert-with} := \text{function}(f, \text{map}, \text{k}, \text{v})$
\{ 
    let $\text{old} := \text{map}(\text{k})$
    let $\text{new} := \text{if}(\text{old}) \text{ then } f(\text{v}, \text{old}) \text{ else } \text{v}$
    \text{return map:merge}((\text{map}, \text{map:entry}(\text{k}, \text{new})))
\}

let $\text{map} := \text{map} \{ \text{foo}': 1 \}$
let $\text{add} := \text{insert-with}(\text{function}(\text{a}, \text{b}) \{ \text{a} + \text{b} \})$, ?, ?, ?
let $\text{ins} := \text{insert-with}(\text{hof:const#2}, ?, ?, ?)$
return (  
    $\text{add}(\text{map}, \text{foo}', 2)\text{('foo')},$
    $\text{ins}(\text{map}, \text{foo}', 42)\text{('foo')}$
)
returns 3 42

### \texttt{hof:fold-left1}

**Signatures**
\texttt{hof:fold-left1}($\text{seq}$ as item()+, $f$ as function(item()*, item()) as item()*) as item()*

**Summary**
Works the same as \texttt{fn:fold-left}, but doesn't need a seed, because the sequence must be non-empty.

**Examples**
• \texttt{hof:fold-left1}(1 to 10, function($a$, $b$) \{ $a + b$ \}) returns 55.
• \texttt{hof:fold-left1}(), function($a$, $b$) \{ $a + b$ \}) throws \texttt{XPTY0004}, because $\text{seq}$ has to be non-empty.

### \texttt{hof:until}

**Signatures**
\texttt{hof:until}($\text{pred}$ as function(item()*) as xs:boolean, $f$ as function(item()*) as item()*, $\text{start}$ as item()*) as item()*

**Summary**
Applies the function $f$ to the initial value $\text{start}$ until the predicate $\text{pred}$ applied to the result returns \texttt{true()}.  

**Examples**
• \texttt{hof:until}(function($x$) \{ $x \geq 1000$ \}, function($y$) \{ 2 * $y$ \}, 1) returns 1024.
• Calculating the square-root of a number by iteratively improving an initial guess:

\begin{verbatim}
let $\text{sqrt} := \text{function}(\text{x} as \text{xs:double}) as \text{xs:double} \{
    \text{hof:until}(  
        function($\text{res}$) \{ \text{abs}($\text{res} \ast \text{res} - \text{x}) < 0.00001 \},
        function($\text{guess}$) \{ ($\text{guess} + \text{x} \text{div} \text{guess}) \text{div} 2 \},
        $\text{x}$
    )
\}
return $\text{sqrt}(25)$
\end{verbatim}

returns 5.000000000053722.

### \texttt{hof:scan-left}

**Signatures**
\texttt{hof:scan-left}($\text{seq}$ as item()*, $\text{start}$ as item()*, $f$ as function(item()*, item()) as item()*) as item()*

**Summary**
This function is similar to \texttt{fn:fold-left}, but it returns a list of successive reduced values from the left. It is equivalent to:
### Higher-Order Functions Module

#### hof:scan-left

```xml
declare function hof:scan-left($seq, $acc, $f) {
  if(empty($seq)) then $acc else (
    $acc,
    hof:scan-left(tail($seq), $f($acc, head($seq)), $f)
  );
}
```

**Examples**  
- Returns triangular numbers:

  ```xml
  hof:scan-left(1 to 10, 0, function($a, $b) { $a + $b })
  ```

#### hof:take-while

**Signatures**

```xml
hof:take-while($seq as item()* as item(), $pred as function(item()) as xs:boolean) as item()*
```

**Summary**  
The function returns items of $seq as long as the predicate $pred is satisfied. It is equivalent to:

```xml
declare function hof:take-while($seq, $pred) {
  if(empty($seq) or not($pred(head($seq)))) then () else (
    head($seq),
    hof:take-while(tail($seq), $pred)
  );
}
```

**Examples**  
- Computes at most 100 random integers, but stops if an integer is smaller than 10:

  ```xml
  hof:take-while(  
    (1 to 100) ! random:integer(50),
    function($x) { $x >= 10 }
  )
  ```

#### hof:top-k-by

**Signatures**

```xml
hof:top-k-by($seq as item()*, $sort-key as function(item()) as item(), $k as xs:integer) as item()*
```

**Summary**  
The function returns the $k items in $seq that are greatest when sorted by the result of $f applied to the item. The function is a much more efficient implementation of the following scheme:

```xml
(for $x in $seq
 order by $sort-key($x) descending
 return $x
)[position() <= $k]
```

**Examples**  
- hof:top-k-by(1 to 1000, hof:id#1, 5) returns 1000 999 998 997 996
- hof:top-k-by(1 to 1000, function($x) { -$x }, 3) returns 1 2 3
- hof:top-k-by(<x a='1' b='2' c='3'/>&(*, xs:integer#1, 2)/node-name()) returns c b

#### hof:top-k-with

**Signatures**

```xml
hof:top-k-with($seq as item()*, $lt as function(item(), item()) as xs:boolean, $k as xs:integer) as item()*
```

**Summary**  
The function returns the $k items in $seq that are greatest when sorted in the order of the less-than predicate $lt. The function is a general version of hof:top-k-by($seq, $sort-key, $k).
**Examples**

- `hof:top-k-with(1 to 1000, function($a, $b) { $a lt $b }, 5)` returns 1000 999 998 997 996
- `hof:top-k-with(-5 to 5, function($a, $b) { abs($a) gt abs($b) }, 5)` returns 0 1 -1 2 -2

**Changelog**

Version 8.1

- Added: `hof:scan-left, hof:take-while`

Version 7.2

- Added: `hof:top-k-by, hof:top-k-with`
- Removed: `hof:iterate`

Version 7.0

- `module added`
Chapter 48. HTML Module

This XQuery Module provides functions for converting HTML to XML. Conversion will only take place if TagSoup is included in the classpath (see HTML Parsing for more details).

Conventions

All functions in this module are assigned to the http://basex.org/modules/html namespace, which is statically bound to the html prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

Functions

html:parser

Signatures

html:parser() as xs:string

Summary

Returns the name of the applied HTML parser (currently: TagSoup). If an empty string is returned, TagSoup was not found in the classpath, and the input will be treated as well-formed XML.

html:parse

Signatures

html:parse($input as xs:anyAtomicType) as document-node()

html:parse($input as xs:anyAtomicType, $options as map(xs:string, item())) as document-node()

Summary

Converts the HTML document specified by $input to XML, and returns a document node:

- The input may either be a string or a binary item (xs:hexBinary, xs:base64Binary).
- If the input is passed on in its binary representation, the HTML parser will try to automatically choose the correct encoding.

The $options argument can be used to set TagSoup Options.

Errors

BXHL0001: the input cannot be converted to XML.

Examples

Basic Example

The following query converts the specified string to an XML document node.

Query

```
html:parse("<html>")
```

Result

```
<html xmlns="http://www.w3.org/1999/xhtml"/>
```

Specifying Options

The next query creates an XML document without namespaces:

Query

```
```
html:parse("<a href='ok.html'/>", map { 'nons': true() })

Result

<html>
  <body>
    <a shape="rect" href="ok.html"/>
  </body>
</html>

Parsing Binary Input

If the input encoding is unknown, the data to be processed can be passed on in its binary representation. The HTML parser will automatically try to detect the correct encoding:

Query


Result

<html xmlns="http://www.w3.org/1999/xhtml" class="client-nojs" dir="ltr" lang="en">
  <head>
    <title>Wikipedia, the free encyclopedia</title>
    <meta charset="UTF-8"/>
...

Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXHL0001</td>
<td>The input cannot be converted to XML.</td>
</tr>
</tbody>
</table>

Changelog

The module was introduced with Version 7.6.
Chapter 49. HTTP Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains a single function to send HTTP requests and handle HTTP responses. The function `send-request` is based on the EXPath HTTP Client Module. It gives full control over the available request and response parameters. For simple GET requests, the Fetch Module may be sufficient.

Conventions

All functions in this module are assigned to the `http://expath.org/ns/http-client` namespace, which is statically bound to the `http` prefix. All errors are assigned to the `http://expath.org/ns/error` namespace, which is statically bound to the `exerr` prefix.

Functions

http:send-request

**Signatures**

```
http:send-request($request as element(http:request)?, $href as xs:string?, $bodies as item(*)*) as item()+
http:send-request($request as element(http:request)) as item()+
http:send-request($request as element(http:request)?, $href as xs:string?)
```

**Summary**

Sends an HTTP request and interprets the corresponding response. `$request` contains the parameters of the HTTP request such as HTTP method and headers. In addition to this it can also contain the URI to which the request will be sent and the body of the HTTP method. If the URI is not given with the parameter `$href`, its value in `$request` is used instead. The structure of `http:request` element follows the EXPath specification. Both basic and digest authentication is supported.

**Errors**

HC0001: an HTTP error occurred. HC0002: error parsing the entity content as XML or HTML. HC0003: with a multipart response, the override-media-type must be either a multipart media type or application/octet-stream. HC0004: the src attribute on the body element is mutually exclusive with all other attribute (except the media-type). HC0005: the request element is not valid. HC0006: a timeout occurred waiting for the response.

Examples

Status Only

Simple GET request. As the attribute `status-only` is set to true, only the response element is returned.

**Query:**

```
http:send-request(<http:request method='get' status-only='true'/>, 'http://basex.org')
```

**Result:**

```
<http:response status="200" message="OK">
  <http:header name="Date" value="Mon, 14 Mar 2011 20:55:53 GMT"/>
  <http:header name="Content-Length" value="12671"/>
  <http:header name="Set-Cookie" value="fe_typo_user=d10c9552f9a784d1a73f8b6ebdf5ce63; path=/"/>
  <http:header name="Connection" value="close"/>
  <http:header name="Content-Type" value="text/html; charset=utf-8"/>
  <http:header name="Server" value="Apache/2.2.16"/>
</http:response>
```
Google Homepage

Retrieve Google search home page. TagSoup must be contained in the class path in order to parse html.

Query:

http:send-request(<http:request method='get' href='http://www.google.com'/>)

Result:

The response content type can also be overwritten in order to retrieve HTML pages and other textual data as plain string (using text/plain) or in its binary representation (using application/octet-stream). With the http:header element, a custom user agent can be set. See the following example:

Query:

let $binary := http:send-request{
http:request method='get'
override-media-type='application/octet-stream'
href='http://www.google.com'>
http:header name="User-Agent" value="Opera"/>
</http:request>
}])
return try {
html:parse($binary)
catch * {
'Conversion to XML failed: ' || $err:description
}

SVG Data

Content-type ending with +xml, e.g. image/svg+xml.

Query:
HTTP Module

http:send-request(<http:request method='get' />, 'http://upload.wikimedia.org/wikipedia/commons/6/6b/Bitmap_VS_SVG.svg')

Result:

<http:response status="200" message="OK">
  <http:header name="ETag" value="W/11b6d-4ba15ed4"/>
  <http:header name="Date" value="Mon, 14 Mar 2011 19:17:10 GMT"/>
  <http:header name="Age" value="9260"/>
  <http:header name="Content-Type" value="image/svg+xml"/>
  <http:header name="Server" value="Sun-Java-System-Web-Server/7.0"/>
  <http:header name="Content-Length" value="72557"/>
  <http:header name="X-Cache-Lookup" value="MISS from knsq22.knams.wikimedia.org:80"/>
  <http:header name="X-Cache" value="MISS from knsq22.knams.wikimedia.org"/>
  <http:header name="Connection" value="keep-alive"/>
  <http:header name="X-Forwarded-For" value=""/>
  <http:header name="X-Real-IP" value=""/>
  <http:header name="User-Agent" value=""/>
  <http:header name="User-Agent" value="Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/58.0.3029.110 Safari/537.36"/>
  <http:body media-type="image/svg+xml"/>
</http:response>

<svg xmlns="http://www.w3.org/2000/svg" xmlns:xlink="http://www.w3.org/1999/xlink" version="1.1" width="1063" height="638">
  <defs>
    <linearGradient id="lg0">
      <stop stop-color="#3333ff" offset="0"/>
      <stop stop-color="#3f3fff" stop-opacity="0" offset="1"/>
    </linearGradient>
  
  </svg>

POST Request

POST request to the BaseX REST Service, specifying a username and password.

Query:

let $request :=
  <http:request href='http://localhost:8984/rest' method='post' username='admin' password='admin' send-authorization='true'>
    <http:body media-type='application/xml'>
      <query xmlns="http://basex.org/rest">
        <text><![CDATA[
          <html>
            for $i in 1 to 3
              return <div>Section {$i}</div>
          </html>
        ]]></text>
      </query>
    </http:body>
  </http:request>
return http:send-request($request)

Result:

  <http:header name="Content-Length" value="135"/>
  <http:header name="Content-Type" value="application/xml"/>
  <http:header name="Server" value="Jetty(6.1.26)"/>
  <http:header name="Server" value="Sun-Java-System-Web-Server/7.0"/>
  <http:header name="Content-Length" value="135"/>
  <http:header name="Content-Type" value="application/xml"/>
  <http:body media-type="application/xml"/>
</http:response>

<html>
  <div>Section 1</div>
  <div>Section 2</div>
</html>
Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HC0001</td>
<td>An HTTP error occurred.</td>
</tr>
<tr>
<td>HC0002</td>
<td>Error parsing the entity content as XML or HTML.</td>
</tr>
<tr>
<td>HC0003</td>
<td>With a multipart response, the override-media-type must be either a multipart media type or application/octet-stream.</td>
</tr>
<tr>
<td>HC0004</td>
<td>The src attribute on the body element is mutually exclusive with all other attribute (except the media-type).</td>
</tr>
<tr>
<td>HC0005</td>
<td>The request element is not valid.</td>
</tr>
<tr>
<td>HC0006</td>
<td>A timeout occurred waiting for the response.</td>
</tr>
</tbody>
</table>

Changelog

Version 8.0

- Added: digest authentication

Version 7.6

- Updated: http:send-request: HC0002 is raised if the input cannot be parsed or converted to the final data type.
- Updated: errors are using text/plain as media-type.
Chapter 50. Index Module

Read this entry online in the BaseX Wiki.

This XQuery Module provides functions for displaying information stored in the database index structures.

Conventions

All functions in this module are assigned to the http://basex.org/modules/index namespace, which is statically bound to the index prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

Functions

index:facets

Signatures

<table>
<thead>
<tr>
<th>Signature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>index:facets($db as xs:string) as xs:string</td>
<td>Returns information about all facets and facet values of the database $db in document structure format. If $type is specified as flat, the function returns this information in a flat summarized version. The returned data is derived from the Path Index.</td>
</tr>
</tbody>
</table>

Summary

Returns information about all facets and facet values of the database $db in document structure format. If $type is specified as flat, the function returns this information in a flat summarized version. The returned data is derived from the Path Index.

Errors

BXDB0002: The addressed database does not exist or could not be opened.

Examples

• index:facets("DB") returns information about facets and facet values on the database DB in document structure.
• index:facets("DB", "flat") returns information about facets and facet values on the database DB in a summarized flat structure.

index:texts

Signatures

<table>
<thead>
<tr>
<th>Signature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>index:texts($db as xs:string) as element(value)*</td>
<td>Returns all strings stored in the Text Index of the database $db, along with their number of occurrences. If $prefix is specified, the returned entries will be refined to the ones starting with that prefix. If $start and $ascending are specified, all nodes will be returned after or before the specified start entry.</td>
</tr>
</tbody>
</table>

Summary

Returns all strings stored in the Text Index of the database $db, along with their number of occurrences. If $prefix is specified, the returned entries will be refined to the ones starting with that prefix. If $start and $ascending are specified, all nodes will be returned after or before the specified start entry.

Errors

BXDB0002: The addressed database does not exist or could not be opened.BXDB0004: The index is not available.

index:attributes

Signatures

<table>
<thead>
<tr>
<th>Signature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>index:attributes($db as xs:string) as element(value)*</td>
<td>Returns all strings stored in the Attribute Index of the database $db, along with their number of occurrences. If $prefix is specified, the returned entries will be refined to the ones starting with that prefix. If $start and $ascending are specified, all nodes will be returned after or before the specified start entry.</td>
</tr>
</tbody>
</table>

Summary

Returns all strings stored in the Attribute Index of the database $db, along with their number of occurrences. If $prefix is specified, the returned entries will be refined to the ones starting with that prefix. If $start and $ascending are specified, all nodes will be returned after or before the specified start entry.

Errors

BXDB0002: The addressed database does not exist or could not be opened.BXDB0004: The index is not available.
Index Module

index:tokens

Signatures
index:tokens($db as xs:string) as element(value)*

Summary
Returns all strings stored in the Token Index of the database $db, along with their number of occurrences.

Errors
BXDB0002: The addressed database does not exist or could not be opened.
BXDB0004: the index is not available.

index:element-names

Signatures
index:element-names($db as xs:string) as element(value)*

Summary
Returns all element names stored in the Name Index of the database $db, along with their number of occurrences.

Errors
BXDB0002: The addressed database does not exist or could not be opened.

index:attribute-names

Signatures
index:attribute-names($db as xs:string) as element(value)*

Summary
Returns all attribute names stored in the Name Index of the database $db, along with their number of occurrences.

Errors
BXDB0002: The addressed database does not exist or could not be opened.

Changelog

Version 8.4
• Added: index:token

Version 7.7
• Updated: the functions no longer accept Database Nodes as reference. Instead, the name of a database must now be specified.

Version 7.3
• Updated: index:texts, index:attributes: signature with three arguments added.

The module was introduced with Version 7.1.
Chapter 51. Inspection Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions for extracting internal information about modules and functions and generating documentation.

Conventions

All functions in this module are assigned to the http://basex.org/modules/inspect namespace, which is statically bound to the inspect prefix. xqDoc document instances are assigned to the http://www.xqdoc.org/1.0 namespace, which is statically bound to the xqdoc prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

Reflection

inspect:functions

| Signatures | inspect:functions() as function(*)* inspect:functions($uri as xs:string) as function(*)* |
| Summary | Returns function items for all user-defined functions (both public and private) that are known in the current query context. If a $uri is specified, the addressed module will be compiled, and its functions will be added to the query context and returned to the user. |
| Examples | Invokes the declared functions and returns their values: |
| |
| declare %private function local:one() { 12 }; |
| declare %private function local:two() { 34 }; |
| for $f in inspect:functions() return $f() |

Compiles all functions in code.xqm and invokes the function named run:

| let $uri := 'code.xqm' |
| let $name := "run" |
| for $f in inspect:functions($uri) |
| where local-name-from-QName(function-name($f)) = $name |
| return $f() |

Documentation

inspect:function

| Signatures | inspect:function($function as function(*)) as element(function) |
| Summary | Inspects the specified $function and returns an element that describes its structure. The output of this function is similar to eXist-db’s inspect:inspect-function function. |
| Examples | The query inspect:function(count#1) yields: |
| |
| <function name="count" uri="http://www.w3.org/2005/xpath-functions"> |
| <argument type="item()" occurrence="*"/> |
| <return type="xs:integer"/> |
| </function> |

The function…
(::~
  : This function simply returns the specified integer.
  : @param $number number to return
  : @return specified number
):}
declare %private function local:same($number as xs:integer) as xs:integer {
  $number
};

...is represented by inspect:function(local:same#1) as...

<function name="local:same" uri="http://www.w3.org/2005/xquery-local-functions">
  <argument type="xs:integer" name="number">number to return</argument>
  <annotation name="private" uri="http://www.w3.org/2012/xquery"/>
  <description>This function simply returns the specified integer.</description>
  <return type="xs:integer">specified number</return>
</function>

**inspect:context**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>inspect:context() as element(context)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Generates an element that describes all variables and functions in the current query context.</td>
</tr>
<tr>
<td>Examples</td>
<td>Evaluate all user-defined functions with zero arguments in the query context:</td>
</tr>
<tr>
<td></td>
<td>inspect:context()/function ! function-lookup(QName(@uri, @name), 0) ! . ()</td>
</tr>
<tr>
<td></td>
<td>Return the names of all private functions in the current context:</td>
</tr>
</tbody>
</table>
|              | for $f in inspect:context()/function where $f/annotation/@name = 'private'
|              | return $f/@name/string() |

**inspect:module**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>inspect:module($uri as xs:string) as element(module)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Retrieves the resource located at the specified $uri, parses it as XQuery module, and generates an element that describes the module's structure.</td>
</tr>
<tr>
<td>Errors</td>
<td>FODC0002: the addressed resource cannot be retrieved.</td>
</tr>
<tr>
<td>Examples</td>
<td>An example is shown below.</td>
</tr>
</tbody>
</table>

**inspect:xqdoc**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>inspect:xqdoc($uri as xs:string) as element(xqdoc:xqdoc)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Retrives the resource located at the specified $uri, parses it as XQuery module, and generates an xqDoc element. xqDoc provides a simple vendor-neutral solution for generating documentation from XQuery modules. The documentation conventions have been inspired by the JavaDoc standard. Documentation comments begin with (::~ and end with :), and tags start with @. xqDoc comments can be specified for main and library modules and variable and function declarations. We have slightly extended the xqDoc conventions to do justice to the current status of XQuery (Schema: xqdoc-1.1.30052013.xsd):</td>
</tr>
</tbody>
</table>
Inspection Module

- an `<xqdoc:annotations/>` node is added to each variable or function that uses annotations. The `xqdoc:annotation` child nodes may have additional `xqdoc:literal` elements with `type` attributes (xs:string, xs:integer, xs:decimal, xs:double) and values.

- a single `<xqdoc:namespaces/>` node is added to the root element, which summarizes all prefixes and namespace URIs used or declared in the module.

- name and type elements are added to variables.

Errors

- FODC0002: the addressed resource cannot be retrieved.

Examples

An example is shown below.

Examples

This is the sample.xqm library module:

```xml
(:~
    This module provides some sample functions to demonstrate
    the features of the Inspection Module.
    :
    :@author BaseX Team
    :@see http://docs.basex.org/wiki/XQDoc_Module
    :@version 1.0
    :)
module namespace samples = 'http://basex.org/modules/samples';

(:~
    This is a sample string. :) 
declare variable $samples:test-string as xs:string := 'this is a string';

(:~
    This function simply returns the specified integer. 
    :@param $number number to return 
    :@return specified number 
    :)
declare %private function samples:same($number as xs:integer) as xs:integer { 
    $number 
};
```

If `inspect:module('sample.xqm')` is run, the following output will be generated:

```xml
<module prefix="samples" uri="http://basex.org/modules/samples">
    <description>This module provides some sample functions to demonstrate the features of the Inspection Module.</description>
    <author>BaseX Team</author>
    <see>http://docs.basex.org/wiki/XQDoc_Module</see>
    <version>1.0</version>
    <variable name="samples:test-string" uri="http://basex.org/modules/samples" type="xs:string">
        <description>This is a sample string.</description>
    </variable>
    <function name="samples:same" uri="http://basex.org/modules/samples">
        <argument name="number" type="xs:integer">number to return</argument>
        <annotation name="private" uri="http://www.w3.org/2012/xquery"/>
        <description>This function simply returns the specified integer.</description>
        <return type="xs:integer">specified number</return>
    </function>
</module>
```

The output looks as follows if `inspect:xqdoc('sample.xqm')` is called:

```xml
<xqdoc:xqdoc xmlns:xqdoc="http://www.xqdoc.org/1.0">
```
This module provides some sample functions to demonstrate the features of the Inspection Module.

**Function: samples:same**

This function simply returns the specified integer.

```xml
<xqdoc:function arity="1">
  <xqdoc:comment>
    This function simply returns the specified integer.
  </xqdoc:comment>
  <xqdoc:name>samples:same</xqdoc:name>
  <xqdoc:parameters>
    <xqdoc:parameter>
      <xqdoc:name>number</xqdoc:name>
      <xqdoc:type>xs:integer</xqdoc:type>
    </xqdoc:parameter>
  </xqdoc:parameters>
  <xqdoc:return>
    <xqdoc:type>xs:integer</xqdoc:type>
  </xqdoc:return>
</xqdoc:function>
```

**Changelog**

Version 7.9

- Updated: a query URI can now be specified with `inspect:functions`.

This module was introduced with Version 7.7.
Chapter 52. JSON Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions to parse and serialize JSON documents. JSON (JavaScript Object Notation) is a popular data exchange format for applications written in JavaScript. As there are notable differences between JSON and XML, no mapping exists that guarantees a lossless, bidirectional conversion between JSON and XML. For this reason, we offer various mappings, all of which are suited to different use cases.

Conventions

All functions in this module are assigned to the http://basex.org/modules/json namespace, which is statically bound to the json prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

Conversion Formats

Direct

The direct conversion format allows a lossless conversion from JSON to XML and back. The transformation is based on the following rules:

- The resulting document has a <json> root node.
- Object pairs are represented via elements. The name of a pair is rewritten to an element name:
  - Empty names are represented by a single underscore (_). Existing underscores are rewritten to two underscores (___), and characters that are not valid in element names are rewritten to an underscore and the character’s four-digit Unicode.
  - If the lax option is set to true, invalid characters are simply replaced with underscores or (when invalid as first character of an element name) prefixed with an underscore. The resulting names are better readable, but cannot always be converted back to their original form.
- Array entries are also represented via elements. _ is used as element name.
- Object and array values are stored in text nodes.
- The types of values are represented via type attributes:
  - The existing types are string, number, boolean, null, object, and array.
- As most values are strings, the string type is by default omitted.

Attributes

The attributes format is lossless, too. The transformation based on the following rules:

- The resulting document has a <json> root node.
- Object pairs are represented via <pair> elements. The name of a pair is stored in a name attribute.
- Array entries are represented via <item> elements.
- Object and array values are stored in text nodes.
- The types of values are represented via type attributes:
  - The existing types are string, number, boolean, null, object, and array.
As most values are strings, the string type is by default omitted.

Map

The map format is lossless, too. It converts a JSON document to an XQuery map and vice versa. The conversion rules are the same as for fn:parse-json.

Basic

The basic format is another lossless format. It converts a JSON document to an XML node and vice versa. The conversion rules are the same as for fn:json-to-xml.

JsonML

The jsonml format is designed to convert XML to JSON and back, using the JsonML dialect. JsonML allows the transformation of arbitrary XML documents, but namespaces, comments and processing instructions will be discarded in the transformation process. More details are found in the official JsonML documentation.

A little advice: in the Database Creation dialog of the GUI, if you select JSON Parsing and switch to the Parsing tab, you can see the effects of some of the conversion options.

Options

The following options are available (the Direction column indicates if an option applies to parsing, serialization, or both operations):

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
<th>Allowed</th>
<th>Default</th>
<th>Direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>format</td>
<td>Specifies the format for converting JSON data.</td>
<td>direct, attributes, jsonml, map</td>
<td>direct</td>
<td>parse, serialize</td>
</tr>
<tr>
<td>liberal</td>
<td>Determines if minor deviations from RFC 7159 will be ignored.</td>
<td>yes, no</td>
<td>no</td>
<td>parse</td>
</tr>
<tr>
<td>merge</td>
<td>This option is considered when direct or attributes conversion is used:</td>
<td>yes, no</td>
<td>no</td>
<td>parse, serialize</td>
</tr>
<tr>
<td></td>
<td>• If a name has the same type throughout the document, the type attribute will be omitted. Instead, the name will be listed in additional, type-specific attributes in the root node</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The attributes are named by their type in plural (numbers, booleans, nulls, objects and arrays), and the attribute value contains all names with that type, separated by whitespaces.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>strings</td>
<td>Indicates if type attributes will be added for strings.</td>
<td>yes, no</td>
<td>yes</td>
<td>parse, serialize</td>
</tr>
<tr>
<td>lax</td>
<td>Specifies if a lax approach is used to convert QNames to JSON names.</td>
<td>yes, no</td>
<td>no</td>
<td>parse, serialize</td>
</tr>
<tr>
<td>escape</td>
<td>Indicates if escaped characters are expanded (for example, \n becomes a single x0A character, while \u20AC becomes the character €).</td>
<td>yes, no</td>
<td>yes</td>
<td>parse</td>
</tr>
<tr>
<td>escape</td>
<td>Indicates if characters are escaped whenever the JSON syntax requires it. This option can be set to no if strings are already in escaped form and no further escaping is permitted.</td>
<td>yes, no</td>
<td>yes</td>
<td>serialize</td>
</tr>
</tbody>
</table>
Indent | Indicates if whitespace should be added to the output with the aim of improving human legibility. If the parameter is set as in the query prolog, it overrides the `indent serialization parameter`.
---|---
yes, no | yes | serialize

## Functions

### json:parse

#### Signatures

- `json:parse($input as xs:string) as element(json)`
- `json:parse($input as xs:string, $options as map(xs:string, xs:string)) as item()`

#### Summary

Converts the JSON document specified by `$input` to an XML document or a map. If the input can be successfully parsed, it can be serialized back to the original JSON representation. The `$options` argument can be used to control the way the input is converted.

#### Errors

- BXJS0001: the specified input cannot be parsed as JSON document.

### json:serialize

#### Signatures

- `json:serialize($input as item()? as xs:string)`
- `json:serialize($input as item(), $options as map(xs:string, xs:string)) as xs:string`

#### Summary

Serializes the node specified by `$input` as JSON, and returns the result as `xs:string` instance. The node is expected to conform to the output created by the `json:parse()` function. All other items will be serialized as specified for the `json output method` of the official specification. Items can also be serialized as JSON if the Serialization Parameter method is set to `json`. The `$options` argument can be used to control the way the input is serialized.

#### Errors

- BXJS0002: the specified node cannot be serialized as JSON document.

## Examples

### BaseX Format

#### Example 1: Adds all JSON documents in a directory to a database

**Query:**

```
let $database := "database"
for $name in file:list('.', false(), '*.json')
let $file := file:read-text($name)
let $json := json:parse($file)
return db:add($database, $json, $name)
```

#### Example 2: Converts a simple JSON string to XML and back

**Query:**

```
json:parse('{}')
```

**Result:**

```
<json type="object"/>
```

**Query:**

---

254
Example 3: Converts a JSON string with simple objects and arrays

Query:

```
json:parse('{
  "title": "Talk On Travel Pool",
  "link": "http://www.flickr.com/groups/talkontravel/pool/",
  "description": "Travel and vacation photos from around the world.",
  "modified": "2014-02-02T11:27Z",
  "generator": "http://www.flickr.com/
}
')
```

Result:

```
<json type="object">
  <title>Talk On Travel Pool</title>
  <link>http://www.flickr.com/groups/talkontravel/pool/</link>
  <description>Travel and vacation photos from around the world.</description>
  <modified>2014-02-02T11:27Z</modified>
  <generator>http://www.flickr.com/</generator>
</json>
```

Example 4: Converts a JSON string with different data types

Query:

```
let $options := map { 'merge': true() }
return json:parse('{
  "first_name": "John",
  "last_name": "Smith",
  "age": 25,
  "address": {
    "street": "21 2nd Street",
    "city": "New York",
    "code": 10021
  },
  "phone": [
    {
      "type": "home",
      "number": "212 555-1234"
    },
    {
      "type": "mobile",
      "number": 1327724623
    }
  ],
}', $options)
```

Result:

```
<json numbers="age code" arrays="phone" objects="json address value">
  <first__name>John</first__name>
</json>
```
<last_name>Smith</last_name>
<age>25</age>
<address>
  <street>21 2nd Street</street>
  <city>New York</city>
  <code>10021</code>
</address>
<phone>
  <type>home</type>
  <number>212 555-1234</number>
</phone>
<phone>
  <type>mobile</type>
  <number>1327724623</number>
</phone>
</json>

JsonML Format

Example 1: Converts all XML documents in a database to JsonML and writes them to disk

Query:

```xml
for $doc in collection('json')
let $name := document-uri($doc)
let $json := json:serialize($doc)
return file:write($name, $json)
```

Example 2: Converts a simple XML fragment to the JsonML format

Query:

```xml
json:serialize(<xml/>, map { 'format': 'jsonml' })
```

Result:

`["xml"]`

Example 3: Converts an XML document with elements and text

Query:

```xml
json:serialize(doc('flickr.xml'), map { 'format': 'jsonml' })
```

flickr.xml:

```xml
<flickr>
  <title>Talk On Travel Pool</title>
  <link>http://www.flickr.com/groups/talkontravel/pool/</link>
  <description>Travel and vacation photos from around the world.</description>
  <modified>2014-02-02T11:02:27Z</modified>
  <generator>http://www.flickr.com/</generator>
</flickr>
```

Result:

`["flickr",`
Example 4: Converts a document with nested elements and attributes

```json
Example 4: Converts a document with nested elements and attributes

Query:

```json
json:serialize(doc('input.xml'), map { 'format': 'jsonml' })
```input.xml:

```xml
<?xml version='1.0'?>
<address id='1'>
 <!-- comments will be discarded -->
 <last_name>Smith</last_name>
 <age>25</age>
 <address xmlns='will be dropped as well'>
  <street>21 2nd Street</street>
  <city>New York</city>
  <code>10021</code>
 </address>
 <phone type='home'>212 555-1234</phone>
</address>
```

Result:

```json
["address", {"id":"1"},
  ["last_name", "Smith"],
  ["age", "25"],
  ["address", ["street", "21 2nd Street"],
   ["city", "New York"],
   ["code", "10021"]],
  ["phone", {"type":"home"}, "212 555-1234"]]
```

Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXJS0001</td>
<td>The specified input cannot be parsed as JSON document.</td>
</tr>
<tr>
<td>BXJS0002</td>
<td>The specified node cannot be serialized as JSON document.</td>
</tr>
</tbody>
</table>

Changelog

Version 8.4

- Updated: unescape changed to escape.
Version 8.2

• Added: Conversion format basic.

Version 8.0

• Updated: Serialization aligned with the json output method of the official specification.

• Added: liberal option.

• Removed: spec option.

Version 7.8


• Updated: json:parse now returns a document node instead of an element, or an XQuery map if format is set to .map.

Version 7.7.2

• Updated: $options argument added to json:parse and json:serialize.

• Updated: json:parse-ml and json:serialize-ml are now deprecated.

The module was introduced with Version 7.0.
Chapter 53. Map Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions for manipulating maps, which will officially be introduced with XQuery 3.1. Please note that the functions are subject to change until the specification has reached its final stage.

Conventions

All functions in this module are assigned to the http://www.w3.org/2005/xpath-functions/map namespace, which is statically bound to the map prefix.

Functions

Some examples use the map $week defined as:


declare variable $week as map(*) := map {
  0: "Sonntag",
  1: "Montag",
  2: "Dienstag",
  3: "Mittwoch",
  4: "Donnerstag",
  5: "Freitag",
  6: "Samstag"
};

map:contains

Signatures

map:contains($input as map(*), $key as xs:anyAtomicType) as xs:boolean

Summary

Returns true if the map supplied as $input contains an entry with a key equal to the supplied value of $key; otherwise it returns false. No error is raised if the map contains keys that are not comparable with the supplied $key. If the supplied key is xs:untypedAtomic, it is compared as an instance of xs:string. If the supplied key is the xs:float or xs:double value NaN, the function returns true if there is an entry whose key is NaN, or false otherwise.

Examples

• map:contains($week, 2) returns true().
• map:contains($week, 9) returns false().
• map:contains(map {}, "xyz") returns false().
• map:contains(map { "xyz": 23 }, "xyz") returns true().

map:entry

Signatures

map:entry($key as xs:anyAtomicType, $value as item()* ) as map(*)

Summary

Creates a new map containing a single entry. The key of the entry in the new map is $key, and its associated value is $value. The function map:entry is intended primarily for use in conjunction with the function map:merge. For example, a map containing seven entries may be constructed like this:

map:merge(
  map:entry("Su", "Sunday"),
  map:entry("Mo", "Monday"),
  map:entry("Tu", "Tuesday"),
)
map:entry("We", "Wednesday"),
map:entry("Th", "Thursday"),
map:entry("Fr", "Friday"),
map:entry("Sa", "Saturday")
)

Unlike the map { ... } expression, this technique can be used to construct a map with a variable number of entries, for example:

map:merge(for $b in //book return map:entry($b/isbn, $b))

Examples

map:entry("M", "Monday") creates map { "M": "Monday" }.

map:for-each

Signatures

map:for-each($input as map(*), $fun as function(xs:anyAtomicType, item()*) as item()*) as item()*

Summary

Applies a function to every entry of the map $input and returns the results as a sequence. The function supplied as $fun takes two arguments. It is called supplying the key of the map entry as the first argument, and the associated value as the second argument.

Examples

The following query adds the keys and values of all map entries and returns (3, 7):

map:for-each(
  map { 1: 2, 3: 4 },
  function($a, $b) { $a + $b }
)

map:get

Signatures

map:get($input as map(*), $key as xs:anyAtomicType) as item()*

Summary

Returns the value associated with a supplied key in a given map. This function attempts to find an entry within the map supplied as $input that has a key equal to the supplied value of $key. If there is such an entry, the function returns the associated value; otherwise it returns an empty sequence. No error is raised if the map contains keys that are not comparable with the supplied $key. If the supplied key is xs:untypedAtomic, it is converted to xs:string. A return value of () from map:get could indicate that the key is present in the map with an associated value of (), or it could indicate that the key is not present in the map. The two cases can be distinguished by calling map:contains. Invoking the map as a function item has the same effect as calling get; that is, when $input is a map, the expression $input($K) is equivalent to get($input, $K). Similarly, the expression get(get(get($input, 'employee'), 'name'), 'first') can be written as $input('employee')('name')('first').

Examples

• map:get($week, 4) returns "Donnerstag".
• map:get($week, 9) returns (). (When the key is not present, the function returns an empty sequence.).
• map:get(map:entry(7, ()), 7) returns (). (An empty sequence as the result can also signify that the key is present and the associated value is an empty sequence.).

map:keys

Signatures

map:keys($input as map(*)) as xs:anyAtomicType*

Summary

Returns a sequence containing all the key values present in a map. The function takes any map as its $input argument and returns the keys that are present in the map as a sequence of atomic values. The order may differ from the order in which entries were inserted in the map.

Examples

• map:keys(map { 1: "yes", 2: "no" }) returns (1, 2).
### map:merge

**Signatures**
\[
\text{map:merge($input \text{ as map(*)*) as map(*)}}
\]

**Summary**
Constructs and returns a new map. The map is formed by combining the contents of the maps supplied in the $input argument. The maps are combined as follows:

1. There is one entry in the new map for each distinct key value present in the union of the input maps, where keys are considered distinct according to the rules of the distinct-values function.
2. The associated value for each such key is taken from the last map in the input sequence $input that contains an entry with this key.

There is no requirement that the supplied input maps should have the same or compatible types. The type of a map (for example map(xs:integer, xs:string)) is descriptive of the entries it currently contains, but is not a constraint on how the map may be combined with other maps.

**Examples**
- `map:merge(())` creates an empty map.
- `map:merge((map:entry(0, "no"), map:entry(1, "yes")))` creates `map { 0: "no", 1: "yes" }`.

### map:put

**Signatures**
\[
\text{map:put($input \text{ as map(*)}, $key as xs:anyAtomicType, $value as item*) as map(*)}}
\]

**Summary**
Creates a new map, containing the entries of the $input argument and a new entry composed by $key and $value. The semantics of this function are equivalent to `map:merge(($input, map { $key, $value }))`.

### map:remove

**Signatures**
\[
\text{map:remove($input \text{ as map(*)}, $key as xs:anyAtomicType) as map(*)}}
\]

**Summary**
Constructs a new map by removing an entry from an existing map. The entries in the new map correspond to the entries of $input, excluding any entry whose key is equal to $key. No failure occurs if the input map contains no entry with the supplied key; the input map is returned unchanged.

**Examples**

### map:size

**Signatures**
\[
\text{map:size($input \text{ as map(*)}) as xs:integer}}
\]

**Summary**
Returns a number of entries in the supplied map. The function takes any `map` as its $input argument and returns the number of entries that are present in the map.
Examples

- `map:size(map:merge(()))` returns 0.
- `map:size(map { "true": 1, "false": 0 })` returns 2.

Changelog

Version 8.4

- Removed: `map:serialize` (use `fn:serialize` instead)

Version 8.0

- Added: `map:for-each`, `map:merge`, `map:put`
- Removed: support for collations (in accordance with the XQuery 3.1 spec).
- Removed: `map:new` (replaced with `map:merge`)
- Updated: aligned with latest specification: compare keys of type `xs:untypedAtomic` as `xs:string` instances, store `xs:float` or `xs:double` value NaN.
- Introduction on maps is now found in the article on XQuery 3.1.

Version 7.8

- Updated: map syntax `map { 'key': 'value' }`
- Added: `map:serialize`

Version 7.7.1

- Updated: alternative map syntax without `map` keyword and `:` as key/value delimiter (e.g.: `{ 'key': 'value' }`)
Chapter 54. Math Module

Read this entry online in the BaseX Wiki.

The math XQuery Module defines functions to perform mathematical operations, such as pi, asin and acos. Most functions are specified in the Functions and Operators Specification of the upcoming XQuery 3.0 Recommendation, and some additional ones have been added in this module.

Conventions

All functions in this module are assigned to the http://www.w3.org/2005/xpath-functions/math namespace, which is statically bound to the math prefix.

W3 Functions

<table>
<thead>
<tr>
<th>Function</th>
<th>Signature</th>
<th>Summary</th>
<th>Examples</th>
</tr>
</thead>
</table>
| math:pi  | `math:pi()` as xs:double | Returns the xs:double value of the mathematical constant \( \pi \) whose lexical representation is \( 3.141592653589793 \). | - \( 2 \times \text{math:pi()} \) returns \( 6.283185307179586 \times 10^0 \).  
  - \( 60 \times (\text{math:pi()} \text{div} 180) \) converts an angle of 60 degrees to radians. |
| math:sqrt | `math:sqrt($arg as xs:double?)` as xs:double? | Returns the square root of \$arg\.
If \$arg\ is the empty sequence, the empty sequence is returned.
Otherwise the result is the xs:double value of the mathematical square root of \$arg. |
| math:sin | `math:sin($arg as xs:double?)` as xs:double? | Returns the sine of the $arg, expressed in radians.
If $arg is the empty sequence, the empty sequence is returned.
Otherwise the result is the sine of $arg, treated as an angle in radians. |
| math:cos | `math:cos($arg as xs:double?)` as xs:double? | Returns the cosine of $arg, expressed in radians.
If $arg is the empty sequence, the empty sequence is returned.
Otherwise the result is the cosine of $arg, treated as an angle in radians. |
If $arg is the empty sequence, the empty sequence is returned.
Otherwise the result is the tangent of $arg, treated as an angle in radians. |
| math:asin | `math:asin($arg as xs:double?)` as xs:double? | |
Summary
Returns the arc sine of $arg$. If $arg$ is the empty sequence, the empty sequence is returned. Otherwise the result is the arc sine of $arg$, returned as an angle in radians in the range $-\pi/2$ to $+\pi/2$.

**Math:acos**

Signatures
`math:acos($arg as xs:double?) as xs:double?`

Summary
Returns the arc cosine of $arg$. If $arg$ is the empty sequence, the empty sequence is returned. Otherwise the result is the arc cosine of $arg$, returned as an angle in radians in the range $0$ to $+\pi$.

**Math:atan**

Signatures
`math:atan($arg as xs:double?) as xs:double?`

Summary
Returns the arc tangent of $arg$. If $arg$ is the empty sequence, the empty sequence is returned. Otherwise the result is the arc tangent of $arg$, returned as an angle in radians in the range $-\pi/2$ to $+\pi/2$.

**Math:atan2**

Signatures
`math:atan2($arg1 as xs:double?, $arg2 as xs:double) as xs:double?`

Summary
Returns the arc tangent of $arg1$ divided by $arg2$, the result being in the range $-\pi/2$ to $+\pi/2$ radians. If $arg1$ is the empty sequence, the empty sequence is returned. Otherwise the result is the arc tangent of $arg1$ divided by $arg2$, returned as an angle in radians in the range $-\pi$ to $+\pi$.

**Math:pow**

Signatures
`math:pow($arg1 as xs:double?, $arg2 as xs:double) as xs:double?`

Summary
Returns $arg1$ raised to the power of $arg2$. If $arg1$ is the empty sequence, the empty sequence is returned. Otherwise the result is the $arg1$ raised to the power of $arg2$.

Examples
• `math:pow(2, 3)` returns 8.

**Math:exp**

Signatures
`math:exp($arg as xs:double?) as xs:double?`

Summary
Returns $e$ raised to the power of $arg$. If $arg$ is the empty sequence, the empty sequence is returned. Otherwise the result is the value of $e$ raised to the power of $arg$.

Examples
• `math:exp(1)` returns $e$.

**Math:log**

Signatures
`math:log($arg as xs:double?) as xs:double?`

Summary
Returns the natural logarithm of $arg$. If $arg$ is the empty sequence, the empty sequence is returned. Otherwise the result is the natural logarithm (base $e$) of $arg$.

Examples
• `math:log(math:e())` returns 1.

**Math:log10**

Signatures
`math:log10($arg as xs:double?) as xs:double?`

Summary
Returns the base 10 logarithm of $arg$. If $arg$ is the empty sequence, the empty sequence is returned. Otherwise the result is the base 10 logarithm of $arg$.

Examples
• `math:log10(100)` returns 2.
## Additional Functions

### math:e

<table>
<thead>
<tr>
<th>Signatures</th>
<th>math:e() as xs:double</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the <code>xs:double</code> value of the mathematical constant $e$ whose lexical representation is 2.718281828459045.</td>
</tr>
<tr>
<td>Examples</td>
<td>5*math:e() returns 13.591409142295225.</td>
</tr>
</tbody>
</table>

### math:sinh

<table>
<thead>
<tr>
<th>Signatures</th>
<th>math:sinh($arg as xs:double?) as xs:double?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the hyperbolic sine of $arg$. If $arg$ is the empty sequence, the empty sequence is returned. Otherwise the result is the hyperbolic sine of $arg$.</td>
</tr>
<tr>
<td>Examples</td>
<td>math:sinh(0) returns 0.</td>
</tr>
</tbody>
</table>

### math:cosh

<table>
<thead>
<tr>
<th>Signatures</th>
<th>math:cosh($arg as xs:double?) as xs:double?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the hyperbolic cosine of $arg$. If $arg$ is the empty sequence, the empty sequence is returned. Otherwise the result is the hyperbolic cosine of $arg$.</td>
</tr>
<tr>
<td>Examples</td>
<td>math:cosh(0) returns 1.</td>
</tr>
</tbody>
</table>

### math:tanh

<table>
<thead>
<tr>
<th>Signatures</th>
<th>math:tanh($arg as xs:double?) as xs:double?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the hyperbolic tangent of $arg$. If $arg$ is the empty sequence, the empty sequence is returned. Otherwise the result is the hyperbolic tangent of $arg$.</td>
</tr>
<tr>
<td>Examples</td>
<td>math:tanh(100) returns 1.</td>
</tr>
</tbody>
</table>

### math:crc32

<table>
<thead>
<tr>
<th>Signatures</th>
<th>math:crc32($str as xs:string) as xs:hexBinary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Calculates the CRC32 check sum of the given string $str$.</td>
</tr>
<tr>
<td>Examples</td>
<td>math:crc32(&quot;&quot;) returns '00000000'.</td>
</tr>
<tr>
<td></td>
<td>math:crc32(&quot;BaseX&quot;) returns '4C06FC7F'.</td>
</tr>
</tbody>
</table>

## Changelog

Version 7.5

- Moved: `math:random` and `math:uuid` have been moved to Random Module.

Version 7.3

- Added: `math:crc32` and `math:uuid` have been adopted from the obsolete Utility Module.
Chapter 55. Output Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions for simplifying formatted data output.

Conventions

All functions in this module are assigned to the http://basex.org/modules/out namespace, which is statically bound to the out prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

Functions

out:nl

<table>
<thead>
<tr>
<th>Signatures</th>
<th>out:nl() as xs:string</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns a single newline character (
).</td>
</tr>
</tbody>
</table>

out:tab

<table>
<thead>
<tr>
<th>Signatures</th>
<th>out:tab() as xs:string</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns a single tabulator character (	).</td>
</tr>
</tbody>
</table>

out:format

<table>
<thead>
<tr>
<th>Signatures</th>
<th>out:format($format as xs:string, $item1 as item(), ...) as xs:string</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns a formatted string. $item1 and all following items are applied to the $format string, according to Java’s printf syntax.</td>
</tr>
<tr>
<td>Errors</td>
<td>BXOU0001: The specified format is not valid.</td>
</tr>
<tr>
<td>Examples</td>
<td>• out:format(&quot;%b&quot;, true()) returns true.</td>
</tr>
<tr>
<td></td>
<td>• out:format(&quot;%06d&quot;, 256) returns 000256.</td>
</tr>
<tr>
<td></td>
<td>• out:format(&quot;%e&quot;, 1234.5678) returns 1.234568e+03.</td>
</tr>
</tbody>
</table>

Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXOU0001</td>
<td>The specified format is not valid.</td>
</tr>
</tbody>
</table>

Changelog

Introduced with Version 7.3. Functions have been adopted from the obsolete Utility Module.
Chapter 56. Process Module

Read this entry online in the BaseX Wiki.

This XQuery Module provides functions for executing system commands from XQuery.

Conventions

All functions in this module are assigned to the http://basex.org/modules/proc namespace, which is statically bound to the proc prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

Functions

**proc:system**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>proc:system($cmd as xs:string) as xs:string</th>
<th>proc:system($cmd as xs:string, $args as xs:string*) as xs:string</th>
<th>proc:system($cmd as xs:string, $args as xs:string*, $encoding as xs:string) as xs:string</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Executes the specified command in a separate process and returns the result as string. $cmd is the name of the command. Arguments to the command may be specified via $args. The result can be explicitly converted to a specified $encoding. If no encoding is specified, the system’s default encoding is used.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Errors</td>
<td>BXPRnnnn: If the command results in an error, an XQuery error will be raised. Its code will consist of the letters BXPR and four digits with the command’s exit code. BXPR9999: the specified encoding does not exist or is not supported.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Examples</td>
<td>• proc:system('date') returns the current date on a Linux system.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The following example returns &quot;Command not found&quot;, if the command &quot;xyz&quot; cannot be located or executed:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>try {</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>proc:system('xyz')</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>} catch bxerr:BXPR0002 {</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>'Command not found.'</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>}</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**proc:execute**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>proc:execute($cmd as xs:string) as element(result)</th>
<th>proc:execute($cmd as xs:string, $args as xs:string*) as element(result)</th>
<th>proc:execute($cmd as xs:string, $args as xs:string*, $encoding as xs:string) as element(result)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Executes the specified command in a separate process and returns the result as element. $cmd is the name of the command. Arguments to the command may be specified via $args. The result can be explicitly converted to a specified $encoding. If no encoding is specified, the system’s default encoding is used. A result has the following structure:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt;result&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt;output&gt;...result output...&lt;/output&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt;error&gt;...error output...&lt;/error&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt;code&gt;0&lt;/code&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt;/result&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Errors</td>
<td>BXPR9999: the specified encoding does not exist or is not supported.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Examples

• `proc:execute('dir', '\')` returns the files of the root directory of a Windows system.
• `proc:execute('ls', ('-l', '-a'))` executes the `ls -la` command on Unix systems.

`proc:property`

**Signatures**
`proc:property($name as xs:string) as xs:string?`

**Summary**
Returns the system property, specified by $name, or a context parameter of the web.xml file with that name (see Web Applications). An empty sequence is returned if the property does not exist. For environment variables of the operating system, please use `fn:environment-variable`.

**Examples**
• `map:merge(proc:property-names() ! map:entry(., proc:property(.)))` returns a map with all system properties.

`proc:property-names`

**Signatures**
`proc:property-names() as xs:string*`

**Summary**
Returns the names of all Java system properties and context parameters of the web.xml file (see Web Applications). For environment variables of the operating system, please use `fn:available-environment-variables`.

**Examples**
• `proc:property('java.runtime.version')` returns the version of the Java runtime engine.

**Errors**

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXPR9999</td>
<td>The specified encoding does not exist or is not supported.</td>
</tr>
</tbody>
</table>

**Changelog**

Version 8.3

• Added: `proc:property`, `proc:property-names`.

The module was introduced with Version 7.3.
# Chapter 57. Profiling Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains various testing, profiling and helper functions.

## Conventions

All functions in this module are assigned to the `http://basex.org/modules/prof` namespace, which is statically bound to the `prof` prefix. All errors are assigned to the `http://basex.org/errors` namespace, which is statically bound to the `bxerr` prefix.

## Functions

### prof:time

**Signatures**

- `prof:time($expr as item()) as item()*`
- `prof:time($expr as item(), $cache as xs:boolean) as item()*`
- `prof:time($expr as item(), $cache as xs:boolean, $label as xs:string) as item()*`

**Summary**

Measures the time needed to evaluate `$expr` and sends it to standard error or, if the GUI is used, to the Info View. If `$cache` is set to `true()`, the result will be temporarily cached. This way, a potential iterative execution of the expression (which often yields different memory usage) is blocked. A third, optional argument `$label` may be specified to tag the profiling result.

**Properties**

The function is *non-deterministic*; evaluation order will be preserved by the compiler.

**Examples**

- `prof:time("1 to 100000")` may output `25.69 ms`.
- `prof:time("1 to 100000", true())` may output `208.12 ms`.

### prof:mem

**Signatures**

- `prof:mem($expr as item()) as item()*`
- `prof:mem($expr as item(), $cache as xs:boolean) as item()*`
- `prof:mem($expr as item(), $cache as xs:boolean, $label as xs:string) as item()*`

**Summary**

Measures the memory allocated by evaluating `$expr` and sends it to standard error or, if the GUI is used, to the Info View. If `$cache` is set to `true()`, the result will be temporarily cached. This way, a potential iterative execution of the expression (which often yields different memory usage) is blocked. A third, optional argument `$label` may be specified to tag the profiling result.

**Properties**

The function is *non-deterministic*; evaluation order will be preserved by the compiler.

**Examples**

- `prof:mem("1 to 100000")` may output `0 Bytes`.
- `prof:mem("1 to 100000", true())` may output `26.678 mb`.

### prof:sleep

**Signatures**

- `prof:sleep($ms as xs:integer) as empty-sequence()`

**Summary**

Sleeps for the specified number of milliseconds.

**Properties**

The function is *non-deterministic*; evaluation order will be preserved by the compiler.

### prof:human

**Signatures**

- `prof:human($number as xs:integer) as xs:string`

**Summary**

Returns a human-readable representation of the specified `$number`. 
Example

• prof:human(16384) returns 16K.

prof:dump

Signatures

prof:dump($expr as item()) as empty-sequence() prof:dump($expr as item(), $label as xs:string) as empty-sequence()

Summary

Dumps a serialized representation of $expr to STDERR, optionally prefixed with $label, and returns an empty sequence. If the GUI is used, the dumped result is shown in the Info View.

Properties

In contrast to fn:trace(), the consumed expression will not be passed on.

prof:variables

Signatures

prof:variables() as empty-sequence()

Summary

Prints a list of all current local and global variable assignments to standard error or, if the GUI is used, to the Info View. As every query is optimized before being evaluated, not all of the original variables may be visible in the output. Moreover, many variables of function calls will disappear because functions are inlined. Function inlining can be turned off by setting the INLINELIMIT option to 0.

Properties

The function is non-deterministic: evaluation order will be preserved by the compiler.

Examples

• for $x in 1 to 2 return prof:variables() will dump the values of $x to standard error.

prof:current-ms

Signatures

prof:current-ms() as xs:integer

Summary

Returns the number of milliseconds passed since 1970/01/01 UTC. The granularity of the value depends on the underlying operating system and may be larger. For example, many operating systems measure time in units of tens of milliseconds.

Properties

In contrast to fn:current-time(), the function is non-deterministic, as it returns different values every time it is called. Its evaluation order will be preserved by the compiler.

prof:current-ns

Signatures

prof:current-ns() as xs:integer

Summary

Returns the current value of the most precise available system timer in nanoseconds.

Properties

In contrast to fn:current-time(), the function is non-deterministic, as it returns different values every time it is called. Its evaluation order will be preserved by the compiler.

prof:void

Signatures

prof:void($value as item()*) as empty-sequence()

Summary

Swallows all items of the specified $value and returns an empty sequence. This function is helpful if some code needs to be evaluated and if the actual result is irrelevant.

Properties

The function is non-deterministic: evaluation order will be preserved by the compiler.

Examples

• prof:void(fetch:binary('http://my.rest.service')) performs an HTTP request and ignores the result.

Changelog

Version 8.1

• Added: prof:variables
Profiling Module

Version 7.7
• Added: prof:void

Version 7.6
• Added: prof:human

Version 7.5
• Added: prof:dump, prof:current-ms, prof:current-ns

This module was introduced with Version 7.3.
## Chapter 58. Random Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions for computing random values. All functions except for random:seeded-double and random:seeded-integer are non-deterministic, i.e., they return different values for each call.

### Conventions

All functions in this module are assigned to the http://basex.org/modules/random namespace, which is statically bound to the random prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

### Functions

#### random:double

<table>
<thead>
<tr>
<th>Signatures</th>
<th>random:double() as xs:double</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns a double value between 0.0 (inclusive) and 1.0 (exclusive).</td>
</tr>
</tbody>
</table>

#### random:integer

<table>
<thead>
<tr>
<th>Signatures</th>
<th>random:integer() as xs:integer random:integer($max as xs:integer) as xs:integer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns an integer value, either in the whole integer range or between 0 (inclusive) and the given maximum (exclusive)</td>
</tr>
<tr>
<td>Errors</td>
<td>BXRA0001: the maximum value is out of bounds.</td>
</tr>
</tbody>
</table>

#### random:seeded-double

<table>
<thead>
<tr>
<th>Signatures</th>
<th>random:seeded-double($seed as xs:integer, $num as xs:integer) as xs:double*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns a sequence with $num double values between 0.0 (inclusive) and 1.0 (exclusive). The random values are created using the initial seed given in $seed.</td>
</tr>
</tbody>
</table>

#### random:seeded-integer

<table>
<thead>
<tr>
<th>Signatures</th>
<th>random:seeded-integer($seed as xs:integer, $num as xs:integer) as xs:integer* random:seeded-integer($seed as xs:integer, $num as xs:integer, $max as xs:integer) as xs:integer*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns a sequence with $num integer values, either in the whole integer range or between 0 (inclusive) and the given maximum (exclusive). The random values are created using the initial seed given in $seed.</td>
</tr>
<tr>
<td>Errors</td>
<td>BXRA0001: the maximum value is out of bounds. BXRA0002: the number of values to be returned is negative.</td>
</tr>
</tbody>
</table>

#### random:gaussian

<table>
<thead>
<tr>
<th>Signatures</th>
<th>random#gaussian($num as xs:integer) as xs:double*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns a sequence with $num double values. The random values are Gaussian (i.e. normally) distributed with the mean 0.0. and the derivation 1.0.</td>
</tr>
</tbody>
</table>
**random:uuid**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>random:uuid() as xs:string</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Summary</strong></td>
<td>Creates a random universally unique identifier (UUID), represented as 128-bit value.</td>
</tr>
<tr>
<td><strong>Examples</strong></td>
<td>* random:uuid() eq random:uuid() will (most probably) return the boolean value false.</td>
</tr>
</tbody>
</table>

**Errors**

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXRA0001</td>
<td>The specified maximum value is out of bounds.</td>
</tr>
<tr>
<td>BXRA0002</td>
<td>The specified number of values to be returned is negative.</td>
</tr>
</tbody>
</table>

**Changelog**

Version 8.0

- Updated: random:integer, random:seeded-integer raise error for invalid input.

The module was introduced with Version 7.5. It includes some functionality which was previously located in the Math Module.
Chapter 59. Repository Module

This XQuery Module contains functions for installing, listing and deleting modules contained in the Repository.

Conventions

All functions in this module are assigned to the http://basex.org/modules/repo namespace, which is statically bound to the repo prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

Functions

repo:install

Signatures
repo:install($path as xs:string) as empty-sequence()

Summary
Installs a package or replaces an existing package. The parameter $path indicates the path to the package.

Errors
BXRE0001: the package does not exist.
BXRE0002: a package uses an unsupported namespace URL.
BXRE0003: the package to be installed requires a package which is still not installed.
BXRE0004: the package descriptor is invalid.
BXRE0005: the module contained in the package to be installed is already installed as part of another package.
BXRE0006: the package cannot be parsed.
BXRE0009: the package version is not supported.
BXRE0010: the package contains an invalid JAR descriptor.
BXRE0011: the package contains a JAR descriptor but it cannot be read.

repo:delete

Signatures
repo:delete($pkg as xs:string) as empty-sequence()

Summary
Deletes a package. The parameter $pkg indicates either the package name as specified in the package descriptor or the name, suffixed with a hyphen and the package version.

Errors
BXRE0007: the package cannot be deleted.
BXRE0008: another package depends on the package to be deleted.

repo:list

Signatures
repo:list() as element(package)*

Summary
Lists the names and versions of all currently installed packages.

Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXRE0001</td>
<td>The addressed package does not exist.</td>
</tr>
<tr>
<td>BXRE0002</td>
<td>A package uses an unsupported namespace URI.</td>
</tr>
<tr>
<td>BXRE0003</td>
<td>The package to be installed requires a package which is not installed yet.</td>
</tr>
<tr>
<td>BXRE0004</td>
<td>The package descriptor is invalid.</td>
</tr>
<tr>
<td>BXRE0005</td>
<td>The module contained in the package to be installed is already installed as part of another package.</td>
</tr>
<tr>
<td>BXRE0006</td>
<td>The package cannot be parsed.</td>
</tr>
<tr>
<td>BXRE0007</td>
<td>The package cannot be deleted.</td>
</tr>
</tbody>
</table>
Repository Module

- BXRE0008 Another package depends on the package to be deleted
- BXRE0009 The package version is not supported.
- BXRE0010 The package contains an invalid JAR descriptor.
- BXRE0011 The package contains a JAR descriptor but it cannot be read.

**Changelog**

Version 7.2.1

- Updated: repo:install: existing packages will be replaced
- Updated: repo:delete: remove specific version of a package

Version 7.2

- Updated: repo:list now returns nodes

The module was introduced with Version 7.1.
Chapter 60. Request Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions for retrieving information on an HTTP request that has triggered the query. It is mainly useful in the context of Web Applications.

The module is related to Adam Retter’s EXQuery Request Module draft.

Conventions

• The basex-api package must be included in the classpath. This is always the case if you use one of the complete distributions (zip, exe, war) of BaseX.

• All functions are assigned to the http://exquery.org/ns/request namespace. The module must be imported in the query prolog:

  import module namespace request = "http://exquery.org/ns/request";

• In this documentation, the namespace is bound to the request prefix.

• The following example demonstrates what components a URI may consist of (the example is derived from RFC 3986):

  foo://example.com:8042/over/there?name=ferret
  
  scheme   hostname  port    path      query

General Functions

request:method

Signatures  request:method() as xs:string

Summary    Returns the Method of the HTTP request.

request:attribute

Signatures  request:attribute($name as xs:string) as xs:string

Summary    Returns the value of an attribute of the HTTP request. If the attribute does not exist, an empty sequence is returned.

Example    • request:attribute("javax.servlet.error.request_uri") returns the original URI of a caught error.
            • request:attribute("javax.servlet.error.message") returns the error message of a caught error.

URI Functions

request:scheme

Signatures  request:scheme() as xs:string
### Request Module

<table>
<thead>
<tr>
<th>Summary</th>
<th>Returns the Scheme component of the URI of an HTTP request.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example</td>
<td>For the example given in the introduction, this function would return foo.</td>
</tr>
</tbody>
</table>

**request:hostname**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>request:hostname() as xs:string</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the Hostname component of the URI of an HTTP request.</td>
</tr>
<tr>
<td>Example</td>
<td>For the example given in the introduction, this function would return example.com.</td>
</tr>
</tbody>
</table>

**request:port**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>request:port() as xs:integer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the Port component of the URI of an HTTP request, or a default port if it has not been explicitly specified in the URI.</td>
</tr>
<tr>
<td>Example</td>
<td>For the example given in the introduction, this function would return 8042.</td>
</tr>
</tbody>
</table>

**request:path**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>request:path() as xs:string</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the Path component of the URI of an HTTP request.</td>
</tr>
<tr>
<td>Example</td>
<td>For the example given in the introduction, this function would return /over/there.</td>
</tr>
</tbody>
</table>

**request:query**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>request:query() as xs:string?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the Query component of the URI of an HTTP request. If no query has been specified, an empty sequence is returned.</td>
</tr>
<tr>
<td>Example</td>
<td>For the example given in the introduction, this function would return name=ferret.</td>
</tr>
</tbody>
</table>

**request:uri**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>request:uri() as xs:anyURI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the complete URI of an HTTP request as it has been specified by the client.</td>
</tr>
<tr>
<td>Example</td>
<td>For the example given in the introduction, this method would return foo://example.com:8042/over/there?name=ferret.</td>
</tr>
</tbody>
</table>

**request:context-path**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>request:context-path() as xs:string</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the context of the request. For servlets in the default (root) context, this method returns an empty string.</td>
</tr>
</tbody>
</table>

### Connection Functions

**request:address**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>request:address() as xs:string</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the IP address of the server.</td>
</tr>
</tbody>
</table>
### request:remote-hostname

**Signatures**
`request:remote-hostname() as xs:string`

**Summary**
Returns the fully qualified hostname of the client that sent the request.

### request:remote-address

**Signatures**
`request:remote-address() as xs:string`

**Summary**
Returns the IP address of the client that sent the request.

### request:remote-port

**Signatures**
`request:remote-port() as xs:string`

**Summary**
Returns the TCP port of the client socket that triggered the request.

### Parameter Functions

### request:parameter-names

**Signatures**
`request:parameter-names() as xs:string*`

**Summary**
Returns the names of all query and form field parameters available from the HTTP request. With RESTXQ, this function can be used to access parameters that have not been statically bound by `%rest:query-param`.

**Example**
For the example given in the introduction, this function would return `name`.

### request:parameter

**Signatures**
`request:parameter($name as xs:string) as xs:string*`
`request:parameter($name as xs:string, $default as xs:string) as xs:string*`

**Summary**
Returns the value of the named query or form field parameter in an HTTP request. If the parameter does not exist, an empty sequence or the optionally specified default value is returned instead. If both query and form field parameters with the same name exist, the form field values will be attached to the query values.

**Example**
For the example given in the introduction, the function call `request:parameter('name')` would return `ferret`.

### Header Functions

### request:header-names

**Signatures**
`request:header-names() as xs:string*`

**Summary**
Returns the names of all headers available from the HTTP request. If RESTXQ is used, this function can be used to access headers that have not been statically bound by `%rest:header-param`.

### request:header

**Signatures**
`request:header($name as xs:string) as xs:string?`
`request:header($name as xs:string, $default as xs:string) as xs:string`

**Summary**
Returns the value of the named header in an HTTP request. If the header does not exist, an empty sequence or the optionally specified default value is returned instead.
## Cookie Functions

### request:cookie-names

**Signatures**  
`request:cookie-names() as xs:string*`

**Summary**  
Returns the names of all cookies in the HTTP headers available from the HTTP request. If RESTXQ is used, this function can be used to access cookies that have not been statically bound by `%rest:cookie-param`.

### request:cookie

**Signatures**  
`request:cookie($name as xs:string) as xs:string*`

```
request:cookie($name as xs:string, $default as xs:string) as xs:string
```

**Summary**  
Returns the value of the named Cookie in an HTTP request. If there is no such cookie, an empty sequence or the optionally specified default value is returned instead.

## Changelog

**Version 7.9**

- Updated: The returned values of `request:parameter-names`, `request:parameter` now also include form field parameters.

**Version 7.8**

- Added: `request:context-path`

**Version 7.7**

- Added: `request:attribute`

This module was introduced with Version 7.5.
Chapter 61. RESTXQ Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains helper functions for the RESTXQ API, some of which are defined in the RESTXQ Draft.

Conventions

• The basex-api package must be included in the classpath. This is always the case if you use one of the complete distributions (zip, exe, war) of BaseX.

• All functions are assigned to the http://exquery.org/ns/restxq namespace. The module must be imported in the query prolog:

```xml
import module namespace rest = "http://exquery.org/ns/restxq";
```

• In this documentation, the namespace is bound to the rest prefix, and the http://wadl.dev.java.net/2009/02 namespace is bound to the wadl prefix.

• If any of the functions is called outside the servlet context, the error BXSE0003: is raised.

General Functions

rest:base-uri

<table>
<thead>
<tr>
<th>Signatures</th>
<th>rest:base-uri() as xs:anyURI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>This function returns the implementation-defined base URI of the resource function.</td>
</tr>
</tbody>
</table>

rest:uri

<table>
<thead>
<tr>
<th>Signatures</th>
<th>rest:uri() as xs:anyURI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>This function returns the complete URI that addresses the Resource Function. This is the result of rest:base-uri appended with the path from the path annotation of the resource function.</td>
</tr>
</tbody>
</table>

rest:wadl

<table>
<thead>
<tr>
<th>Signatures</th>
<th>rest:wadl() as element(wadl:application)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>This (unofficial) function returns a WADL description of all available REST services.</td>
</tr>
</tbody>
</table>

Changelog

This module was introduced with Version 7.7.
Chapter 62. Session Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions for accessing and modifying server-side session information. This module is mainly useful in the context of Web Applications.

Conventions

- The basex-api package must be included in the classpath. This is always the case if you use one of the complete distributions (zip, exe, war) of BaseX.

- All functions are assigned to the http://basex.org/modules/session namespace. The module must be imported in the query prolog:

  ```xml
  import module namespace session = "http://basex.org/modules/session";
  ...
  ```

- In this documentation, the namespace is bound to the session prefix.

- Errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

- If any of the functions is called outside the servlet context, the error BXSE0003: is raised.

- As sessions are side-effecting operations, all functions are flagged as non-deterministic. This means that the functions will not be reordered by the compiler.

Functions

session:id

**Signatures**  
`session:id() as xs:string`

**Summary**  
Returns the session ID of a servlet request.

**Examples**  
Running the server-side XQuery file `id.xq` via `http://localhost:8984/id.xq`:

```xml
import module namespace session = "http://basex.org/modules/session";
'Session ID: ' || session:id()
```

session:created

**Signatures**  
`session:created() as xs:dateTime`

**Summary**  
Returns the creation time of a session.

session:accessed

**Signatures**  
`session:accessed() as xs:dateTime`

**Summary**  
Returns the last access time of a session.

session:names

**Signatures**  
`session:names() as xs:string*`

**Summary**  
Returns the names of all variables bound to the current session.
Examples
Running the server-side XQuery file names.xq via http://localhost:8984/names.xq:

```xquery
import module namespace session = "http://basex.org/modules/session";
session:names() ! element variable { . }
```

### session:get

**Signatures**

```
session:get($key as xs:string) as item()*
```

```
session:get($key as xs:string, $default as item()) as item()*
```

**Summary**
Returns the value of a variable bound to the current session. If the key is unknown, an empty sequence or the optionally specified default value is returned instead.

**Errors**

- BXSE0002: the value of a session variable could not be retrieved.

**Examples**
Running the server-side XQuery file get.xq via http://localhost:8984/get.xq?
key=user:

```xquery
import module namespace session = "http://basex.org/modules/session";
'Value of ' || $key || ': ' || session:get($key)
```

### session:set

**Signatures**

```
session:set($key as xs:string, $value as item()) as empty-sequence()
```

**Summary**
Binds the specified key/value pair to a session.

**Errors**

- BXSE0001: a function item was specified as value of a session variable.

**Examples**
Running the server-side XQuery file set.xq via http://localhost:8984/set.xq?
key=user&value=john:

```xquery
import module namespace session = "http://basex.org/modules/session";
session:set($key, $value), 'Variable was set.'
```

### session:delete

**Signatures**

```
session:delete($key as xs:string) as empty-sequence()
```

**Summary**
Deletes a session variable.

**Examples**
Running the server-side XQuery file delete.xq via http://localhost:8984/delete.xq?
key=user:

```xquery
import module namespace session = "http://basex.org/modules/session";
session:delete($key), 'Variable was deleted.'
```

### session:close

**Signatures**

```
session:close() as empty-sequence()
```

**Summary**
Unregisters a session and all data associated with it.

**Errors**

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXSE0001</td>
<td>A function item was specified as value of a session attribute.</td>
</tr>
<tr>
<td>BXSE0002</td>
<td>An error occurred while retrieving the value of a session attribute.</td>
</tr>
</tbody>
</table>
BXSE0003 | A function was called outside the servlet context.

**Changelog**

Version 8.0

- Updated: Allow sequences as session values.

This module was introduced with Version 7.5.
Chapter 63. Sessions Module

Read this entry online in the BaseX Wiki.

This XQuery Module can only be called from users with Admin permissions. It contains functions for accessing and modifying all registered server-side sessions. This module is mainly useful in the context of Web Applications.

Conventions

- The basex-api package must be included in the classpath. This is always the case if you use one of the complete distributions (zip, exe, war) of BaseX.

- All functions are assigned to the http://basex.org/modules/sessions namespace. The module must be imported in the query prolog:

```xquery
import module namespace sessions = "http://basex.org/modules/sessions";
```

- In this documentation, the namespace is bound to the sessions prefix.

- Errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

- If any of the functions is called outside the servlet context, the error BXSE0003: is raised.

- As sessions are side-effecting operations, all functions are flagged as non-deterministic. This means that the functions will not be reordered by the compiler.

Functions

sessions:ids

<table>
<thead>
<tr>
<th>Signatures</th>
<th>sessions:ids() as xs:string</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the IDs of all registered sessions.</td>
</tr>
</tbody>
</table>

sessions:created

<table>
<thead>
<tr>
<th>Signatures</th>
<th>sessions:created($id as xs:string) as xs:dateTime</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the creation time of the session specified by $id.</td>
</tr>
</tbody>
</table>

sessions:accessed

<table>
<thead>
<tr>
<th>Signatures</th>
<th>sessions:accessed($id as xs:string) as xs:dateTime</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the last access time of the session specified by $id.</td>
</tr>
</tbody>
</table>

sessions:names

<table>
<thead>
<tr>
<th>Signatures</th>
<th>sessions:names($id as xs:string) as xs:string*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the names of all variables bound to the session specified by $id.</td>
</tr>
</tbody>
</table>

sessions:get

<table>
<thead>
<tr>
<th>Signatures</th>
<th>sessions:get($id as xs:string, $key as xs:string) as item()*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>sessions:get($id as xs:string, $key as xs:string, $default as item()<em>) as item()</em></td>
</tr>
</tbody>
</table>
Sessions Module

Summary

Returns the value of a variable bound to the session specified by $id. If the variable does not exist, or if the key is unknown, an empty sequence or the optionally specified default value is returned instead.

Errors

BXSE0002: the value of a session variable could not be retrieved.

sessions: set

Signatures

sessions: set($id as xs:string, $key as xs:string, $value as item()* as empty-sequence() )

Summary

Binds the specified key/value pair to the session specified by $id.

Errors

BXSE0001: a function item was specified as value of a session variable.

sessions: delete

Signatures

sessions: delete($id as xs:string, $key as xs:string) as empty-sequence()

Summary

Deletes a variable bound to the session specified by $id.

sessions: close

Signatures

sessions: close($id as xs:string) as empty-sequence()

Summary

Unregisters the session specified by $id.

Errors

Code Description
BXSE0001 A function item was specified as value of a session attribute.
BXSE0002 An error occurred while retrieving the value of a session attribute.
BXSE0003 A function was called outside the servlet context.
BXSE0004 The specified session was not found.

Changelog

Version 8.4

- Updated: Allow sequences as session values.

This module was introduced with Version 7.5.
Chapter 64. SQL Module

This XQuery Module contains functions to access relational databases from XQuery using SQL. With this module, you can execute query, update and prepared statements, and the result sets are returned as sequences of XML elements representing tuples. Each element has children representing the columns returned by the SQL statement.

This module uses JDBC to connect to a SQL server. Hence, your JDBC driver will need to be added to the classpath, too. If you work with the full distributions of BaseX, you can copy the driver into the \lib\ directory. To connect to MySQL, for example, download the Connector/J Driver and extract the archive into this directory.

Conventions

All functions in this module are assigned to the http://basex.org/modules/sql namespace, which is statically bound to the sql prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

Functions

sql:init

Signatures

sql:init($class as xs:string) as empty-sequence()

Summary

This function initializes a JDBC driver specified via $class. This step might be superfluous if the SQL database is not embedded.

Errors

BXSQ0007: the specified driver class is not found.

sql:connect

Signatures

sql:connect($url as xs:string) as xs:integer

Summary

This function establishes a connection to a relational database. As a result a connection handle is returned. The parameter $url is the URL of the database and shall be of the form: jdbc:<driver name>://[<server>[/<database>]]. If the parameters $user and $password are specified, they are used as credentials for connecting to the database. The $options parameter can be used to set connection options.

Errors

BXSQ0001: an SQL exception occurs, e.g. missing JDBC driver or not existing relation.

Examples

Connects to an SQL Server and sets autocommit to true:

sql:connect('jdbc:sqlserver://DBServer', map { 'autocommit': true() })

sql:execute

Once a connection is established, the returned connection handle can be used to execute queries on the database. Our SQL module supports both direct queries and prepared statements.

Signatures

sql:execute($connection as xs:integer, $query as xs:string) as element()*

Summary

This function executes a query or update statement. The parameter $connection specifies a connection handle. The parameter $query is a string representing an SQL statement.
Errors

<table>
<thead>
<tr>
<th>SQL Module</th>
</tr>
</thead>
</table>

BXSQ0001: an SQL exception occurs, e.g. not existing relation is retrieved. BXSQ0002: a wrong connection handle is passed.

**sql:execute-prepared**

**Signatures**

```xml
sql:execute-prepared($id as xs:integer, $params as element(sql:parameters)) as element(*)*
```

**Summary**

This function executes a prepared statement. The parameter $id specifies a prepared statement handle. The optional parameter $params is an element `<sql:parameters/>` representing the parameters for a prepared statement along with their types and values. The following schema shall be used:

```xml
element sql:parameters {
  element sql:parameter {
    attribute type { "int"|"string"|"boolean"|"date"|"double"|"float"|"short"|"time"|"timestamp" },
    attribute null { "true"|"false" },
    text }+ }
```

**Errors**

BXSQ0001: an SQL exception occurs, e.g. not existing relation is retrieved. BXSQ0002: a wrong prepared statement handle is passed. BXSQ0003: the number of `<sql:parameter/>` elements in `<sql:parameters/>` differs from the number of placeholders in the prepared statement. BXSQ0004: the type of a parameter for a prepared statement is not specified. BXSQ0005: an attribute different from type and null is set for a `<sql:parameter/>` element. BXSQ0006: a parameter is from type date, time or timestamp and its value is in an invalid format.

**sql:prepare**

**Signatures**

```xml
sql:prepare($connection as xs:integer, $statement as xs:string) as xs:integer
```

**Summary**

This function prepares a statement and returns a handle to it. The parameter $connection indicates the connection handle to be used. The parameter $statement is a string representing an SQL statement with one or more '?' placeholders. If the value of a field has to be set to NULL, then the attribute null of the element `<sql:parameter/>` has to be true.

**Errors**

BXSQ0001: an SQL exception occurs. BXSQ0002: a wrong connection handle is passed.

**sql:commit**

**Signatures**

```xml
sql:commit($connection as xs:integer) as empty-sequence()
```

**Summary**

This function commits the changes made to a relational database. $connection specifies the connection handle.

**Errors**

BXSQ0001: an SQL exception occurs. BXSQ0002: a wrong connection handle is passed.

**sql:rollback**

**Signatures**

```xml
sql:rollback($connection as xs:integer) as empty-sequence()
```

**Summary**

This function rolls back the changes made to a relational database. $connection specifies the connection handle.

**Errors**

BXSQ0001: an SQL exception occurs. BXSQ0002: a wrong connection handle is passed.

**sql:close**

**Signatures**

```xml
sql:close($connection as xs:integer) as empty-sequence()
```
Summary
This function closes a connection to a relational database. $connection specifies the connection handle.

Errors
BXSQ0001: an SQL exception occurs. BXSQ0002: a wrong connection handle is passed.

Examples

Direct queries
A simple select statement can be executed on the following way:

```$conn := sql:connect("jdbc:postgresql://localhost:5432/coffeehouse")
return sql:execute($conn, "SELECT * FROM coffees WHERE price < 10")
```

The result will look like:

```
<sql:row xmlns:sql="http://basex.org/modules/sql">
  <sql:column name="cof_name">French_Roast</sql:column>
  <sql:column name="sup_id">49</sql:column>
  <sql:column name="price">9.5</sql:column>
  <sql:column name="sales">15</sql:column>
  <sql:column name="total">30</sql:column>
</sql:row>
```

Prepared Statements
A prepared select statement can be executed in the following way:

```
(let $conn := sql:connect("jdbc:postgresql://localhost:5432/coffeehouse")
let $prep := sql:prepare($conn, "SELECT * FROM coffees WHERE price < ? AND cof_name = ?")
(let $params := <sql:parameters>
  <sql:parameter type='double'>10</sql:parameter>
  <sql:parameter type='string'>French_Roast</sql:parameter>
</sql:parameters>)
return sql:execute-prepared($prep, $params)
```

SQLite
The following expression demonstrates how SQLite can be addressed using the Xerial SQLite JDBC driver:
(! Initialize driver !)
sql:init("org.sqlite.JDBC"),
(! Establish a connection !)
let $conn := sql:connect("jdbc:sqlite:database.db")
return (
  (! Create a new table !)
  sql:execute($conn, "drop table if exists person"),
  sql:execute($conn, "create table person (id integer, name string)"),
  (! Run 10 updates !)
  for $i in 1 to 10
  let $q := "insert into person values(" || $i || ", " || $i || ")"
  return sql:execute($conn, $q),
  (! Return table contents !)
  sql:execute($conn, "select * from person"),
  sql:close($conn)
)

Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXSQ0001</td>
<td>An SQL exception occurred (e.g.: a non-existing relation is retrieved).</td>
</tr>
<tr>
<td>BXSQ0002</td>
<td>A wrong connection handle or prepared statement handle is passed.</td>
</tr>
<tr>
<td>BXSQ0003</td>
<td>The number of <a href="">sql:parameter/</a> elements in <a href="">sql:parameters/</a> differs from the</td>
</tr>
<tr>
<td></td>
<td>number of placeholders in the prepared statement.</td>
</tr>
<tr>
<td>BXSQ0004</td>
<td>The type of a parameter for a prepared statement is not specified.</td>
</tr>
<tr>
<td>BXSQ0005</td>
<td>An attribute different from type and null is set for a <a href="">sql:parameter/</a> element.</td>
</tr>
<tr>
<td>BXSQ0006</td>
<td>A parameter is from type date, time or timestamp and its value is in an invalid format.</td>
</tr>
<tr>
<td>BXSQ0007</td>
<td>A specified database driver class is not found.</td>
</tr>
</tbody>
</table>

Changelog

Version 7.5

- Updated: prepared statements are now executed via sql:execute-prepared

The module was introduced with Version 7.0.
Chapter 65. Streaming Module

This XQuery Module contains functions for handling streamable items.

In contrast to standard XQuery items, a streamable item contains only a reference to the actual data. The data itself will be retrieved if it is requested by an expression, or if the item is to be serialized. Hence, a streamable item only uses a few bytes, and no additional memory is occupied during serialization.

The following BaseX functions return streamable items:

- Streamable Base64 binaries:
  - `db:retrieve`
  - `fetch:binary`
  - `file:read-binary`

- Streamable strings:
  - `fetch:text`
  - `file:read-text`

Some functions are capable of consuming items in a streamable fashion: data will never be cached, but instead passed on to another target (file, the calling expression, etc.). The following streaming functions are currently available:

- `convert:binary-to-bytes`
- `db:store`
- `file:write-binary`
- `file:write-text`

The XQuery expression below serves as an example on how large files can be downloaded and written to a file with constant memory consumption:

```xquery
```

Conventions

All functions in this module are assigned to the `http://basex.org/modules/stream` namespace, which is statically bound to the `stream` prefix. All errors are assigned to the `http://basex.org/errors` namespace, which is statically bound to the `bxerr` prefix.

Functions

**stream:materialize**

| Signatures | stream:materialize($value as item()* as item()*) |
| Summary    | Returns a materialized instance of the specified $value: |
|            |   • if an item is streamable, its value will be retrieved, and a new item containing the value will be returned. |
Streaming Module

- other, non-streamable items will simply be passed through.

Materialization is advisable if a value is to be processed more than once, and is expensive to retrieve. It is get mandatory whenever a value is invalidated before it is requested (see the example below).

**Example**

In the following example, a file will be deleted before its content is returned. To avoid a "file not found" error, the content will first be materialized:

```plaintext
let $file := 'data.txt'
let $data := stream:materialize(file:read-text($file))
return (file:delete($file), $data)
```

**stream:is-streamable**

**Signatures**

`stream:is-streamable($item as item()) as item()`

**Summary**

Checks whether the specified $item is streamable.

**Changelog**

Version 8.0

- Update: `stream:materialize` extended to sequences.

This module was introduced with Version 7.7.
Chapter 66. Strings Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions for string computations.

Conventions

All functions in this module and errors are assigned to the http://basex.org/modules/strings namespace, which is statically bound to the strings prefix.

Functions

strings:levenshtein

Signatures

| strings:levenshtein($string1 as xs:string, $string2 as xs:string) as xs:double |

Summary

Computes the Damerau-Levenshtein Distance for two strings and returns a double value (0.0 - 1.0). The returned value is computed as follows:

- 1.0 – distance / max(length of strings)
- 1.0 is returned if the strings are equal; 0.0 is returned if the strings are too different.

Examples

- strings:levenshtein("flower", "flower") returns 1
- strings:levenshtein("flower", "lewes") returns 0.5
- In the following query, the input is first normalized (words are stemmed, converted to lower case, and diacritics are removed). It returns 1:

  let $norm := ft:normalize(?, map { 'stemming': true() })
  return strings:levenshtein($norm("HOUSES"), $norm("house"))

strings:soundex

Signatures

| strings:soundex($string as xs:string) as xs:string |

Summary

Computes the Soundex value for the specified string. The algorithm can be used to find and index English words with similar pronunciation.

Examples

- strings:soundex("Michael") returns M240
- strings:soundex("OBrien") = strings:soundex("O'Brien") returns true

strings:cologne-phonetic

Signatures

| strings:cologne-phonetic($string as xs:string) as xs:string |

Summary

Computes the Kölner Phonetik value for the specified string. Similar to Soundex, the algorithm is used to find similarly pronounced words, but for the German language. As the first returned digit can be 0, the value is returned as string.

Examples

- strings:cologne-phonetic("Michael") returns 645
- every $s in ("Mayr", "Maier", "Meier") satisfies strings:cologne-phonetic($s) = "67" returns true
Changelog

The Module was introduced with Version 8.3.
Chapter 67. Unit Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains annotations and functions for performing XQUnit tests.

Introduction

The more complex a software application grows, the more error-prone it gets. This is why testing frameworks have been developed, which provide a standardized, automated way of testing software. The XUnit frameworks (such as SUnit or JUnit) allow testing of atomic units of a program, such as single functions and algorithms.

This module borrows heavily from the existing frameworks: it provides various annotations for testing XQuery functions. Unit functions are provided to assert the validity of arbitrary conditions expressed in XQuery and to raise errors whenever a condition is not satisfied. Some additional functions exist to run all unit tests of the current module or a set of specified library modules.

Usage

Tests are started via the TEST command. It compiles all XQuery modules in a given file or directory and runs all functions that are annotated with %unit:test. A test report is generated and returned, which resembles the format returned by other xUnit testing frameworks, such as the Maven Surefire Plugin (see below).

Conventions

All annotations, functions and errors in this module are assigned to the http://basex.org/modules/unit namespace, which is statically bound to the unit prefix.

Annotations

%unit:test

Syntax

```
%unit:test %unit:test("expected", CODE)
```

Summary

With this annotation, a function can be marked as unit test. It will be evaluated if a test report is created for the module in which this function is located. error can be supplied as additional string argument. It is followed by CODE, which must be a valid EQName string. If the function expression does not raise that error, the test will fail.

Examples

- The following test does will be successful, as it does nothing (and, hence, nothing wrong):

  declare %unit:test function local:void() { () };

- The following test will be successful, as the function body will raise err:XPTY0004:

  declare %unit:test('expected', "err:XPTY0004") function local:add() {
  123 + 'strings and integers cannot be added'
};

%unit:before

Syntax

```
%unit:before %unit:before(FUNCTION)
```

Summary

A function decorated with this annotation will be evaluated before each unit test. FUNCTION can be supplied as additional argument. It must be a valid EQName string. If specified, the function will only be evaluated before a function with the given name is tested. This extension is e.g. helpful if the results of updates need to be tested.
Examples

• The first function will be evaluated before the actual test:

```%unit:before
function

declare %updating %unit:before("local:check") function local:before-check () {
  db:create('test-db')
};
declare %updating %unit:test function local:check () {
  unit:assert(db:exists('test-db'))
};
```

%unit:after

Syntax `%unit:after %unit:after(FUNCTION)`

Summary A function decorated with this annotation will be evaluated after each unit test. `FUNCTION` can be supplied as additional argument. It must be a valid EQName string. If specified, the function will only be evaluated after a function with the given name is tested.

%unit:before-module

Syntax `%unit:before-module`  

Summary If a function is decorated with this annotation, it will be evaluated before all unit tests in the current module.

%unit:after-module

Syntax `%unit:after-module`  

Summary If a function is decorated with this annotation, it will be evaluated after all unit tests in the current module.

%unit:ignore

Syntax `%unit:ignore %unit:ignore(MESSAGE)`  

Summary If a function is decorated with this annotation, it will temporarily be ignored by the test suite runner.

Functions

unit:assert

Signatures `unit:assert($test as item()*) as empty-sequence()` `unit:assert($test as item()*, $info as item()) as empty-sequence()`  

Summary Asserts that the effective boolean value of the specified `$test` is true and returns an empty sequence. Otherwise, raises an error. The effective boolean value of an expression can be explicitly computed by using the `fn:boolean` function. The default failure message can be overridden with the `$info` argument.

Errors UNIT0001: the assertion failed, or an error was raised.

unit:assert-equals

Signatures `unit:assert-equals($returned as item()*, $expected as item()*) as empty-sequence()` `unit:assert-equals($returned as item()*, $expected as item()*, $info as item()) as empty-sequence()`  

Summary Asserts that the specified arguments are equal according to the rules of the `fn:deep-equal` function. Otherwise, raises an error. The default failure message can be overridden with the `$info` argument.
Errors | UNIT0001: the assertion failed, or an error was raised.

unit:fail

Signatures | unit:fail() as empty-sequence() unit:fail($info as item()) as empty-sequence()

Summary | Raises a unit error. The default failure message can be overridden with the $info argument.

Example

The following XQUnit module tests.xqm contains all available unit annotations:

Query

```xml
module namespace test = 'http://basex.org/modules/xqunit-tests';

(:~ Initializing function, which is called once before all tests. :) declare %unit:before-module function test:before-all-tests() {
  ()
};

(:~ Initializing function, which is called once after all tests. :) declare %unit:after-module function test:after-all-tests() {
  ()
};

(:~ Initializing function, which is called before each test. :) declare %unit:before function test:before() {
  ()
};

(:~ Initializing function, which is called after each test. :) declare %unit:after function test:after() {
  ()
};

(:~ Function demonstrating a successful test. :) declare %unit:test function test:assert-success() {
  unit:assert(<a/>)
};

(:~ Function demonstrating a failure using unit:assert. :) declare %unit:test function test:assert-failure() {
  unit:assert((), 'Empty sequence.')
};

(:~ Function demonstrating a failure using unit:assert-equals. :) declare %unit:test function test:assert-equals-failure() {
  unit:assert-equals(4 + 5, 6)
};

(:~ Function demonstrating an unexpected success. :) declare %unit:test("expected", "err:FORG0001") function test:unexpected-success() {
  ()
};

(:~ Function demonstrating an expected failure. :) declare %unit:test("expected", "err:FORG0001") function test:expected-failure() {
  1 + <a/>
};
```
Unit Module

(:~ Function demonstrating the creation of a failure. :) 
declare %unit:test function test:failure() {
    unit:fail("Failure!")
};

(:~ Function demonstrating an error. :) 
declare %unit:test function test:error() {
    1 + a/
};

(:~ Skipping a test. :) 
declare %unit:test %unit:ignore("Skipped!") function test:skipped() {
    ()
};

By running TEST tests.xqm, the following report will be generated (timings may differ):

Result

<testsuites time="PT0.256S">
    <testsuite name="file:///C:/Users/user/Desktop/test.xqm" time="PT0.212S" tests="8" failures="4" errors="1" skipped="1">
        <testcase name="assert-success" time="PT0.016S"/>
        <testcase name="assert-failure" time="PT0.005S">
            <failure line="30" column="15">
                <info>Empty sequence.</info>
            </failure>
        </testcase>
        <testcase name="assert-equals-failure" time="PT0.006S">
            <failure line="35" column="22">
                <returned item="1" type="xs:integer">9</returned>
                <expected item="1" type="xs:integer">6</expected>
                <info>Item 1: 6 expected, 9 returned.</info>
            </failure>
        </testcase>
        <testcase name="unexpected-success" time="PT0.006S">
            <failure>
                <expected>FORG0001</expected>
            </failure>
        </testcase>
        <testcase name="expected-failure" time="PT0.004S"/>
        <testcase name="failure" time="PT0.004S">
            <failure line="50" column="13">
                <info>Failure!</info>
            </failure>
        </testcase>
        <testcase name="error" time="PT0.004S">
            <error line="55" column="6" type="FORG0001">
                <info>Cannot cast to xs:double: "".</info>
            </error>
        </testcase>
        <testcase name="skipped" skipped="Skipped!" time="PT0S"/>
    </testsuite>
</testsuites>

Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNIT0001</td>
<td>An assertion failed, or an error was raised.</td>
</tr>
<tr>
<td>UNIT0002</td>
<td>A test function must have no arguments.</td>
</tr>
</tbody>
</table>
UNIT0003 A test function is not public.
UNIT0004 An annotation was declared twice.
UNIT0005 An annotation has invalid arguments.

**Changelog**

Version 8.0.2

- Updated: (expected) errors are compared by QNames instead of local names (including namespaces).

Version 8.0

- Deleted: UNIT0006 (ignore results returned by functions).
- Added: unit:fail, 0-argument signature.
- Updated: the info argument of functions can now be an arbitrary item.
- Updated: infos are now represented in an info child element.
- Updated: unit:before and unit:after can be extended by a filter argument.

Version 7.9

- Added: TEST command
- Removed: unit:test, unit:test-uris

Version 7.8

- Added: unit:assert-equals
- Updated: enhanced test report output

This module was introduced with Version 7.7.
Chapter 68. User Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions for creating and administering database users. The User Management article gives more information on database users and permissions.

Conventions

All functions in this module and errors are assigned to the http://basex.org/modules/user namespace, which is statically bound to the user prefix.

Read Operations

user:current

Signatures
user:current() as xs:string

Summary
Returns the name of the currently logged in user.

Examples
* If the GUI or the standalone mode is used, user:current() always returns admin.

user:list

Signatures
user:list() as xs:string*

Summary
Returns the names of all registered users.

Examples
• After installing BaseX, user:list() returns admin.

user:list-details

Signatures
user:list-details() as element(user)*
user:list-details($name as xs:string) as element(user)*

Summary
Returns an element sequence, containing all registered users and their permissions. In addition to the SHOW USERS command, encoded password strings and database permissions will be output. A user $name can be specified to filter the results in advance.

Examples
• By default, user:list-details() returns the following output:

```
<user name="admin" permission="admin">
  <password algorithm="digest">
    <hash>304bdfb0383c16f070a897fc1eb25cb4</hash>
  </password>
  <password algorithm="salted-sha256">
    <salt>871602799292195</salt>
    <hash>a065ca66fa3d6da5762c227587f1c8258c6dc08ee867e44a605a72da115dcf41</hash>
  </password>
</user>
```

Errors
unknown: The specified user name is unknown.

user:exists

Signatures
user:exists($name as xs:string) as xs:boolean

Summary
Checks if a user with the specified $name exists.
### User Module

#### Examples
- `user:exists('admin')` will always yield true.

#### Errors
- `name`: The specified user name is invalid.

### Updates

**Important note:** All functions in this section are updating functions: they will not be immediately executed, but queued on the Pending Update List, which will be processed after the actual query has been evaluated. This means that the order in which the functions are specified in the query does usually not reflect the order in which the code will be evaluated.

#### user:create

**Signatures**
```
user:create($name as xs:string, $password as xs:string) as empty-sequence()
user:create($name as xs:string, $password as xs:string, $permissions as xs:string*) as empty-sequence()
user:create($name as xs:string, $password as xs:string, $permissions as xs:string*, $patterns as xs:string*) as empty-sequence()
```

**Summary**
Creates a new user with the specified $name, $password, and $permissions. Local permissions are granted with non-empty glob $patterns. The default global permission (none) can be overwritten with an empty pattern or by omitting the last argument. Existing users will be overwritten.

**Examples**
- `user:create('John', '7e$j#!1', 'admin')` creates a new user 'John' with admin permissions.
- `user:create('Jack', 'top!secret', 'read', 'index*')` creates a new user 'Jack' with no permissions, but write permissions for databases starting with the letters 'index'.

**Errors**
- `name`: The specified user name is invalid.
- `permission`: The specified permission is invalid.
- `admin`: The "admin" user cannot be modified.
- `logged-in`: The specified user is currently logged in.
- `update`: The operation can only be performed once per user or database pattern.

#### user:grant

**Signatures**
```
user:grant($name as xs:string, $permissions as xs:string*) as empty-sequence()
user:grant($name as xs:string, $permissions as xs:string*, $patterns as xs:string*) as empty-sequence()
```

**Summary**
Grants global or local $permissions to a user with the specified $name. Local permissions are granted with non-empty glob $patterns.

**Examples**
- `user:grant('John', 'create')` grants create permissions to the user 'John'.
- `user:grant('John', ('read','write'), ('index*','unit*'))` allows John to read all databases starting with the letters 'index', and to write to all databases starting with 'unit'.

**Errors**
- `unknown`: The specified user name is unknown.
- `name`: The specified user name is invalid.
- `pattern`: The specified database pattern is invalid.
- `permission`: The specified permission is invalid.
- `admin`: The "admin" user cannot be modified.
- `local`: A local permission can only be 'none', 'read' or 'write'.
- `logged-in`: The specified user is currently logged in.
- `update`: The operation can only be performed once per user or database pattern.

#### user:drop

**Signatures**
```
user:drop($name as xs:string) as empty-sequence()
user:drop($name as xs:string, $patterns as xs:string*) as empty-sequence()
```

**Summary**
Drops a user with the specified $name. If non-empty glob $patterns are specified, only the database patterns will be dropped.
Examples

- `user:drop('John')` drops the user 'John'.
- `user:grant('John', 'unit*')` removes the 'unit*' database pattern. If John accesses any of these database, his global permission will be checked again.

Errors

unknown: The specified user name is unknown.
name: The specified user name is invalid.
pattern: The specified database pattern is invalid.
admin: The "admin" user cannot be modified.
logged-in: The specified user is currently logged in.
update: The operation can only be performed once per user or database pattern.
conflict: A user cannot be both altered and dropped.

user:alter

Signatures

user:alter($name as xs:string, $newname as xs:string) as empty-sequence()

Summary

Renames a user with the specified $name to $newname.

Examples

- `user:rename('John', 'Jack')` renames the user 'John' to 'Jack'.

Errors

unknown: The specified user name is unknown.
name: The specified user name is invalid.
admin: The "admin" user cannot be modified.
logged-in: The specified user is currently logged in.
update: The operation can only be performed once per user or database pattern.
conflict: A user cannot be both altered and dropped.

user:password

Signatures

user:password($name as xs:string, $password as xs:string) as empty-sequence()

Summary

Changes the password of a user with the specified $name.

Examples

- `user:password('John', )` assigns user 'John' an empty password string.

Errors

unknown: The specified user name is unknown.
name: The specified user name is invalid.
update: The operation can only be performed once per user or database pattern.

Errors

Code | Description
--- | ---
name | The specified user name is invalid.
pattern | The specified database name is invalid.
permission | The specified permission is invalid.
unknown | The specified user does not exist.
admin | The "admin" user cannot be modified.
equal | Name of old and new user is equal.
local | A local permission can only be 'none', 'read' or 'write'.
logged-in | The specified user is currently logged in.
update | The operation can only be performed once per user or database pattern.
conflict | A user cannot be both altered and dropped.

Changelog

Version 8.4

- Updated: `user:create`, `user:grant`, `user:drop`: extended support for database patterns.
Version 8.1

- Added: user:current.

The Module was introduced with Version 8.0.
Chapter 69. Validation Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions to perform validations against DTDs, XML Schema and RelaxNG. For Schematron validation, check out Vincent Lizzi’s Schematron XQuery Module.

Conventions

All functions in this module are assigned to the http://basex.org/modules/validate namespace, which is statically bound to the validate prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

DTD Validation

Checks whether an XML document validates against a DTD. The input document can be specified as:

- an xs:string, containing the path to the resource,
- an xs:string, containing the resource in its string representation, or
- a node(), containing the resource itself.

If no DTD is supplied in a function, the XML document is expected to contain an embedded DTD doctype declaration.

**validate:dtd**

**Signatures**

validate:dtd($input as item()) as empty-sequence()

validate:dtd($input as item(), $schema as xs:string?) as empty-sequence()

**Summary**

Validates the XML $input document against a $schema and returns an empty sequence or an error.

**Errors**

BXVA0001: the validation fails. BXVA0002: the validation process cannot be started. BXVA0003: no DTD validator is available.

**Examples**

- validate:dtd('doc.xml', 'doc.dtd') validates the document doc.xml against the specified DTD file doc.dtd.
- The following example validates an invalid document against a DTD, which is specified as string:

```xml
try {
    let $doc := <invalid/>
    let $schema := '<!ELEMENT root (#PCDATA)>'
    return validate:dtd($doc, $schema)
} catch bxerr:BXVA0001 {
    'DTD Validation failed.'
}
```

**validate:dtd-info**

**Signatures**

validate:dtd-info($input as item()) as xs:string*

validate:dtd-info($input as item(), $schema as xs:string?) as xs:string*

**Summary**

Validates the XML $input document against a $schema and returns warnings, errors and fatal errors in a string sequence.

**Errors**

BXVA0002: the validation process cannot be started. BXVA0003: no DTD validator is available.
Examples

validate:dtd-info(<invalid/>, '<!ELEMENT root (#PCDATA)>') returns: 2:11: Element type "invalid" must be declared.

validate:dtd-report

| width='100%' | width='120' |

Signatures

validate:dtd-report($input as item()) as element(report) validate:dtd-report($input as item(), $schema as xs:string?) as element(report) -

Summary

Validates the XML $input document against a $schema and returns warnings, errors and fatal errors as XML. -

Errors

BXVA0002: the validation process cannot be started.

Examples

• validate:dtd-report(<invalid/>, '<!ELEMENT root (#PCDATA)>') returns:

<report>
    <status>invalid</status>
    <message level="Error" line="2" column="11">Element type "invalid" must be declared.</message>
</report>

XML Schema Validation

Checks whether an XML document validates against an XML Schema. The input document and the schema can be specified as:

• `xs:string`, containing the path to the resource,
• `xs:string`, containing the resource in its string representation, or
• `node()`, containing the resource itself.

If no schema is given, the input is expected to contain an `xsi:(noNamespace)schemaLocation` attribute, as defined in W3C XML Schema.

Different XML Schema implementations can be applied for validation:

• By default, the Java implementation of XML Schema 1.0 is used (it is based on an old version of Apache Xerces).
• If Xerces2 exists in the classpath, it will be used instead. If you want to manually add Xerces2, you should download the Xerces2 Java ..... (XML Schema 1.1) (Beta) archive and copy at least four libraries (cupv10k-
runtime.jar, org.eclipse.wst.xml.xpath2.processor_1.1.0.jar, xercesImpl.jar, xml-apis.jar) to the lib directory of the full distribution of BaseX.
• The Saxon Enterprise Edition (saxon9ee.jar) is detected as well.

The XML Schema version can be specified as well (the supplied string must be `1.0` or `1.1`). Please note that Version 1.1 is only available if Saxon EE or Xerces2 is added to the classpath.

validate:xsd

| width='100%' | width='120' |

Signatures

validate:xsd($input as item()) as empty-sequence() validate:xsd($input as item(), $schema as item()?) as empty-sequence() validate:xsd($input as item(), $schema as item()?, $version as xs:string) as empty-sequence()

Summary

Validates the XML $input document against a $schema.

Errors

BXVA0001: the validation fails.BXVA0002: the validation process cannot be started.BXVA0003: no XML Schema validator is available.BXVA0004: no validator is found for the specified version.
Examples

• validate:xsd('doc.xml', 'doc.xsd') validates the document doc.xml against the specified schema doc.xsd.

• The following example demonstrates how a document can be validated against a schema without resorting to local or remote URIs:

```xml
let $schema :=
  <xs:schema xmlns:xs='http://www.w3.org/2001/XMLSchema'
    targetNamespace='http://baseX.org/simple'>
    <xs:element name='root'/>
  </xs:schema>
return validate:xsd($doc, $schema)
```

validate:xsd-info

Signatures

validate:xsd-info($input as item()) as xs:string*
validate:xsd-info($input as item(), $schema as item()) as xs:string*

Summary

Validates the XML $input document against a $schema and returns warnings, errors and fatal errors in a string sequence.

Errors

BXVA0002: the validation process cannot be started.
BXVA0003: no XML Schema validator is available.
BXVA0004: no validator is found for the specified version.

validate:xsd-report

Signatures

validate:xsd-report($input as item()) as element(report)
validate:xsd-report($input as item(), $schema as xs:string?) as element(report)

Summary

Validates the XML $input document against a $schema and returns warnings, errors and fatal errors as XML.

Errors

BXVA0002: the validation process cannot be started.
BXVA0003: no XML Schema validator is available.
BXVA0004: no validator is found for the specified version.

RelaxNG Validation

Checks whether an XML document validates against a RelaxNG schema. The input document and the schema can be specified as:

• xs:string, containing the path to the resource,

• xs:string, containing the resource in its string representation, or

• node(), containing the resource itself.

RelaxNG validation will be available if Jing exists in the classpath. The latest version, jing-20091111.jar, is included in the full distributions of BaseX.

validate:rng

Signatures

validate:rng($input as item(), $schema as item()) as empty-sequence()
validate:rng($input as item(), $schema as item(), $compact as xs:boolean) as empty-sequence()

Summary

Validates the XML $input document against a $schema, using the XML or $compact notation.

Errors

BXVA0001: the validation fails.
BXVA0002: the validation process cannot be started.
BXVA0003: the RelaxNG validator is not available.
Examples

- `validate:rng('doc.xml', 'doc.rng')` validates the document `doc.xml` against the specified schema `doc.rng`.

### validate:rng-info

**Signatures**

```xml
validate:rng-info($input as item(), $schema as item()) as xs:string*
validate:rng-info($input as item(), $schema as item(), $compact as xs:boolean) as xs:string*
```

**Summary**

Validates the XML `$input` document against a `$schema`, using the XML or `$compact` notation, and returns warnings, errors and fatal errors in a string sequence.

**Errors**

- BXVA0002: the validation process cannot be started.
- BXVA0003: the RelaxNG validator is not available.

### validate:rng-report

**Signatures**

```xml
validate:rng-report($input as item(), $schema as xs:string) as element/report
validate:rng-report($input as item(), $schema as xs:string, $compact as xs:boolean) as element/report
```

**Summary**

Validates the XML `$input` document against a `$schema`, using the XML or `$compact` notation, and returns warnings, errors and fatal errors as XML.

**Errors**

- BXVA0002: the validation process cannot be started.
- BXVA0003: The RelaxNG validator is not available.

### Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXVA0001</td>
<td>The document cannot be validated against the specified DTD or XML Schema.</td>
</tr>
<tr>
<td>BXVA0002</td>
<td>The validation cannot be started.</td>
</tr>
<tr>
<td>BXVA0003</td>
<td>No validator is available.</td>
</tr>
<tr>
<td>BXVA0004</td>
<td>No validator is found for the specified version.</td>
</tr>
</tbody>
</table>

### Changelog

**Version 8.3**

- Added: `validate:rng, validate:rng-info`

**Version 7.6**


The module was introduced with Version 7.3.
Chapter 70. Web Module

Read this entry online in the BaseX Wiki.

This XQuery Module provides convenience functions for building web applications with RESTXQ.

Conventions

All functions in this module are assigned to the http://basex.org/modules/web namespace, which is statically bound to the web prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

Functions

web:content-type

Signatures
web:content-type($path as xs:string) as xs:string

Summary
Returns the content type of a path by analyzing its file suffix. application/octet-stream is returned if the file suffix is unknown.

Examples
• web:content-type("sample.mp3") returns audio/mpeg

web:create-url

Signatures
web:create-url($url as xs:string, $parameters as map(item(), item(*))) as xs:string

Summary
Creates a new URL from the specified $url string and the $parameters specified in a map. The keys and values of the map entries will be converted to strings, URL-encoded (see web:encode-url), and appended to the url as query parameters. If a map entry has more than a single item, all of them will be appended as single parameters.

Examples
• web:create-url('http://find.me', map { 'q': 'dog' }) returns http://find.me?q=dog

web:encode-url

Signatures
web:encode-url($string as xs:string) as xs:string

Summary
Encodes a string to a URL. Spaces are rewritten to +, *, - , and _ are adopted; and all other non-ASCII characters and special characters are percent-encoded.

Examples
• web:encode-url("this is a test!.html") returns this+is+a+test%21.html.

web:decode-url

Signatures
web:decode-url($string as xs:string) as xs:string

Summary
Decodes a URL to the original string. Percent-encoded characters are decoded to their UTF8 codepoints, and + characters are rewritten to spaces.

Examples
• web:decode-url("%E6%97%A5%E6%9C%AC%E8%AA%9E") returns ###.

Errors
BXWE0001: Specified URI is invalid.
**web:redirect**

**Signatures**
- `web:redirect($location as xs:string) as element(rest:response)`
- `web:redirect($location as xs:string, $parameters as map(item(), item()*)) as element(rest:response)`

**Summary**
Creates a RESTXQ redirection to the specified location. The returned response will only work if no other items are returned by the RESTXQ function. If `$parameters` are specified, they will be appended as query parameters to the URL as described for `web:create-url`.

**Examples**
The query `web:redirect('/a/b')` returns the following response element:

```xml
<rest:response xmlns:rest="http://exquery.org/ns/restxq">
    <http:header name="location" value="/a/b"/>
  </http:response>
</rest:response>
```

**web:response-header**

**Signatures**
- `web:response-header() as element(rest:response)`
- `web:response-header($output as map(*)) as element(rest:response)`
- `web:response-header($output as map(*), $headers as map(xs:string, xs:string)) as element(rest:response)`

**Summary**
Creates a RESTXQ response header with a default caching directive and default serialization parameters. Serialization parameters can be supplied via the `$output` argument. Empty string values can be specified to invalidate default values. By default, the following serialization parameters will be returned:

- `media-type`: `application/octet-stream`

Header options can be supplied via the `$headers` argument. Empty string values can be specified to invalidate default values. By default, the following header options will be returned:

- `Cache-Control`: `max-age=3600,public`

**Examples**
- The function call `web:response-header()` returns the following response element:

```xml
<rest:response xmlns:rest="http://exquery.org/ns/restxq">
    <http:header name="Cache-Control" value="max-age=3600,public"/>
  </http:response>
  <output:serialization-parameters xmlns:output="http://www.w3.org/2010/xslt-xquery-serialization">
    <output:media-type value="application/octet-stream"/>
  </output:serialization-parameters>
</rest:response>
```
- If the following RESTXQ function is called by a web browser...

```xml
declare %rest:path('media/{$file}')
function local:get($file)
{
  let $path := 'path/to/' || $file
  return {
    web:response-header(map { 'media-type': web:content-type($path) }),
    file:read-binary($path)
  }
```

---

**Errors**
BXWE0002: Specified URI contains invalid XML characters.
...a file resource with the correct content-type will be returned to user (provided that it exists in the web server's file system).

## Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXWE0001</td>
<td>Specified URL is invalid.</td>
</tr>
<tr>
<td>BXWE0002</td>
<td>Specified URL contains invalid XML characters.</td>
</tr>
</tbody>
</table>

## Changelog

Version 8.4

- Updated: `web:response-header`: serialization method `raw` was removed (now obsolete).

Version 8.2


The module was introduced with Version 8.1.
Chapter 71. XQuery Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions for evaluating XQuery strings and modules at runtime.

Conventions

All functions in this module are assigned to the http://basex.org/modules/xquery namespace, which is statically bound to the xquery prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

Functions

xquery:eval

Updated with Version 8.5: base-uri option added.

Signatures

xquery:eval($query as xs:string) as item()*

xquery:eval($query as xs:string, $bindings as map(*)) as item()*

xquery:eval($query as xs:string, $bindings as map(*), $options as map(xs:string, xs:string)) as item()*

Summary

Evaluates the supplied $query string as XQuery expression and returns the resulting items. The evaluated query has its own query context. If a returned node is stored in a database, a main-memory copy will be returned as result, because the referenced database is closed after query execution and will not be accessible anymore. Variables and context items can be declared via $bindings. The specified keys must be QNames or strings:

- If a key is a QName, it will be directly adopted as variable name.
- It a key is a string, it may be prefixed with a dollar sign. Namespace can be specified using the Clark Notation.
- If the specified string is empty, the value will be bound to the context item.

The $options parameter contains evaluation options:

- permission: the query will be evaluated with the specified permissions (see User Management).
- timeout: query execution will be interrupted after the specified number of seconds.
- memory: query execution will be interrupted if the specified number of megabytes will be exceeded. This check works best if only one process is running at the same time.
- base-uri: set base-uri property for the query. This URI will be used when resolving relative URIs by functions such as fn:doc. Added with Version 8.5.

Errors

BXXQ0001: the query contains updating expressions. BXXQ0003: insufficient permissions for evaluating the query. BXXQ0004: query execution exceeded timeout or memory constraints. FOTY0013: the expression yields function items. Any other error that may occur while evaluating the query.

Examples

- xquery:eval("1+3") returns 4.
- You can bind the context and e.g. operate on a certain database only:

```xquery
xquery:eval("//country", map { '': db:open('factbook') })
```
The following expressions use strings as keys. All of them return 'XML':

```xquery
xquery:eval("." , map { '' : 'XML' } ),
xquery:eval("declare variable $xml external; $xml", map { 'xml': 'XML' } ),
xquery:eval("declare namespace pref='URI';
declare variable $pref:xml external; $pref:xml",
map { '{URI}xml': 'XML' } )
```

The following expressions use QNames as keys. All of them return 'XML':

```xquery
declare namespace pref = 'URI';
let $query := "declare namespace pref='URI';
declare variable $pref:xml external; $pref:xml"
let $vars := map { xs:QName('pref:xml'): 'XML' }
return xquery:eval($query, $vars)
```

**xquery:update**

**Signatures**

```xquery
xquery:update($query as xs:string) as item()*
xquery:update($query as xs:string, $bindings as map(*)) as item()*
xquery:update($query as xs:string, $bindings as map(*), $options as map(xs:string, xs:string)) as item()
```

**Summary**

Evaluates `$query` as updating XQuery expression at runtime. All updates will be added to the Pending Update List of the main query and performed after the evaluation of the main query. The semantics of the `$bindings` and `$options` parameters is the same as for `xquery:eval`.

**Errors**

BXXQ0002: the query contains no updating expressions. BXXQ0003: insufficient permissions for evaluating the query. BXXQ0004: query execution exceeded timeout or memory constraints. FOTY0013: the expression yields function items. Any other error that may occur while evaluating the query.

**xquery:parse**

**Signatures**

```xquery
xquery:parse($query as xs:string) as item()*
xquery:parse($query as xs:string, $options as map(*)) as item()
```

**Summary**

Parses the specified `$query` string as XQuery module and returns the resulting query plan. The `$options` parameter influences the output:

- `compile`: additionally compiles the query after parsing it. By default, this option is `false`.
- `plan`: returns an XML representation of the internal query plan. By default, this option is `true`. The naming of the expressions in the query plan may change over time.
- `pass`: passes on the original error info (line and column number, optional file uri). By default, this option is `false`.

**Errors**

Any error that may occur while parsing the query.

**Examples**

- `xquery:parse("1 + 3")` returns:
xquery:parse-uri

Signatures  xquery:parse-uri($uri as xs:string) as item()*

Summary  Parses the XQuery module located at $uri and returns the resulting query plan. The $options parameter is the same as for xquery:parse.

Errors  Any error that may occur while parsing the query.

xquery:invoke

Signatures  xquery:invoke($uri as xs:string) as item()*

Summary  Evaluates the XQuery module located at $uri at runtime and returns the resulting items. Database nodes in the result will be copied and returned instead. The semantics of the $bindings and $options parameters is the same as for xquery:eval.

Errors  BXXQ0001: the expression contains updating expressions. BXXQ0003: insufficient permissions for evaluating the query. BXXQ0004: query execution exceeded timeout. FOTY0013: the expression yields function items. Any other error that may occur while evaluating the query.

xquery:type

Signatures  xquery:type($expr as item()) as item()*

Summary  Similar to fn:trace($expr, $msg), but instead of a user-defined message, it emits the compile-time type and estimated result size of its argument.

Errors  

Code  Description

BXXQ0001  The specified query contains updating expressions.

BXXQ0002  The specified query contains no updating expressions.

BXXQ0003  Insufficient permissions for evaluating the query.

BXXQ0004  Query execution exceeded timeout.

Changelog

Version 8.5

• Updated: xquery:eval: base-uri option added

Version 8.4

• Added: xquery:parse-uri

• Updated: xquery:parse: pass option added
XQuery Module

Version 8.0

• Added: xquery:update, xquery:parse

• Deleted: xquery:evaluate (opened databases will now be closed by main query)

Version 7.8.2

• Added: $options argument

Version 7.8

• Added: xquery:evaluate

• Updated: used variables must be explicitly declared in the query string.

This module was introduced with Version 7.3. Functions have been adopted from the obsolete Utility Module.
Chapter 72. XSLT Module

This XQuery Module contains functions and variables to perform XSLT transformations. By default, this module uses Java's XSLT 1.0 Xalan implementation to transform documents. XSLT 2.0 is used instead if Version 9.x of the Saxon XSLT Processor (saxon9he.jar, saxon9pe.jar, saxon9ee.jar) is found in the classpath. A custom transformer can be specified by overwriting the system property javax.xml.transform.TransformerFactory, as shown in the following Java example:

```java
Context ctx = new Context();
String result = new XQuery("xslt:transform('...', '...')").execute(ctx);
...
ctx.close();
```

Conventions

All functions in this module are assigned to the http://basex.org/modules/xslt namespace, which is statically bound to the xslt prefix. All errors are assigned to the http://basex.org/errors namespace, which is statically bound to the bxerr prefix.

Functions

**xslt:processor**

**Signatures**

xslt:processor() as xs:string

**Summary**

Returns the name of the applied XSLT processor, or the path to a custom implementation (currently: "Java", "Saxon EE", "Saxon PE", or "Saxon HE").

**xslt:version**

**Signatures**

xslt:version() as xs:string

**Summary**

Returns the supported XSLT version (currently: "1.0" or "2.0"). "Unknown" is returned if a custom implementation was chosen.

**xslt:transform**

**Signatures**

xslt:transform($input as item(), $stylesheet as item()) as node()
xslt:transform($input as item(), $stylesheet as item(), $params as map(xs:string, xs:string)) as node()

**Summary**

Transforms the document specified by $input, using the XSLT template specified by $stylesheet, and returns the result as node. $input and $stylesheet can be specified as

- xs:string, containing the path to the document,
- xs:string, containing the document in its string representation, or
- node(), containing the document itself.

The $params argument can be used to bind variables to a stylesheet. Only strings are supported when using Saxon (XSLT 2.0).

**Error**

BXSL0001: an error occurred during the transformation process.
**xslt:transform-text**

**Signatures**

```xml
xslt:transform-text($input as item(), $stylesheet as item()) as xs:string
xslt:transform-text($input as item(), $stylesheet as item(), $params as map(xs:string, xs:string)) as xs:string
```

**Summary**

Transforms the document specified by $input, using the XSLT template specified by $stylesheet, and returns the result as string. The semantics of $params is the same as for xslt:transform.

**Error**

BXSL0001: an error occurred during the transformation process.

**Examples**

**Example 1: Basic XSL transformation with dummy document and without parameters**

**Query:**

```xml
xslt:transform-text(<dummy/>, 'basic.xslt')
```

**Result:**

```
123
```

**Example 2: XSLT transformation of an input document**

**Query:**

```xml
(: Outputs the result as html. :) 
declare option output:method 'html';
(: Turn whitespace chopping off. :) 
declare option db:chop 'no';

let $in :=
    <books>
        <book>
            <title>XSLT Programmer’s Reference</title>
            <author>Michael H. Kay</author>
        </book>
        <book>
            <title>XSLT</title>
            <author>Doug Tidwell</author>
            <author>Simon St. Laurent</author>
            <author>Robert Romano</author>
        </book>
    </books>

let $style :=
    <xsl:stylesheet version='2.0' xmlns:xsl='http://www.w3.org/1999/XSL/Transform'>
        <xsl:output method='xml'/>
    </xsl:stylesheet>

<html>
    <body>
        <div>
            <xsl:for-each select='books/book'>
                • <b><xsl:apply-templates select='title'/></b>: <xsl:value-of select='author'/><br/>
            </xsl:for-each>
        </div>
    </body>
</html>
```

```xml
123
• XSLT Programmer’s Reference: Michael H. Kay
• XSLT: Doug Tidwell
  Simon St. Laurent
  Robert Romano
```
Example 3: Assigning a variable to an XSLT stylesheet

Query:

let $in := <dummy/>
let $style := doc('variable.xsl')
return (  
xslt:transform($in, $style, <xslt:parameters><xslt:v>1</xslt:v></xslt:parameters>),
  xslt:transform($in, $style, map { "v": 1 })
 )

variable.xslt

<xsl:stylesheet version='1.0' xmlns:xsl='http://www.w3.org/1999/XSL/Transform'>
  <xsl:param name='v'/>
  <xsl:template match='/'
    <v><xsl:value-of select='$v'/></v>
  </xsl:template>
</xsl:stylesheet>

Result:

<v>1</v>
<v>1</v>

Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXSL0001</td>
<td>An error occurred during the transformation process.</td>
</tr>
</tbody>
</table>

Changelog

Version 7.6

- Added: xslt:transform-text
- Updated: xslt:transform returned error code
Version 7.3

- Updated: $xslt:processor → xslt:processor, $xslt:version → xslt:version
Chapter 73. ZIP Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions to handle ZIP archives. The contents of ZIP files can be extracted and listed, and new archives can be created. The module is based on the EXPath ZIP Module. Please note that the ZIP module is not being actively maintained but is still distributed for compatibility with older applications. We recommend you use the Archive Module wherever possible.

Conventions

All functions in this module are assigned to the http://expath.org/ns/zip namespace, which is statically bound to the zip prefix. All errors are assigned to the http://expath.org/ns/error namespace, which is statically bound to the experr prefix.

Functions

zip:binary-entry

<table>
<thead>
<tr>
<th>Signatures</th>
<th>zip:binary-entry($uri as xs:string, $path as xs:string) as xs:base64Binary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Extracts the binary file at $path within the ZIP file located at $uri and returns it as an xs:base64Binary item.</td>
</tr>
<tr>
<td>Errors</td>
<td>ZIP0001: the specified path does not exist.ZIP0003: the operation fails for some other reason.</td>
</tr>
</tbody>
</table>

zip:text-entry

<table>
<thead>
<tr>
<th>Signatures</th>
<th>zip:text-entry($uri as xs:string, $path as xs:string) as xs:string</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>zip:text-entry($uri as xs:string, $path as xs:string, $encoding as xs:string) as xs:string</td>
</tr>
<tr>
<td>Summary</td>
<td>Extracts the text file at $path within the ZIP file located at $uri and returns it as an xs:string item. An optional encoding can be specified via $encoding.</td>
</tr>
<tr>
<td>Errors</td>
<td>ZIP0001: the specified path does not exist.ZIP0003: the operation fails for some other reason.</td>
</tr>
</tbody>
</table>

zip:xml-entry

<table>
<thead>
<tr>
<th>Signatures</th>
<th>zip:xml-entry($uri as xs:string, $path as xs:string) as document-node()</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Extracts the XML file at $path within the ZIP file located at $uri and returns it as a document node.</td>
</tr>
<tr>
<td>Errors</td>
<td>FODC0006: the addressed file is not well-formed.ZIP0001: the specified path does not exist.ZIP0003: the operation fails for some other reason.</td>
</tr>
</tbody>
</table>

zip:html-entry

<table>
<thead>
<tr>
<th>Signatures</th>
<th>zip:html-entry($uri as xs:string, $path as xs:string) as document-node()</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Extracts the HTML file at $path within the ZIP file located at $uri and returns it as a document node. The file is converted to XML first if Tagsoup is found in the classpath.</td>
</tr>
<tr>
<td>Errors</td>
<td>FODC0006: the addressed file is not well-formed, or cannot be converted to correct XML.ZIP0001: the specified path does not exist.ZIP0003: the operation fails for some other reason.</td>
</tr>
</tbody>
</table>
### zip:entries

**Signatures**

```xml
zip:entries($uri as xs:string) as element(zip:file)
```

**Summary**

Generates an **ZIP XML Representation** of the hierarchical structure of the ZIP file located at $uri and returns it as an element node. The file contents are not returned by this function.

**Errors**

- ZIP0001: the specified path does not exist.
- ZIP0003: the operation fails for some other reason.

**Examples**

If the ZIP archive `archive.zip` is empty, `zip:entries('archive.zip')` returns:

```xml
```

### zip:zip-file

**Signatures**

```xml
zip:zip-file($zip as element(zip:file)) as empty-sequence()
```

**Summary**

Creates a new ZIP archive with the characteristics described by $zip, the **ZIP XML Representation**.

**Errors**

- ZIP0001: an addressed file does not exist.
- ZIP0002: entries in the ZIP archive description are unknown, missing, or invalid.
- ZIP0003: the operation fails for some other reason.
- Serialization Errors: an inlined XML fragment cannot be successfully serialized.

**Examples**

The following function creates a file `archive.zip` with the file `file.txt` inside:

```xml
zip:zip-file(
    <file xmlns="http://expath.org/ns/zip" href="archive.zip">
        <entry src="file.txt"/>
    </file>)
```

The following function creates a file `archive.zip`. It contains one file `readme` with the content "thanks":

```xml
zip:zip-file(
    <file xmlns="http://expath.org/ns/zip" href="archive.zip">
        <entry name="readme">thanks</entry>
    </file>)
```

### zip:update-entries

**Signatures**

```xml
zip:update-entries($zip as element(zip:file), $output as xs:string) as empty-sequence()
```

**Summary**

Updates an existing ZIP archive or creates a modified copy, based on the characteristics described by $zip, the **ZIP XML Representation**. The $output argument is the URI where the modified ZIP file is copied to.

**Errors**

- ZIP0001: an addressed file does not exist.
- ZIP0002: entries in the ZIP archive description are unknown, missing, or invalid.
- ZIP0003: the operation fails for some other reason.
- Serialization Errors: an inlined XML fragment cannot be successfully serialized.

**Examples**

The following function creates a copy `new.zip` of the existing `archive.zip` file:

```xml
zip:update-entries(zip:entries('archive.zip'), 'new.zip')
```

The following function deletes all PNG files from `archive.zip`:

```xml
declare namespace zip = "http://expath.org/ns/zip";
copy $doc := zip:entries('archive.zip')
modify delete node $doc//zip:entry[ends-with(lower-case(@name), '.png')]
return zip:update-entries($doc, 'archive.zip')
```
## Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZIP0001</td>
<td>A specified path does not exist.</td>
</tr>
<tr>
<td>ZIP0002</td>
<td>Entries in the ZIP archive description are unknown, missing, or invalid.</td>
</tr>
<tr>
<td>ZIP0003</td>
<td>An operation fails for some other reason.</td>
</tr>
</tbody>
</table>
Part VII. Developing
Chapter 74. Developing

Read this entry online in the BaseX Wiki.

This page is one of the Main Sections of the documentation. It provides useful information for developers. Here you can find information on various alternatives to integrate BaseX into your own project.

Integrate & Contribute

- **Eclipse**: Compile and run BaseX from within Eclipse
- **Git**: Learn how to work with Git
- **Maven**: Embed BaseX into your own projects
- **Releases**: Official releases, snapshots, old versions
- **Translations**: Contribute a new translation to BaseX!

Project (Code, Features, Questions)

- The **Source Code** can be browsed online
- The **Issue Tracker** contains feature requests and bug reports
- For questions on the project, please write to our **mailing list**

Web Application

- **RESTXQ**: Write web services with XQuery
- **REST**: Access and update databases via HTTP requests
- **WebDAV**: Access databases from your filesystem
- **XForms**: Build browser forms with XML technologies

APIs

- **Clients**: Communicate with BaseX using C#, PHP, Python, Perl, C, ...
- **Java Examples**: Code examples for developing with BaseX
- **Docker**: Isolate BaseX in a docker container
- **XQJ API**, implemented by Charles Foster
- **XQuery for Scala API**, based on XQJ and written by Dino Fancellu
Chapter 75. Developing with Eclipse

Read this entry online in the BaseX Wiki.

This page is part of the Developer Section. It describes how to get the BaseX sources compiled and running on your system.

Another article in the documentation describes how to use BaseX as a query processor in Eclipse.

Prerequisites

BaseX is developed with the Eclipse environment (other IDEs like IntelliJ IDEA can be used as well). The Eclipse IDE for Java Developers includes the EGit plugin (for Git) and the m2e plugin (for Maven).

Other Eclipse plugins we use are:

<table>
<thead>
<tr>
<th>Project</th>
<th>Description</th>
<th>Update URL</th>
<th>Eclipse Marketplace</th>
</tr>
</thead>
<tbody>
<tr>
<td>FindBugs</td>
<td>Analyze project at byte code level</td>
<td><a href="http://findbugs.cs.umd.edu/eclipse">http://findbugs.cs.umd.edu/eclipse</a></td>
<td>install</td>
</tr>
<tr>
<td>UCDetector</td>
<td>Unnecessary code detector</td>
<td><a href="http://ucdetector.sourceforge.net/update">http://ucdetector.sourceforge.net/update</a></td>
<td>install</td>
</tr>
</tbody>
</table>

Check Out

Our Git Tutorial explains how BaseX can be checked out from the GitHub Repository and embedded in Eclipse with EGit. The article also demonstrates how git can be used on command-line.

The basex repository contains the following sub-directories:

1. basex-core is the main project
2. basex-api contains the BaseX APIs (XML:DB, bindings in other languages) and HTTP Services (REST, RESTXQ, WebDAV)
3. basex-examples includes some examples code for BaseX
4. basex-tests contains several unit and stress tests

If the "Problems" View contains errors or warnings, you may need to switch to Java 7 (Windows → Preferences → Installed JREs). With the Maven plugin from Eclipse, it sometimes requires several attempts to get all dependencies updated. This loop can be avoided if the sources are precompiled via Maven on command-line.

Start in Eclipse

1. Press Run → Run...
2. Create a new "Java Application" launch configuration
3. Select "basex" as "Project"
4. Choose a "Main class" (e.g., org.basex.BaseXGUI for the graphical user interface)
5. Launch the project via *Run*

**Alternative**

You may as well use the standalone version of *Maven* to compile and run the project, use other IDEs such as *IntelliJ IDEA*. 
Chapter 76. Git

Read this entry online in the BaseX Wiki.

This page is part of the Developer Section. It describes how to use git to manage the BaseX sources.

Using Git to contribute to BaseX

Our team uses git and GitHub to manage the source code. All team members have read+write access to the repository, and external contributors are welcome to fork the project.

Git makes it easy to retain a full copy of the repository for yourself. To get started and running, simply fork BaseX:

1. Head over to https://github.com and create an account
2. Fork https://github.com/BaseXdb/basex, so you have a version on your own
3. The forked project can then be cloned on your local machine, and changes can be pushed back to your remote repository
4. Open Eclipse
5. Install egit (Eclipse: Help → Marketplace → Search for egit or get it from http://www.eclipse.org/egit/)

Using Git & Eclipse

To clone the project from within Eclipse, you may need to install EGit first (Eclipse: Help → Marketplace → Search for egit or get it from http://www.eclipse.org/egit/).

Clone

- In the Package Explorer to the left use right-click and choose Import...
- Select "Projects from Git" and click Next >
- Click "Clone..." to create a local copy of the remote repository. This copy will include the full project history
- Copy & Paste the GitHub URI in the Location field. If you want to use SSH make sure you provided GitHub with your public key to allow write-access. If in doubt use the HTTPS URI and authenticate yourself with your GitHub credentials. The read-only URI of the repository is https://github.com/BaseXdb/basex.git.
- Select the master branch (or arbitrary branches you like)
- Now choose a location where the local repository is stored: Create <workspace>/repos/BaseX and click "Finish".

Create the project

- Select our newly cloned repository and click Next
- Select "Import Existing Projects" and depending on your Eclipse version enable automatic sharing. More
recent versions will not offer this feature as sharing is enabled by default.

- Click next to select the Project to import
- Check "basex" to checkout and click finish
- You are now ready to contribute.

EGit & SSH
EGit uses the JSch library which is, however, reported to have problems with RSA SSH keys in Linux and possibly other platforms. A solution would be to use the variable GIT_SSH and assign it a path to the native SSH executable. According to this change in EGit, the plugin will try to use a native SSH implementation instead of JSch (this, however, may not always work either :( ).

Using Git on Command-Line

Note: this is not intended to be a complete git reference; it's purpose is to quickly introduce BaseX developers to the most commonly used git commands in the context of the BaseX project.

Preparation
1. Create a GitHub user account: here (your github user name will be referenced as $username)
2. Set up SSH access to GitHub as described here
3. Create a fork of one of the BaseXdb projects (it will be referenced as $project)
4. Choose a directory where the project will be created and make it your working directory (e.g. /home/user/myprojects)

Clone Your Personal Repository

```
$ git clone git@github.com:$username/$project.git
Cloning into $project...
Enter passphrase for key '/home/user/.ssh/id_rsa': ...
```

```
$ ls -d -l $PWD/*
/home/user/myprojects/$project
```

Note that git automatically creates a directory where the repository content will be checked out.

List Remote Repositories

```
$ git remote -v
origin git@github.com:$username/$project.git (fetch)
```
origin git@github.com:$username/$project.git (push)

Currently, there is only one remote repository; it is automatically registered during the clone operation. Git remembers this repository as the default repository for push/pull operations.

List Local Changes

After some files have been changed locally, the changes can be seen as follows:

```bash
$ git diff
diff --git a/readme.txt b/readme.txt
index fabaeaa..cd09568 100644
--- a/readme.txt
+++ b/readme.txt
@@ -49,6 +49,10 @@ ADDING CHECKSTYLE
--------------------------------------------------------------
- Enter the URL: http://eclipse-cs.sourceforge.net/update
- Follow the installation procedure and restart Eclipse
+USING GIT
----------------------------------------------------------------------
Any kind of feedback is welcome; please check out the online documentation at
Commit to Local Repository

Note: this commit operation does not commit into the remote repository!

First, it is needed to select the modified files which should be committed:

```bash
$ git add readme.txt
```

Then perform the actual commit:

```bash
$ git commit
[master 0fde1fb] Added TODO in section "USING GIT"
1 files changed, 4 insertions(+), 0 deletions(-)
```

Before executing the actual commit, git will open the default shell editor (determined using the $EDITOR variable, usually vi) to enter a message describing the commit changes.

Alternative way is to commit all changed files, i.e. it is not needed to explicitly add the changed files:

```bash
$ git commit -a
```
Pushing Local Changes to Remote Repository

$ git push
Enter passphrase for key '/home/user/.ssh/id_rsa':
Everything up-to-date

Pulling Changes from Remote Repository

$ git pull
Enter passphrase for key '/home/user/.ssh/id_rsa':
Already up-to-date.

Add BaseXdb Upstream Repository

The upstream repository is the one from which the BaseX releases are made and the one from which the personal repository was forked.

$ git remote add upstream
git@github.com:BaseXdb/$project.git

$ git remote -v
origin git@github.com:$username/$project.git (fetch)
origin git@github.com:$username/$project.git (push)
upstream git@github.com:BaseXdb/$project.git (fetch)
upstream git@github.com:BaseXdb/$project.git (push)

Pulling Changes from Upstream to Local Repository

When some changes are made in the upstream repository, they can be pulled to the local repository as follows:

$ git pull upstream master
Enter passphrase for key '/home/user/.ssh/id_rsa':
From github.com:BaseXdb/$project
 * branch            master     -> FETCH_HEAD
Already up-to-date.

The changes can then be pushed in the personal repository:
Developing a new feature or bug fix using git

It is always a good idea to create a new branch for a new feature or a bug fix you are working on. So first, let's make sure you have the most up-to-date source code. We assume, that you added BaseX as upstream repository as described above and you are currently in the *master* branch:

```
$ git pull upstream master
```

Now, we create a new branch, based on the master branch

```
$ git checkout -b new-feature
Switched to a new branch 'new-feature'
```

Your are now automatically switched to the *new-feature* branch. Now you can make all your changes in one or several commits. You can commit all changes using

```
$ git commit -a
```

Now, you want to push these changes to the repository on GitHub. Remember, that up to now your changes just reside on your local drive, so now you want to push it to your remote fork of BaseX. Simply do:

```
$ git push origin new-feature
```

You can now use your web browser and go to your fork of BaseX. You will see the following message:

You can now click the "Compare & pull request" button. You can now review the changes you are going to push.

Please review them carefully. Also, please give a meaningful comment so we can quickly determine what your changes are doing. After clicking the "Create Pull request" button you are done and we will review your changes and either merge the pull request or get back to you.

**Need help using git?**

**Installing**

For information on how to install git on various platforms please refer to: GitHub: git Installation Guide
Documentation

- Comprehensive Getting Starting Guide on GitHub
- The git book
- Gitcasts.com – Video Guides
Chapter 77. Maven

Read this entry online in the BaseX Wiki.

This page is part of the Developer Section. It demonstrates how Maven is used to compile and run BaseX, and embed it into other projects.

Using Maven

If you have cloned our repository and installed Maven on your machine, you can run the following commands from all local repository directories:

• *mvn compile*: the BaseX source files are compiled.

• *mvn package*: JAR archives are created in the *target* class directory, and all relevant libraries are created in the *lib* directory. Packaging is useful if you want to use the start scripts.

• *mvn install*: the JAR archive is installed to the local repository, and made available to other Maven projects. This is particularly useful if you are compiling a beta version of BaseX, for which no archives exist in the repositories.

By adding the flag *-DskipTests* you can skip the JUnit tests and speed up packaging. You may as well use Eclipse and m2eclipse to compile the BaseX sources.

There are several alternatives for starting BaseX:

• type in *java -cp target/classes org.basex.BaseX* in the *basex-core* directory to start BaseX on the command-line mode,

• type in *mvn jetty:run* in the *basex-api* directory to start BaseX with Jetty and the HTTP servers,

• run one of the Start Scripts contained in the etc directory

Artifacts

You can easily embed BaseX into your own Maven projects by adding the following XML snippets to your *pom.xml* file:

```xml
<repositories>
  <repository>
    <id>basex</id>
    <name>BaseX Maven Repository</name>
    <url>http://files.basex.org/maven</url>
  </repository>
</repositories>

<BaseX Main Package>
<dependency>
  <groupId>org.basex</groupId>
  <artifactId>basex</artifactId>
  <version>7.6</version>
</dependency>

<APIs and Services>
...including APIs and the REST, RESTXQ and WebDAV services:
<dependency>
  <groupId>org.basex</groupId>
  <artifactId>basex-api</artifactId>
</dependency>
```
<version>7.6</version>
</dependency>

**XQJ API**

The XQJ API is hosted at [http://xqj.net](http://xqj.net):

```xml
<repository>
  <id>xqj</id>
  <name>XQJ Maven Repository</name>
  <url>http://xqj.net/maven</url>
</repository>
...  
<dependency>
  <groupId>net.xqj</groupId>
  <artifactId>basex-xqj</artifactId>
  <version>1.2.0</version>
</dependency>
<dependency>
  <groupId>com.xqj2</groupId>
  <artifactId>xqj2</artifactId>
  <version>0.1.0</version>
</dependency>
<dependency>
  <groupId>javax.xml.xquery</groupId>
  <artifactId>xqj-api</artifactId>
  <version>1.0</version>
</dependency>
```
Chapter 78. Releases

Read this entry online in the BaseX Wiki.

This page is part of the Developer Section. It lists the official locations of major and minor BaseX versions:

**Official Releases**

Our releases, packaged for various platforms, are linked from our homepage. They are updated every 2-8 weeks:

- [http://basex.org/download](http://basex.org/download)

Our file server contains links to older releases as well (but we recommend everyone to stay up-to-date, as you’ll get faster feedback working with the latest version):

- [http://files.basex.org/releases](http://files.basex.org/releases)

**Stable Snapshots**

If you are a developer, we recommend you to regularly download one of our stable snapshots, which are packaged and uploaded several times a week:

- [http://files.basex.org/releases/latest/](http://files.basex.org/releases/latest/)

Note that the offered snapshot files are replaced as soon as newer versions are available.

**Code Base**

If you always want to be on the cutting edge, you are invited to watch and clone our GitHub repository:

- [https://github.com/BaseXdb/](https://github.com/BaseXdb/)

We do our best to keep our main repository stable as well.

**Maven Artifacts**

The official releases and the current snapshots of both our core and our API packages are also deployed as Maven artifacts on our file server at regular intervals:

- [http://files.basex.org/maven/org/basex/](http://files.basex.org/maven/org/basex/)

**Linux**

BaseX can also be found in some Linux distributions, such as Debian, Ubuntu and archlinux (Suse and other distributions will follow soon):

- Debian: [http://packages.debian.org/sid/basex](http://packages.debian.org/sid/basex)
- Ubuntu: [http://launchpad.net/ubuntu/+source/basex](http://launchpad.net/ubuntu/+source/basex)
Chapter 79. Translations

Read this entry online in the BaseX Wiki.

This page is part of the Developer Section. It describes how to translate BaseX into other (natural) languages.

Thanks to the following contributors, BaseX is currently available in 10 languages:

- **Dutch**: Huib Verweij
- **English**: BaseX Team
- **French**: Maud Ingarao
- **German**: BaseX Team
- **Hungarian**: Kiss-Kálmán Dániel
- **Indonesian**: Andria Arisal
- **Italian**: Massimo Franceschet
- **Japanese**: Toshio HIRAI and Kazuo KASHIMA
- **Mongolian**: Tuguldur Jamiyansharav
- **Romanian**: Adrian Berila
- **Russian**: Oleksandr Shpak and Max Shamaev
- **Spanish**: Carlos Marcos

It is easy to translate BaseX into your native language! This is how you can proceed:

### Working with the sources

If you have downloaded all BaseX sources via Eclipse or Git, you may proceed as follows:

All language files are placed in the `src/main/resources/lang` directory of the main project:

1. Create a copy of an existing translation file (e.g., `English.lang`) and rename it to your target language (e.g., `Hawaiian.lang`).
2. Enter your name and contact information in the second line.
3. If you are using Eclipse, refresh the project (via Project → Refresh); if you are using Maven, type in `mvn compile`. Your new language file will be automatically detected.
4. Start the BaseX GUI, choose your language via Options → Preferences... and close the GUI.
5. Translate the texts in your language file and restart BaseX in order to see the changes.
6. Repeat the last step if you want to revise your translations.

If new strings are added to BaseX, they will automatically be added to your language files in English. The history view in GitHub is helpful to see which strings have recently been updated to a file.

### Updating BaseX.jar

You can directly add new languages to the JAR file. JAR files are nothing else than ZIP archives, and all language files are placed in the `lang` directory into the JAR file:
1. Unzip an existing translation file (e.g., English.lang) and rename it to your target language (e.g., Hawaiian.lang)

2. Enter your name and contact information in the second line and translate the texts

3. Update your JAR file by copying the translated file into the zipped lang directory. Your new language file will be automatically detected.

4. Start BaseX.jar, choose your language via Options → Preferences... and restart BaseX to see the changes

You can also directly assign a language in the .basex configuration file, which is placed in your home directory. The language is assigned to the LANG option. In order to see where the text keys are used within BaseX, you can set LANGKEY to true.
Part VIII. HTTP Services
Chapter 80. RESTXQ

Read this entry online in the BaseX Wiki.

This page presents one of the Web Application services. It describes how to use the RESTXQ API of BaseX.

RESTXQ, introduced by Adam Retter, is an API that facilitates the use of XQuery as a server-side processing language for the Web. RESTXQ has been inspired by Java’s JAX-RS API: it defines a pre-defined set of XQuery 3.0 annotations for mapping HTTP requests to XQuery functions, which in turn generate and return HTTP responses.

Please note that BaseX provides various extensions to the original draft of the specification:

- Multipart types are supported, including `multipart/form-data`
- A `%rest:error` annotation can be used to catch XQuery errors
- Servlet errors can be redirected to other RESTXQ pages
- A RESTXQ Module provides some helper functions
- Parameters are implicitly cast to the type of the function argument
- The Path Annotation can contain regular expressions
- Quality factors in the Accept header will be evaluated
- `%input` annotations, support for input-specific content-type parameters
- `%rest:single` annotation to cancel running RESTXQ functions

Introduction

Preliminaries

The RESTXQ service is accessible via `http://localhost:8984/`

All RESTXQ annotations are assigned to the `http://exquery.org/ns/restxq` namespace, which is statically bound to the `rest` prefix. A Resource Function is an XQuery function that has been marked up with RESTXQ annotations. When an HTTP request comes in, a resource function will be invoked that matches the constraints indicated by its annotations.

Whenever a RESTXQ URL is requested, the RESTXQPATH module directory and its sub-directories will be parsed for functions with RESTXQ annotations in library modules (detected by the extension `.xqm`) and main modules (detected by `.xq`). In main expressions, the main module will never be evaluated. All modules will be cached and parsed again when their timestamp changes.

A sub-directory will not be parsed for RESTXQ files if it contains a file named `.ignore`. Module caching can be turned on by setting `CACHERESTXQ` to true. The option is helpful in productive environments with a high load, but files should not be replaced while the web server is running. If files are still replaced, the RESTXQ module cache must be invalidated by calling the static root path `/init`.

Examples

A first RESTXQ function is shown below:

```xquery
module namespace page = 'http://baseX.org/examples/web-page';
```
declare %rest:path("hello/\{$who\}\") %rest:GET function page:hello($who) {
  <response>
    <title>Hello { $who }!</title>
  </response>
};

If the URI http://localhost:8984/hello/World is accessed, the result will be:

<response>
  <title>Hello World!</title>
</response>

The next function demonstrates a POST request:

declare
  %rest:path("/form")
  %rest:POST
  %rest:form-param("message","\{$message\}", "(no message)"
  %rest:header-param("User-Agent","\{$agent\}"
function page:hello-postman(
  $message as xs:string,
  $agent as xs:string*)
  as element(response) {
  <response type='form'>
    <message>{ $message }</message>
    <user-agent>{ $agent }</user-agent>
  </response>
};

If you post something (e.g. using curl or the embedded form at http://localhost:8984/)...
curly brackets, and map the corresponding segments of the request path to the arguments of the resource function. The first slash in the path is optional.

The following example contains a path annotation with three segments and two templates. One of the function arguments is further specified with a data type, which means that the value for $\text{variable}$ will be cast to an \texttt{xs:integer} before being bound:

```xml
declare %rest:path("/a/path/{$with}/some/{$variable}"
    function page:test($with, $variable as xs:integer) { ... };
```

Variables can be enhanced by regular expressions:

```xml
!(: Matches all paths with "app" as first, a number as second, and "order" as third segment :)
declare %rest:path("app/{$code=[0-9]+}/order")
    function page:order($full-path) { ... };

!(: Matches all other all paths starting with "app/" :)
declare %rest:path("app/{$path=+.}"
    function page:others($path) { ... };
```

**Content Negotiation**

Two following annotations can be used to restrict functions to specific content types:

- **HTTP Content Types**: a function will only be invoked if the HTTP Content-Type header of the request matches one of the given mime types. Example:

  ```xml
  %rest:consumes("application/xml", "text/xml")
  ```

- **HTTP Accept**: a function will only be invoked if the HTTP Accept header of the request matches one of the defined mime types. Example:

  ```xml
  %rest:produces("application/atom+xml")
  ```

By default, both mime types are */*/*. Quality factors supplied by a client will also be considered in the path selection process. If a client supplies the following accept header...

```xml
*/q=0.5,text/html;q=1.0
```

...and if two RESTXQ functions exist with the same path annotation and the produces annotations */* and text/html, respectively, the function with the second annotation will be called, because the quality factor for text/html documents is higher than the one for arbitrary other mime types.

Note that this annotation will not affect the content-type of the HTTP response. Instead, you will need to add a %output:media-type annotation.

**HTTP Methods**

**Default Methods**

The HTTP method annotations are equivalent to all HTTP request methods except TRACE and CONNECT. Zero or more methods may be used on a function; if none is specified, the function will be invoked for each method.

The following function will be called if GET or POST is used as request method:

```xml
declare %rest:GET %rest:POST %rest:path("/post")
    function page:post() { "This was a GET or POST request" };
```
The POST and PUT annotations may optionally take a string literal in order to map the HTTP request body to a function argument. Once again, the target variable must be embraced by curly brackets:

```xml
declare %rest:PUT("{$body}") %rest:path("/put")
function page:put($body) { "Request body: " || $body };
```

**Custom Methods**

Custom HTTP methods can be specified with the `%rest:method` annotation:

```xml
declare %rest:method("RETRIEVE")
function page:retrieve() { "RETRIEVE was specified as request method." };
```

**Content Types**

The body of a POST or PUT request will be converted to an XQuery item. Conversion can be controlled by specifying a content type. It can be further influenced by specifying additional content-type parameters:

<table>
<thead>
<tr>
<th>Content-Type</th>
<th>Parameters (name=value)</th>
<th>Type of resulting XQuery item</th>
</tr>
</thead>
<tbody>
<tr>
<td>text/xml, application/xml</td>
<td></td>
<td>document-node()</td>
</tr>
<tr>
<td>text/*</td>
<td></td>
<td>xs:string</td>
</tr>
<tr>
<td>application/json</td>
<td>JSON Options</td>
<td>document-node() or map(*)</td>
</tr>
<tr>
<td>text/html</td>
<td>HTML Options</td>
<td>document-node()</td>
</tr>
<tr>
<td>text/comma-separated-values</td>
<td>CSV Options</td>
<td>document-node() or map(*)</td>
</tr>
<tr>
<td>others</td>
<td></td>
<td>xs:base64Binary</td>
</tr>
<tr>
<td>multipart/*</td>
<td></td>
<td>sequence (see next paragraph)</td>
</tr>
</tbody>
</table>

For example, if `application/json; lax=yes` is specified as content type, the input will be transformed to JSON, and the lax QName conversion rules will be applied, as described in the JSON Module.

**Input options**

Conversion options for JSON, CSV and HTML can also be specified via annotations using the `input` prefix. The following function treats the first line of the textual input as CSV header:

```xml
declare %rest:path("/store.csv") %rest:POST("{$csv}") %input:csv("header=true")
function page:store-csv($csv as document-node()) { 
  "Number of rows: " || count($csv/csv/record)
};
```

**Multipart Types**

The single parts of a multipart message are represented as a sequence, and each part is converted to an XQuery item as described in the last paragraph.

A function that is capable of handling multipart types is identical to other RESTXQ functions:

```xml
declare %rest:path("/multipart") %rest:POST("{$data}")
```
 PARAMETERS

The following annotations can be used to bind request values to function arguments. Values will implicitly be cast to the type of the argument.

 QUERY PARAMETERS

The value of the first parameter, if found in the query component, will be assigned to the variable specified as second parameter. If no value is specified in the HTTP request, all additional parameters will be bound to the variable (if no additional parameter is given, an empty sequence will be bound):

 ```xml
<result id="{$id}" sum="{ sum($add) }"/>
```

 HTML Form Fields

Form parameters are specified the same way as query parameters. Their values are the result of HTML forms submitted with the content type application/x-www-form-urlencoded.

 ```xml
%rest:form-param("parameter", "{$value}", "default")
```

 FILE Uploads

Files can be uploaded to the server by using the content type multipart/form-data (the HTML5 multiple attribute enables the upload of multiple files):

 ```html
<form action="/upload" method="POST" enctype="multipart/form-data">
  <input type="file" name="files" multiple="multiple"/>
  <input type="submit"/>
</form>
```

The file contents are placed in a map, with the filename serving as key. The following example shows how uploaded files can be stored in a temporary directory:

 ```xml
declare
%rest:POST
%rest:path("/upload")
%rest:form-param("files", "{$files}")
function page:upload($files)
{
  for $name in map:keys($files)
  let $content := $files($name)
  let $path := file:temp-dir() || $name
  return (
    file:write-binary($path, $content),
    <file name="{$name}" size="{ file:size($path) }"/>
  )
}
```
HTTP Headers

Header parameters are specified the same way as query parameters:

```
%rest:header-param("User-Agent", "{$user-agnt}")
%rest:header-param("Referer", "{$referer}", "none")
```

Cookies

Cookie parameters are specified the same way as query parameters:

```
%rest:cookie-param("username", "{$user}")
%rest:cookie-param("authentication", "{$auth}", "no_auth")
```

Query Execution

In many RESTXQ search scenarios, input from browser forms is processed and search results are returned. User experience can generally be made more interactive if an updated search request is triggered with each key click. However, this may lead to many expensive parallel requests, from which only the result of the last request will be relevant for the client.

With the `%rest:single` annotation, it can be enforced that only one instance of a function will be executed for the same client. If the same function will be called for the second time, the already running query will be stopped, and the HTTP error code 410 (Gone) will be returned instead:

```
(! If fast enough, returns the result. Otherwise, if called again, raises 410 :) 
declare 
  %rest:path("/search")
  %rest:query-param("term", "{$term}")
  %rest:single
function page:search($term as xs:string) {
  <ul>{
    for $result in db:open('large-db')//*[text() = $term]
      return <li>{ $result }</li>
  }</ul>
}
```

By specifying a string along with the annotation, functions can be bundled together, and one request can be canceled by calling another one.

This is shown by another example, in which the first function can be interrupted by the second one. If you call both functions in separate browser tabs, you will note that the first tab will return 410, and the second one will return `<xml>stopped</xml>`.

```
declare 
  %rest:path("/compute")
  %rest:single("EXPENSIVE")
function local:compute() {
  (1 to 100000000000000). = 0
}
```

```
declare 
  %rest:path("/stop")
  %rest:single("EXPENSIVE")
function local:stop() {
  <xml>stopped</xml>
}
```

The following things should be noted:
If a query will be canceled, there will be no undesirable side-effects. For example, it won’t be possible to kill a query if it is currently updating the database or performing any other I/O operations. As a result, the termination of a running query can take some more time as expected.

The currently executed function is bound to the current session. This way, a client will not be able to cancel requests from other clients. As a result, functions can only be stopped if there was at least one previous successful response, in which initial session data was returned to the client.

**Response**

By default, a successful request is answered with the HTTP status code **200** (OK) and is followed by the given content. An erroneous request leads to an error code and an optional error message (e.g. **404** for “resource not found”).

**Custom Response**

Custom responses can be built from within XQuery by returning an **rest:response** element, an **http:response** child node that matches the syntax of the **EXPath HTTP Client Module** specification, and more optional child nodes that will be serialized as usual. A function that reacts on an unknown resource may look as follows:

```xml
declare %rest:path("*") function page:error404() {
  <rest:response>
    <http:response status="404" message="I was not found.">
      <http:header name="Content-Language" value="en"/>
      <http:header name="Content-Type" value="text/html; charset=utf-8"/>
    </http:response>
  </rest:response>
};
```

**Forwards and Redirects**

The two XML elements **rest:forward** and **rest:redirect** can be used in the context of RESTXQ. These nodes allow e.g. multiple XQuery Updates in a row by redirecting to the RESTXQ path of updating functions. Both wrap a URL to a RESTXQ path. The wrapped URL should be properly encoded via **fn:encode-for-uri**().

Note that, currently, these elements are not part of RESTXQ specification.

**rest:forward**

Usage: wrap the location as follows

```xml
<rest:forward>{ $location }</rest:forward>
```

This results in a server-side forwarding, which as well reduces traffic among client and server. A forwarding of this kind will not change the URL seen from the client’s perspective.

As an example, returning

```xml
<rest:forward>/hello/universe</rest:forward>
```

would internally forward to **http://localhost:8984/hello/universe**

**rest:redirect**

The function **web:redirect** can be used to create a redirect response element. Alternatively, the following element can be sent:

```xml
<rest:redirect>{ $location }</rest:redirect>
```
It is an abbreviation for:

```xml
<rest:response>
  <http:response status="302">
    <http:header name="location" value="{ $location }">
  </http:response>
</rest:response>
```

The client decides whether to follow this redirection. Browsers usually will, tools like `curl` won't unless `-L` is specified.

**Output**

The content-type of a response can be influenced by the user via Serialization Parameters. The steps are described in the REST chapter. In RESTXQ, serialization parameters can be specified in the query prolog, via annotations, or within the REST response element:

**Query Prolog**

In main modules, serialization parameters may be specified in the query prolog. These parameters will then apply to all functions in a module. In the following example, the content type of the response is overwritten with the media-type parameter:

```restxq
declare option output:media-type 'text/plain';
```

**Annotations**

Global serialization parameters can be overwritten via `%output` annotations. The following example serializes XML nodes as JSON, using the JsonML format:

```restxq
declare
  %rest:path("cities")
  %output:method("json")
  %output:json("format=jsonml")
function page:cities() {
  element cities {
    db:open('factbook')//city/name
  }
}
```

The next function, when called, generates XHTML headers, and `text/html` will be set as content type:

```restxq
declare
  %rest:path("done")
  %output:method("xhtml")
  %output:omit-xml-declaration("no")
  %output:doctype-public("-//W3C//DTD XHTML 1.0 Transitional//EN")
  %output:doctype-system("http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd")
function page:html() {
  <html xmlns="http://www.w3.org/1999/xhtml">
    <body>done</body>
  </html>
}
```
Response Element

Serialization parameters can also be specified in a REST response element in a query. Serialization parameters will be overwritten:

```xquery
declare %rest:path("version3") function page:version3() {
    <rest:response>
    <output:serialization-parameters>
        <output:media-type value='text/plain'/>
    </output:serialization-parameters>
    </rest:response>,
    'Not that simple anymore'
};
```

Error Handling

XQuery Errors

XQuery runtime errors can be processed via error annotations. Error annotations have one or more arguments, which represent the error codes to be caught. The codes equal the names of the XQuery 3.0 try/catch construct:

<table>
<thead>
<tr>
<th>Precedence</th>
<th>Syntax</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>prefix:name Q{uri}name</td>
<td>err:FORG0001 Q{<a href="http://www.w3.org/2005/xqt-errors%7DFORG0001">http://www.w3.org/2005/xqt-errors}FORG0001</a></td>
</tr>
<tr>
<td>2</td>
<td>prefix:* Q{uri}*</td>
<td>err:* Q{<a href="http://www.w3.org/2005/xqt-errors%7D">http://www.w3.org/2005/xqt-errors}</a>*</td>
</tr>
<tr>
<td>3</td>
<td>*:name</td>
<td>*:FORG0001</td>
</tr>
<tr>
<td>4</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

All error codes that are specified for a function must have the same precedence. The following rules apply when catching errors:

- Codes with a higher precedence (smaller number) will be given preference.
- A global RESTXQ error will be raised if two functions with conflicting codes are found.

Similar to try/catch, the pre-defined variables (code, description, value, module, line-number, column-number, additional) can be bound to variables via error parameter annotations, which are specified the same way as query parameters.

Errors may occur unexpectedly. However, they can also be triggered by a query, as demonstrated by the following example:

```xquery
declare %rest:path("/check/($user)")
    function page:check($user) {
        if($user = ('jack', 'lisa'))
            then 'User exists'
        else fn:error(xs:QName('err:user'), $user)
    };

declare %rest:error("err:user")
    %rest:error-param("description", "{$user}")
    function page:user-error($user)
An XQuery error in a RESTXQ context delivers by default a HTTP status code 400 error back to the client. However, you can also define a custom error code by using the third argument of the error function:

```
declare
  %rest:path("/teapot")
  function page:teapot()
  {
    fn:error(xs:QName('error'), "I'm a teapot", 418)
  }
```

### HTTP Errors

Errors that occur outside RESTXQ can be caught by adding `error-page` elements with an error code and a target location to the `web.xml` configuration file (find more details in the Jetty Documentation):

```xml
<error-page>
  <error-code>404</error-code>
  <location>/error404</location>
</error-page>
```

The target location may be another RESTXQ function. The `request:attribute` function can be used to request details on the caught error:

```
declare %rest:path("/error404") function page:error404() {
  "URL: " || request:attribute("javax.servlet.error.request_uri") || ", " || 
  "Error message: " || request:attribute("javax.servlet.error.message")
};
```

### Functions

The Request Module contains functions for accessing data related to the current HTTP request. Two modules exist for setting and retrieving server-side session data of the current user (Session Module) and all users known to the HTTP server (Sessions Module). The RESTXQ Module provides functions for requesting RESTXQ base URIs and generating a WADL description of all services. Please note that the namespaces of all of these modules must be explicitly specified via module imports in the query prolog.

The following example returns the current host name:

```
import module namespace request = "http://exquery.org/ns/request";

declare %rest:path("/host-name") function page:host() {
  'Remote host name: ' || request:remote-hostname()
};
```

### References

Currently, the following external resources on RESTXQ exist:

- RESTXQ Specification, First Draft
- RESTful XQuery, Standardised XQuery 3.0 Annotations for REST, Paper, XMLPrague, 2012
- RESTXQ, Slides, MarkLogic User Group London, 2012
- Web Application Development, Slides from XMLPrague 2013
Changelog

Version 8.4
• Added: %rest:single annotation

Version 8.1
• Added: support for input-specific content-type parameters
• Added: %input annotations

Version 8.0
• Added: Support for regular expressions in the Path Annotation
• Added: Evaluation of quality factors that are supplied in the Accept header

Version 7.9
• Updated: XQuery Errors, extended error annotations
• Added: %rest:method

Version 7.7
• Added: Error Handling, File Uploads, Multipart Types
• Updated: RESTXQ function may now also be specified in main modules (suffix: *.xq).
• Updated: the RESTXQ prefix has been changed from restxq to rest.
• Updated: parameters are implicitly cast to the type of the function argument
• Updated: the RESTXQ root url has been changed to http://localhost:8984/

Version 7.5
• Added: new XML elements <rest:redirect/> and <rest:forward/>
Chapter 81. REST

Read this entry online in the BaseX Wiki.

This page presents one of the Web Application services. It describes how to use the REST API of BaseX.

BaseX offers a RESTful API for accessing database resources via URLs. REST (REpresentational State Transfer) facilitates a simple and fast access to databases through HTTP. The HTTP methods GET, PUT, DELETE, and POST can be used to interact with the database.

Usage

By default, REST services are available at http://localhost:8984/rest/. If no default credentials are specified in the URL or when starting the web application, they will be requested by the client (see further).

A web browser can be used to perform simple GET-based REST requests and display the response. Some alternatives for using REST are listed in the Usage Examples.

URL Architecture

The root URL lists all available databases. The following examples assume that you have created a database instance from the factbook.xml document:

http://localhost:8984/rest

```xml
<rest:databases resources="1" xmlns:rest="http://basex.org/rest">
   <rest:database resources="1" size="1813599">factbook</rest:database>
</rest:databases>
```

The resources of a database can be listed by specifying the database, and potential sub directories, in the URL. In the given example, a single XML document is stored in the factbook database:

http://localhost:8984/rest/factbook

```xml
<rest:database name="factbook" resources="1" xmlns:rest="http://basex.org/rest">
   <rest:resource type="xml" content-type="application/xml" size="77192">factbook.xml</rest:resource>
</rest:databases>
```

The contents of a database can be retrieved by directly addressing the resource:

http://localhost:8984/rest/factbook/factbook.xml

If a resource is not found, an HTTP response will be generated with 404 as status code.

Parameters

The following parameters can be applied to the operations:

- **Variables**: External variables can be bound before a query is evaluated (see below for more).
- **Context**: The context parameter may be used to provide an initial XML context node.
- **Options**: Specified Options are applied before the actual operation will be performed.
- **Serialization**: All Serialization parameters known to BaseX can be specified as query parameters. Parameters that are specified within a query will be interpreted by the REST server before the output is generated.
While **Options** can be specified for all operations, the remaining parameters will only make sense for **Query** and **Run**.

**Request**

**GET Method**

If the GET method is used, all query parameters are directly specified within the URL. Additionally, the following operations can be specified:

- **query**: Evaluate an XQuery expression. If a database or database path is specified in the URL, it is used as initial query context.

- **command**: Execute a single database command.

- **run**: Evaluate an XQuery file or command script located on the server. The file path is resolved against the directory specified by **RESTPATH** (before, it was resolved against **WEBPATH**).

**Examples**

- Lists all resources found in the **tmp** path of the **factbook** database:
  
  http://localhost:8984/rest/factbook/tmp

- Serializes a document as JSONML:
  

- **US-ASCII** is chosen as output encoding, and the query `eval.xq` is evaluated:
  

- The next URL lists all database users that are known to BaseX:
  
  http://localhost:8984/rest?command=show+users

**POST Method**

The body of a POST request is interpreted as XML fragment, which specifies the operation to perform. The body must conform to a given XML Schema.

**Examples**

- The following query returns the first five city names of the **factbook** database:

  ```
  <query xmlns="http://basex.org/rest">
    <text><![CDATA[ (//city/name)[position() &lt;= 5] ]]></text>
  </query>
  ```

- The second query returns the string lengths of all text nodes, which are found in the node that has been specified as initial context node:

  ```
  <rest:query xmlns:rest="http://basex.org/rest">
    <rest:text>for $i in ./text() return string-length($i)</rest:text>
    <rest:context>
      <xml>
        <text>Hello</text>
        <text>World</text>
      </xml>
    </rest:context>
  </rest:query>
  ```

- The following request returns the registered database users encoded in **ISO-8859-1**:
<command xmlns="http://basex.org/rest">
  <text>show users</text>
  <parameter name='encoding' value='ISO-8859-1'/>
</command>

• This example creates a new database from the specified input and retains all whitespaces:

<command xmlns="http://basex.org/rest">
  <text>create db test http://files.basex.org/xml/xmark.xml</text>
  <option name='chop' value='false'/>
</command>

• The last request runs a query query.xq located in the directory specified by WEBPATH:

<run xmlns="http://basex.org/rest">
  <text>query.xq</text>
</run>

PUT Method

The PUT method is used to create new databases, or to add or update existing database resources:

• Create Database : A new database is created if the URL only specifies the name of a database. If the request body contains XML, a single document is created, adopting the name of the database.

• Store Resource : A resource is added to the database if the URL contains a database path. If the addressed resource already exists, it is replaced by the new input.

There are two ways to store non-XML data in BaseX:

• Store as Raw Data : If application/octet-stream is chosen as content-type, the input is added as Binary Data.

• Convert to XML : Incoming data is converted to XML if a parser is available for the specified content-type. The following content types are supported:
  • application/json: Stores JSON as XML.
  • text/plain: Stores plain text input as XML.
  • text/comma-separated-values: Stores CSV text input as XML.
  • text/html: Stores HTML input as XML.

Conversion can be influenced by specifying additional content-type parameters (see RESTXQ for more information).

If raw data is added and if no content type, or a wrong content, is specified, a 400 (BAD REQUEST) error will be raised.

Examples

• A new database with the name XMark is created. If XML input is sent in the HTTP body, the resulting database resource will be called XMark.xml: http://localhost:8984/rest/XMark

• A new database is created, and no whitespaces will be removed from the passed on XML input: http://localhost:8984/rest/XMark?chop=false

• The contents of the HTTP body will be taken as input for the document one.xml, which will be stored in the XMark database: http://localhost:8984/rest/XMark/one.xml
An HTTP response with status code 201 (CREATED) is sent back if the operation was successful. Otherwise, the server will reply with 404 (if a specified database was not found) or 400 (if the operation could not be completed).

Have a look at the usage examples for more detailed examples using Java and shell tools like cURL.

**DELETE Method**

The DELETE method is used to delete databases or resources within a database.

Example

- The factbook database is deleted: http://localhost:8984/rest/factbook
- All resources of the XMark database are deleted that reside in the tmp path: http://localhost:8984/rest/XMark/tmp/

The HTTP status code 404 is returned if no database is specified. 200 (OK) will be sent in all other cases.

**Assigning Variables**

**GET Method**

All query parameters that have not been processed before will be treated as variable assignments:

Example

- The following request assigns two variables to a server-side query file mult.xq placed in the HTTP directory: http://localhost:8984/rest?run=mult.xq&$a=21&$b=2

```
(: XQuery file: mult.xq :)
declare variable $a as xs:integer external;
declare variable $b as xs:integer external;
<mult>{ $a * $b }</mult>
```

The dollar sign can be omitted as long as the variable name does not equal a parameter keyword (e.g.: method).

**POST Method**

If query or run is used as operation, external variables can be specified via the `<variable/>` element:

```
<query xmlns="http://basex.org/rest">
  <text><![CDATA[
    declare variable $x as xs:integer external;
declare variable $y as xs:integer external;
    <mult>{ $x * $y }</mult>
  ]]></text>
  <variable name="a" value="21"/>
  <variable name="b" value="2"/>
</query>
```

**Response**

**Content Type**

As the content type of a REST response cannot necessarily be dynamically determined, it can be enforced by the user. The final content type of a REST response is chosen as follows:

1. If the serialization parameter `media-type` is supplied, it will be adopted as content-type.
2. Otherwise, if the serialization parameter method is supplied, the content-type will be chosen according to the following mapping:

- xml, adaptive, basex → application/xml
- xhtml → text/html
- html → text/html
- text → text/plain
- json → application/json

3. If no media-type or serialization method is supplied, the content type of a response depends on the chosen REST operation:

- Query / Run → application/xml
- Command → text/plain
- Get → application/xml, or content type of the addressed resource

Serialization parameters can either be supplied as query parameters or within the query.

The following three example requests all return `<a/>` with application/xml as content-type:

```plaintext
http://localhost:8984/rest?query=%3Ca/%3E
http://localhost:8984/rest?query=%3Ca/%3E&method=xml
http://localhost:8984/rest?query=%3Ca/%3E&media-type=application/xml
```

### Usage Examples

#### Java

#### Authentication

Most programming languages offer libraries to communicate with HTTP servers. The following example demonstrates how easy it is to perform a DELETE request with Java.

Basic access authentication can be activated in Java by adding an authorization header to the HttpURLConnection instance. The header contains the word Basic, which specifies the authentication method, followed by the Base64-encoded USER:PASSWORD pair. As Java does not include a default conversion library for Base64 data, the internal BaseX class org.basex.util.Base64 can be used for that purpose:

```java
import java.net.*;
import org.basex.util.*;

public final class RESTExample {
    public static void main(String[] args) throws Exception {
        // The java URL connection to the resource.
        URL url = new URL("http://localhost:8984/rest/factbook");

        // Establish the connection to the URL.
        HttpURLConnection conn = (HttpURLConnection) url.openConnection();
        // Set as DELETE request.
        conn.setRequestMethod("DELETE");

        // User and password.
        String user = "bob";
        String pw = "alice";
        // Encode user name and password pair with a base64 implementation.
```
String encoded = Base64.encode(user + "":" + pw);
  // Basic access authentication header to connection request.
  conn.setRequestProperty("Authorization", "Basic " + encoded);

  // Print the HTTP response code.
  System.out.println("HTTP response: " + conn.getResponseCode());

  // Close connection.
  conn.disconnect();
}

// store input as raw
conn.setRequestProperty("Content-Type", "application/octet-stream");

See the PUT Requests section for a description of the possible content-types.

Find Java examples for all methods here: GET, POST, PUT, DELETE.

Command Line

Tools such as the Linux commands Wget or cURL exist to perform HTTP requests (try copy & paste):

GET
  • curl -i "http://localhost:8984/rest/factbook?query=//city/name"

POST
  • curl -i -X POST -H "Content-Type: application/xml" -d "<query xmlns='http://basex.org/rest'><text>//city/name</text></query>" "http://localhost:8984/rest/factbook"
  • curl -i -X POST -H "Content-Type: application/xml" -T query.xml "http://localhost:8984/rest/factbook"

PUT
  • curl -i -X PUT -T "etc/xml/factbook.xml" "http://localhost:8984/rest/factbook"
  • curl -i -X PUT -H "Content-Type: application/json" -T "plain.json" "http://localhost:8984/rest/plain"

DELETE
  • curl -i -X DELETE "http://admin:admin@localhost:8984/rest/factbook"

Changelog

Version 8.1
  • Added: support for input-specific content-type parameters
  • Updated: the run operation now resolves file paths against the RESTPATH option.

Version 8.0
• Removed: \textit{wrap} parameter

Version 7.9

• Updated: Also evaluate command scripts via the \textit{run} operation.

Version 7.2

• Removed: direct evaluation of addresses resources with \texttt{application/xquery} as content type

Version 7.1.1

• Added: \texttt{options} parameter for specifying database options

Version 7.1

• Added: PUT request: automatic conversion to XML if known content type is specified

Version 7.0

• REST API introduced, replacing the old JAX-RX API
Chapter 82. REST: POST Schema

Read this entry online in the BaseX Wiki.

The following schema is used from the REST API to validate POST requests:

```xml
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
  xmlns="http://basex.org/rest"
  targetNamespace="http://basex.org/rest">
  <xs:element name="query">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="text" minOccurs="1" maxOccurs="1"/>
        <xs:element ref="parameter" minOccurs="0" maxOccurs="unbounded"/>
        <xs:element ref="option" minOccurs="0" maxOccurs="unbounded"/>
        <xs:element ref="variable" minOccurs="0" maxOccurs="unbounded"/>
        <xs:element ref="context" minOccurs="0" maxOccurs="1"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
  <xs:element name="run">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="text" minOccurs="1" maxOccurs="1"/>
        <xs:element ref="parameter" minOccurs="0" maxOccurs="unbounded"/>
        <xs:element ref="option" minOccurs="0" maxOccurs="unbounded"/>
        <xs:element ref="variable" minOccurs="0" maxOccurs="unbounded"/>
        <xs:element ref="context" minOccurs="0" maxOccurs="1"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
  <xs:element name="command">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="text" minOccurs="1" maxOccurs="1"/>
        <xs:element ref="parameter" minOccurs="0" maxOccurs="unbounded"/>
        <xs:element ref="option" minOccurs="0" maxOccurs="unbounded"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
  <xs:element name="text" type="xs:string"/>
  <xs:element name="option">
    <xs:complexType>
      <xs:attribute name="name" type="xs:string" use="required"/>
      <xs:attribute name="value" type="xs:string" use="required"/>
    </xs:complexType>
  </xs:element>
  <xs:element name="parameter">
    <xs:complexType>
      <xs:attribute name="name" type="xs:string" use="required"/>
      <xs:attribute name="value" type="xs:string" use="required"/>
    </xs:complexType>
  </xs:element>
  <xs:element name="variable">
    <xs:complexType>
      <xs:attribute name="name" type="xs:string" use="required"/>
      <xs:attribute name="value" type="xs:string" use="required"/>
      <xs:attribute name="type" type="xs:string" use="optional"/>
    </xs:complexType>
  </xs:element>
  <xs:element name="context" type="xs:anyType"/>
</xs:schema>
```
</xs:schema>
Chapter 83. WebDAV

Read this entry online in the BaseX Wiki.

This page presents one of the Web Application services. It describes how to use the WebDAV file system interface.

BaseX offers access to the databases and documents using the WebDAV protocol. WebDAV provides convenient means to access and edit XML documents by representing BaseX databases and documents in the form of a file system hierarchy.

Usage

By default, the BaseX HTTP server makes the WebDAV service accessible at http://localhost:8984/webdav/. If no default credentials are specified, they will be requested by the client (see further). It can be accessed by either http://<httphost>:<httpport>/webdav/ or webdav://<httphost>:<httpport>/webdav/, depending on your WebDAV client.

Please note that the file size of XML documents will be displayed as 0 bytes, as the actual file size can only be determined if the full document is being returned and serialized. This may cause problems with some WebDAV clients (e.g. NetDrive or WebDrive).

Authorization

The WebDAV service uses the database user credentials in order to perform authentication and authorization. If database user and password are explicitly specified when starting the BaseX HTTP Server using the corresponding startup options, WebDAV will not request additional user authentication from the client.

Root Directory

In the WebDAV root directory, all existing databases are listed. As new resources can only be stored within database, it is not possible to store files in the root directory. If a file is copied on top level, a new database will be created, which contains this resource.

Locking

The BaseX WebDAV implementation supports locking. It can be utilized with clients which support this feature (e.g. oXygen Editor). EXCLUSIVE and SHARED locks are supported, as well as WRITE locks.

Note: WebDAV locks are stored in a database called ~webdav. If the database is deleted, it will automatically be recreated along with the next lock operations. If a resource remains locked, it can be unlocking by removing the correspondent <w:lockinfo> entry.

WebDAV Clients

Please check out the following tutorials to get WebDAV running on different operating systems and with oXygen:

- Windows 7 and up
- Windows XP
- Mac OSX 10.4+
- GNOME and Nautilus
- KDE
- oXygen Editor
Changelog

Version 7.7

• Added: Locking

Version 7.0

• WebDAV API introduced
Chapter 84. WebDAV: Windows 7

Read this entry online in the BaseX Wiki.

This page belongs to the WebDAV page. It describes how to get the WebDAV API running with Windows 7.

• Open the Explorer

• Open the "Map network drive..." dialog by right-clicking on "My Computer"

• Click on the link "Connect to a Web site that you can use to store your documents and pictures."

```
What network folder would you like to map?

Specify the drive letter for the connection and the folder that you want to connect to:

Drive: Z:  
Folder:  

Example: \\server\share  
☑ Reconnect at logon  
☐ Connect using different credentials

Connect to a Web site that you can use to store your documents and pictures.

Finish  Cancel
```

• Click "Next", select "Choose a custom network location" and click "Next" again.

```
Where do you want to create this network location?

Choose a custom network location

Specify the address of a website, network location, or FTP site.

Next  Cancel
```

• Enter the URL address of the BaseX WebDAV Server (e.g. http://localhost:8984/webdav) and click "Next".

359
If a message saying that the folder is not valid, this is because Microsoft WebClient is not configured to use Basic HTTP authentication. Please check this Microsoft article in order to enable Basic HTTP authentication.

• Enter a name for the network location and click "Next".

• The BaseX WebDAV can be accessed from the Explorer window.
Chapter 85. WebDAV: Windows XP

Read this entry online in the BaseX Wiki.

This page belongs to the WebDAV page. It describes how to get the WebDAV API running with Windows XP.

• In the "My Network Places" view, double click on "Add Network Place":

![Image of Add Network Place dialog]

• Confirm the upcoming introductory dialog:

![Image of Welcome to the Add Network Place Wizard]

• Select "Choose another network location" in the next dialog:
• Next, specify the BaseX WebDAV URL:

• Enter the user/password combination to connect to the WebDAV service:
• Assign a name to your WebDAV connection:

• Finish the wizard:

• You can now see all BaseX databases in the Windows Explorer:
Chapter 86. WebDAV: Mac OSX

Read this entry online in the BaseX Wiki.

This page belongs to the WebDAV page. It describes how to get the WebDAV API running with Mac OS X 10.4+.

• Mac OS X supports WebDAV since 10.4/Tiger

• Open Finder, choose Go -> Connect to Server:

• Enter BaseX WebDAV URL (eg. http://localhost:8984/webdav) - do not use webdav://-scheme! Press Connect:
WebDAV: Mac OSX

- Enter the user credentials:

- That's it, now the databases can be browsed:
WebDAV: Mac OSX
Chapter 87. WebDAV: GNOME

Read this entry online in the BaseX Wiki.

This page belongs to the WebDAV page. It describes how to get the WebDAV API running with GNOME and Nautilus.

- In Nautilus choose File -> Connect to Server:

- Choose "WebDAV (HTTP)" from the "Type" drop-down and enter the server address, port and user credentials:
• After clicking "Connect" the databases can be browsed:
Chapter 88. WebDAV: KDE

Read this entry online in the BaseX Wiki.

This page belongs to the WebDAV page. It describes how to get the WebDAV API running with KDE.

• KDE SC provides two file managers - Dolphin and Konqueror, which both support WebDAV using the "webdav://" URL prefix. Start Dolphin or Konqueror and enter the BaseX WebDAV URL (eg. webdav://localhost:8984/webdav):

• Enter the user credentials:

• After clicking "OK" the databases can be browsed:
Part IX. Client APIs
Chapter 89. Clients

Read this entry online in the BaseX Wiki.

This page is part of the Developer Section. It describes how to communicate with BaseX from other programming languages.

You can use the following light-weight language bindings to connect to a running BaseX server instance, execute database commands and evaluate XQuery expressions.

Most clients provide two modes:

• **Standard Mode**: connecting to a server, sending commands
• **Query Mode**: defining queries, binding variables, iterative evaluation

Currently, we offer bindings for the following programming languages:

### BaseX 7.x, BaseX 8.x and later

- **Java**: The default implementation
- **C++**: contributed by Jean-Marc Mercier
- **C#**, contributed by the BaseX Team and Martin Ferrari
- **node.js**: contributed by Andy Bunce
- **Perl**, contributed by the BaseX Team
- **PHP**: updated by James Ball
- **Python 3.x, 2.7.3**: contributed by Hiroaki Itoh
- **Python**, using BaseX REST services: contributed by Luca Lianas
- **Ruby**, contributed by the BaseX Team

With **Version 8.0**, authentication has changed. Some of the language bindings have not been updated yet. The update is rather trivial, though (see here for more details); we are looking forward to your patches!

### BaseX 7.x (outdated)

- **ActionScript**: contributed by Manfred Knobloch
- **C**: contributed by the BaseX Team
- **Golang**: contributed by Christian Baune
- **Haskell**: contributed by Leo Wörteler
- **Lisp**: contributed by Andy Chambers
- **node.js**: contributed by Hans Hübner (deviating from client API)
- **Python < 2.7**: improved by Arjen van Elteren
- **Qt**: contributed by Hendrik Strobelt
- **Rebol**: contributed by Sabu Francis
- **Scala**: contributed by Manuel Bernhardt
- **Scala** (simple implementation)
- **VB**: contributed by the BaseX Team

Many of the interfaces contain the following files:

- **BaseXClient** contains the code for creating a session, sending and executing commands and receiving results. An inner **Query** class facilitates the binding of external variables and iterative query evaluation.
- **Example** demonstrates how to send database commands.
- **QueryExample** shows you how to evaluate queries in an iterative manner.
- **QueryBindExample** shows you how to bind a variable to your query and evaluates the query in an iterative manner.
- **CreateExample** shows how new databases can be created by using streams.
- **AddExample** shows how documents can be added to a database by using streams.
Changelog

Version 8.0

• Updated: cram-md5 replaced with digest authentication
Chapter 90. Standard Mode

Read this entry online in the BaseX Wiki.

In the standard mode of the Clients, a database command can be sent to the server using the `execute()` function of the `Session`. This function returns the whole result. With the `info()` function, you can request some information on your executed process. If an error occurs, an exception with the error message will be thrown.

Usage

The standard execution works as follows:

1. Create a new session instance with hostname, port, username and password.
2. Call the `execute()` function of the session with the database commands as argument.
3. Receive the result of a successfully executed command. If an error occurs, an exception is thrown.
4. Optionally, call `info()` to get some process information
5. Continue using the client (back to 2.), or close the session.

Example in PHP

Taken from our repository:

```php
<?php
/*
 * This example shows how database commands can be executed.
 * Documentation: http://baseX.org/apl
 * (C) BaseX Team 2005-15, BSD License
 */
include("BaseXClient.php");

try {
    // initialize timer
    $start = microtime(true);

    // create session
    $session = new Session("localhost", 1984, "admin", "admin");

    // perform command and print returned string
    print $session->execute("xquery 1 to 10");

    // close session
    $session->close();

    // print time needed
    $time = (microtime(true) - $start) * 1000;
    print "$time ms
";
} catch (Exception $e) {
    // print exception
    print $e->getMessage();
}
?>
```
Chapter 91. Query Mode

Read this entry online in the BaseX Wiki.

The query mode of the Clients allows you to bind external variables to a query and evaluate the query in an iterative manner. The query() function of the Session instance returns a new query instance.

Usage

The query execution works as follows:

1. Create a new session instance with hostname, port, username and password.
2. Call query() with your XQuery expression to get a query object.
3. Optionally bind variables to the query with one of the bind() functions.
4. Optionally bind a value to the context item via context().
5. Iterate through the query object with the more() and next() functions.
6. As an alternative, call execute() to get the whole result at a time.
7. info() gives you information on query evaluation.
8. options() returns the query serialization parameters.
9. Don’t forget to close the query with close().

PHP Example

Taken from our repository:

```php
<?php
/*
 * This example shows how queries can be executed in an iterative manner.
 * Documentation: http://basex.org/api
 * (C) BaseX Team 2005-15, BSD License
 */
include("BaseXClient.php");

try {
    // create session
    $session = new Session("localhost", 1984, "admin", "admin");

    try {
        // create query instance
        $input = 'declare variable $name external; '.
            'for $i in 1 to 10 return element { $name } { $i }';
        $query = $session->query($input);

        // bind variable
        $query->bind("$name", "number");

        // print result
        print $query->execute()."\n";

        // close query instance
        $query->close();
    }
}
```
Changelog

Version 7.2

- Added: context() function
Chapter 92. Server Protocol

Read this entry online in the BaseX Wiki.

This page presents the classes and functions of the BaseX Clients, and the underlying protocol, which is utilized for communicating with the database server. A detailed example demonstrates how a concrete byte exchange can look like.

Workflow

- All clients are based on the client/server architecture. Hence, a BaseX database server must be started first.
- Each client provides a session class or script with methods to connect to and communicate with the database server. A socket connection will be established by the constructor, which expects a host, port, user name and password as arguments.
- The `execute()` method is called to launch a database command. It returns the result or throws an exception with the received error message.
- The `query()` method creates a query instance. Variables and the context item can be bound to that instance, and the result can either be requested via `execute()`, or in an iterative manner with the `more()` and `next()` functions. If an error occurs, an exception will be thrown.
- The `create()`, `add()`, `replace()` and `store()` method pass on input streams to the corresponding database commands.
- To speed up execution, an output stream can be specified by some clients; this way, all results will be directed to that output stream.
- Most clients are accompanied by some example files, which demonstrate how database commands can be executed or how queries can be evaluated.

Transfer Protocol

All Clients use the following client/server protocol to communicate with the server. The description of the protocol is helpful if you want to implement your own client.

Conventions

- \( xx \): single byte.
- \{ ... \}: utf8 strings or raw data, suffixed with a \00 byte. To avoid confusion with this end-of-string byte, all transferred \00 and \FF bytes are prefixed by an additional \FF byte.

Authentication

Digest

Digest authentication is used since Version 8.0:

1. Client connects to server socket
2. Server sends a realm and nonce, separated by a colon: \{realm:nonce\}
3. Client sends the user name and a hash value. The hash is composed of the md5 hash of
   a. the md5 hash of the user name, realm, and password (all separated by a colon), and
   b. the nonce: \{username\} \{md5(md5(username:realm:password) + nonce)\}
4. Server replies with \00 (success) or \01 (error)
CRAM-MD5

CRAM-MD5 was discarded, because unsalted md5 hashes could easily be uncovered using rainbow tables. However, most client bindings still provide support for the outdated handshaking, as it only slightly differs from the new protocol:

1. Client connects to server socket
2. Server sends a nonce (timestamp): \{nonce\}
3. Client sends the user name and a hash value. The hash is composed of the md5 hash of
   a. the md5 of the password and
   b. the nonce: \{username\} \{md5(md5(password) + nonce)\}
4. Server replies with \00 (success) or \01 (error)

Clients can easily be implemented to both support digest and cram-md5 authentication: If the first server response contains no colon, cram-md5 should be chosen.

Command Protocol

The following byte sequences are sent and received from the client (please note that a specific client may not support all of the presented commands):

<table>
<thead>
<tr>
<th>Command</th>
<th>Client Request</th>
<th>Server Response</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>COMMAND</td>
<td>{command}</td>
<td>{result} {info} \00</td>
<td>Executes a database command.</td>
</tr>
<tr>
<td>QUERY</td>
<td>\00 {query}</td>
<td>{id} \00</td>
<td>Creates a new query instance and returns its id.</td>
</tr>
<tr>
<td>CREATE</td>
<td>\08 {name} {input} {info} \00</td>
<td>Creates a new database with the specified input (may be empty).</td>
<td></td>
</tr>
<tr>
<td>ADD</td>
<td>\09 {name} {path} {input} {info} \00</td>
<td>Adds a new resource to the opened database.</td>
<td></td>
</tr>
<tr>
<td>REPLACE</td>
<td>\0C {path} {input} {info} \00</td>
<td>Replaces a resource with the specified input.</td>
<td></td>
</tr>
<tr>
<td>STORE</td>
<td>\0D {path} {input} {info} \00</td>
<td>Stores a binary resource in the opened database.</td>
<td></td>
</tr>
<tr>
<td># error</td>
<td></td>
<td>{ partial result } {error} \01</td>
<td>Error feedback.</td>
</tr>
</tbody>
</table>

Query Command Protocol

Queries are referenced via an id, which has been returned by the QUERY command (see above).

<table>
<thead>
<tr>
<th>Query Command</th>
<th>Client Request</th>
<th>Server Response</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLOSE</td>
<td>\02 {id}</td>
<td>\00 \00</td>
<td>Closes and unregisters the query with the specified id.</td>
</tr>
<tr>
<td>BIND</td>
<td>\03 {id} {name} \00 \00 {value} {type}</td>
<td>Binds a value to a variable. The type will be ignored if the string is empty.</td>
<td></td>
</tr>
</tbody>
</table>
| RESULTS       | \04 \{id\} \xx \{item\} ... \xx \{item\} \00 | Returns all resulting items as strings, prefixed by a single byte (\xx) that
**Server Protocol**

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXECUTE</td>
<td>\05 {id} {result} \00 Represents the Type ID. This command is called by the more() function of a client implementation. Executes the query and returns the result as a single string.</td>
</tr>
<tr>
<td>INFO</td>
<td>\06 {id} {result} \00 Returns a string with query compilation and profiling info.</td>
</tr>
<tr>
<td>OPTIONS</td>
<td>\07 {id} {result} \00 Returns a string with all query serialization parameters, which can e.g. be assigned to the SERIALIZER option.</td>
</tr>
<tr>
<td>CONTEXT</td>
<td>\0E {id} {value} \00 \00 {type} Binds a value to the context. The type will be ignored if the string is empty.</td>
</tr>
<tr>
<td>UPDATING</td>
<td>\1E {id} {result} \00 Returns true if the query contains updating expressions; false otherwise.</td>
</tr>
<tr>
<td>FULL</td>
<td>\1F {id} XDM {item} ... XDM {item} \00 Returns all resulting items as strings, prefixed by the XDM Meta Data. This command is e.g. used by the XQJ API.</td>
</tr>
</tbody>
</table>

As can be seen in the table, all results end with a single \00 byte, which indicates that the process was successful. If an error occurs, an additional byte \01 is sent, which is then followed by the error message string.

**Binding Sequences**

Also sequences can be bound to variables and the context:

- `empty-sequence()` must be supplied as type if an empty sequence is to be bound.
- Multiple items are supplied via the \{value\} argument and separated with \01 bytes.
- Item types are specified by appending \02 and the type in its string representation to an item. If no item type is specified, the general type is used.

Some examples for the \{value\} argument:

- the two integers 123 and 789 are encoded as 123, \01, 789 and \00(xs:integer may be specified via the \{type\} argument).
- the two items `xs:integer(123)` and `xs:string('ABC')` are encoded as 123, \02, `xs:integer. \01, ABC, \02, xs:string\ and \00.`

**Example**

In the following example, a client registers a new session and executes the INFO database command. Next, it creates a new query instance for the XQuery expression 1, 2+’3’. The query is then evaluated, and the server returns the result of the first subexpression 1 and an error for the second sub expression. Finally, the query instance and client session are closed.
Server Protocol

- **Client** connects to the database server socket

- **Server** sends realm and timestamp "BaseX:1369578179679": # 42 61 73 65 58 3A 31 33 36 39 35 37 38 31 37 39 36 37 39 00

- **Client** sends user name "jack": 6A 61 63 6B 00 #

- **Client** sends hash: md5(md5("jack:BaseX:topsecret") + "1369578179679") = "ca664a31f8deda9b71ea3e79347f6666": 63 61 36 ... 00 #

- **Server** replies with success code: # 00

- **Client** sends the "INFO" command: 49 4E 46 4F 00 #

- **Server** responds with the result "General Information...": # 47 65 6e 65 ... 00

- **Server** additionally sends an (empty) info string: # 00

- **Client** creates a new query instance for the XQuery "1, 2+'3'": 00 31 2C 20 32 2B 27 33 27 00 #

- **Server** returns query id "1" and a success code: # 31 00 00

- **Client** requests the query results via the RESULTS protocol command and its query id: 04 31 00 #

- **Server** sends the first result ("1", type xs:integer): # 52 31 00

- **Server** sends a single \00 byte instead of a new result, which indicates that no more results can be expected: # 00

- **Server** sends the error code \01 and the error message ("Stopped at..."): # 01 53 74 6f ... 00

- **Client** closes the query instance: 02 31 00 #

- **Server** sends a response (which is equal to an empty info string) and success code: # 00 00

- **Client** closes the socket connection

Constructors and Functions

Most language bindings provide the following constructors and functions:

**Session**

- Create and return session with host, port, user name and password: `Session(String host, int port, String name, String password)`

- Execute a command and return the result: `String execute(String command)`

- Return a query instance for the specified query: `Query query(String query)`

- Create a database from an input stream: `void create(String name, InputStream input)`

- Add a document to the current database from an input stream: `void add(String path, InputStream input)`

- Replace a document with the specified input stream: `void replace(String path, InputStream input)`

- Store raw data at the specified path: `void store(String path, InputStream input)`

- Return process information: `String info()`

- Close the session: `void close()`
Query

- Create query instance with session and query: `Query(Session session, String query)`
- Bind an external variable: `void bind(String name, String value, String type)` The type can be an empty string.
- Bind the context item: `void context(String value, String type)` The type can be an empty string.
- Execute the query and return the result: `String execute()`
- Iterator: check if a query returns more items: `boolean more()`
- Iterator: return the next item: `String next()`
- Return query information: `String info()`
- Return serialization parameters: `String options()`
- Return if the query may perform updates: `boolean updating()`
- Close the query: `void close()`

Changelog

Version 8.2
- Removed: `WATCH` and `UNWATCH` command

Version 8.0
- Updated: cram-md5 replaced with digest authentication
- Updated: `BIND` command: support more than one item

Version 7.2
- Added: Query Commands `CONTEXT`, `UPDATING` and `FULL`
- Added: Client function `context(String value, String type)`
Chapter 93. Server Protocol: Types

This article lists extended type information that is returned by the Server Protocol.

**XDM Meta Data**

In most cases, the XDM meta data is nothing else than the Type ID. There are three exceptions: document-node(), attribute() and xs:QName items are followed by an additional \{URI\} string.

**Type IDs**

The following table lists the type IDs that are returned by the server. Currently, all node kinds are of type `xs:untypedAtomic`:

<table>
<thead>
<tr>
<th>Type ID</th>
<th>Node Kind/Item Type</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>Function item</td>
<td>function</td>
</tr>
<tr>
<td>8</td>
<td>node()</td>
<td>node</td>
</tr>
<tr>
<td>9</td>
<td>text()</td>
<td>node</td>
</tr>
<tr>
<td>10</td>
<td>processing-instruction()</td>
<td>node</td>
</tr>
<tr>
<td>11</td>
<td>element()</td>
<td>node</td>
</tr>
<tr>
<td>12</td>
<td>document-node()</td>
<td>node</td>
</tr>
<tr>
<td>13</td>
<td>document-node(element())</td>
<td>node</td>
</tr>
<tr>
<td>14</td>
<td>attribute()</td>
<td>node</td>
</tr>
<tr>
<td>15</td>
<td>comment()</td>
<td>node</td>
</tr>
<tr>
<td>32</td>
<td>item()</td>
<td>atomic value</td>
</tr>
<tr>
<td>33</td>
<td><code>xs:untyped</code></td>
<td>atomic value</td>
</tr>
<tr>
<td>34</td>
<td><code>xs:anyType</code></td>
<td>atomic value</td>
</tr>
<tr>
<td>35</td>
<td><code>xs:anySimpleType</code></td>
<td>atomic value</td>
</tr>
<tr>
<td>36</td>
<td><code>xs:anyAtomicType</code></td>
<td>atomic value</td>
</tr>
<tr>
<td>37</td>
<td><code>xs:untypedAtomic</code></td>
<td>atomic value</td>
</tr>
<tr>
<td>38</td>
<td><code>xs:string</code></td>
<td>atomic value</td>
</tr>
<tr>
<td>39</td>
<td><code>xs:normalizedString</code></td>
<td>atomic value</td>
</tr>
<tr>
<td>40</td>
<td><code>xs:token</code></td>
<td>atomic value</td>
</tr>
<tr>
<td>41</td>
<td><code>xs:language</code></td>
<td>atomic value</td>
</tr>
<tr>
<td>42</td>
<td><code>xs:NM_TOKEN</code></td>
<td>atomic value</td>
</tr>
<tr>
<td>43</td>
<td><code>xs:Name</code></td>
<td>atomic value</td>
</tr>
<tr>
<td>44</td>
<td><code>xs:NCName</code></td>
<td>atomic value</td>
</tr>
<tr>
<td>45</td>
<td><code>xs:ID</code></td>
<td>atomic value</td>
</tr>
<tr>
<td>46</td>
<td><code>xs:IDREF</code></td>
<td>atomic value</td>
</tr>
<tr>
<td>47</td>
<td><code>xs:ENTITY</code></td>
<td>atomic value</td>
</tr>
<tr>
<td>48</td>
<td><code>xs:float</code></td>
<td>atomic value</td>
</tr>
<tr>
<td>49</td>
<td><code>xs:double</code></td>
<td>atomic value</td>
</tr>
<tr>
<td>50</td>
<td><code>xs:decimal</code></td>
<td>atomic value</td>
</tr>
<tr>
<td>51</td>
<td><code>xs:precisionDecimal</code></td>
<td>atomic value</td>
</tr>
<tr>
<td>No.</td>
<td>Type Name</td>
<td>Description</td>
</tr>
<tr>
<td>-----</td>
<td>-----------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>52</td>
<td>xs:integer</td>
<td>atomic value</td>
</tr>
<tr>
<td>53</td>
<td>xs:nonPositiveInteger</td>
<td>atomic value</td>
</tr>
<tr>
<td>54</td>
<td>xs:negativeInteger</td>
<td>atomic value</td>
</tr>
<tr>
<td>55</td>
<td>xs:long</td>
<td>atomic value</td>
</tr>
<tr>
<td>56</td>
<td>xs:int</td>
<td>atomic value</td>
</tr>
<tr>
<td>57</td>
<td>xs:short</td>
<td>atomic value</td>
</tr>
<tr>
<td>58</td>
<td>xs:byte</td>
<td>atomic value</td>
</tr>
<tr>
<td>59</td>
<td>xs:nonNegativeInteger</td>
<td>atomic value</td>
</tr>
<tr>
<td>60</td>
<td>xs:unsignedLong</td>
<td>atomic value</td>
</tr>
<tr>
<td>61</td>
<td>xs:unsignedInt</td>
<td>atomic value</td>
</tr>
<tr>
<td>62</td>
<td>xs:unsignedShort</td>
<td>atomic value</td>
</tr>
<tr>
<td>63</td>
<td>xs:unsignedByte</td>
<td>atomic value</td>
</tr>
<tr>
<td>64</td>
<td>xs:positiveInteger</td>
<td>atomic value</td>
</tr>
<tr>
<td>65</td>
<td>xs:duration</td>
<td>atomic value</td>
</tr>
<tr>
<td>66</td>
<td>xs:yearMonthDuration</td>
<td>atomic value</td>
</tr>
<tr>
<td>67</td>
<td>xs:dayTimeDuration</td>
<td>atomic value</td>
</tr>
<tr>
<td>68</td>
<td>xs:dateTime</td>
<td>atomic value</td>
</tr>
<tr>
<td>69</td>
<td>xs:dateTimeStamp</td>
<td>atomic value</td>
</tr>
<tr>
<td>70</td>
<td>xs:date</td>
<td>atomic value</td>
</tr>
<tr>
<td>71</td>
<td>xs:time</td>
<td>atomic value</td>
</tr>
<tr>
<td>72</td>
<td>xs:gYearMonth</td>
<td>atomic value</td>
</tr>
<tr>
<td>73</td>
<td>xs:gYear</td>
<td>atomic value</td>
</tr>
<tr>
<td>74</td>
<td>xs:gMonthDay</td>
<td>atomic value</td>
</tr>
<tr>
<td>75</td>
<td>xs:gDay</td>
<td>atomic value</td>
</tr>
<tr>
<td>76</td>
<td>xs:gMonth</td>
<td>atomic value</td>
</tr>
<tr>
<td>77</td>
<td>xs:boolean</td>
<td>atomic value</td>
</tr>
<tr>
<td>78</td>
<td>base6:binary</td>
<td>atomic value</td>
</tr>
<tr>
<td>79</td>
<td>xs:base64Binary</td>
<td>atomic value</td>
</tr>
<tr>
<td>80</td>
<td>xs:hexBinary</td>
<td>atomic value</td>
</tr>
<tr>
<td>81</td>
<td>xs:anyURI</td>
<td>atomic value</td>
</tr>
<tr>
<td>82</td>
<td>xs:QName</td>
<td>atomic value</td>
</tr>
<tr>
<td>83</td>
<td>xs:NOTATION</td>
<td>atomic value</td>
</tr>
</tbody>
</table>
Chapter 94. Java Examples

Read this entry online in the BaseX Wiki.

This page is part of the Developer Section. The following Java code snippets demonstrate how easy it is to run database commands, create collections, perform queries, etc. by integrating the BaseX code. Most examples are taken from our basex-examples repository, in which you will find some more use cases.

Local Examples

The following code snippets work in embedded mode; they do not rely on an additional server instance:

• RunCommands.java creates and drops database and index instances, prints a list of all existing databases.
• RunQueries.java shows three variants of running queries.
• BindContext.java demonstrates how a value can be bound as context item.
• BindVariables.java demonstrates how a value can be bound to a variable.
• CreateCollection.java creates and manages a collection.
• QueryCollection.java creates, runs queries against it and drops a collection.
• WikiExample.java creates a database from an url (wiki instance), runs a query against it and drops the database.

Server Examples

The examples below take advantage of the client/server architecture:

• ServerCommands.java launches server-side commands using a client session.
• ServerAndLocal.java processes server results locally.
• ServerConcurrency.java runs concurrent queries.
• ServerQueries.java shows how iterative queries can be performed.
• UserExample.java manages database users.

XQuery Module Examples

BaseX provides Java Bindings for accessing external Java code via XQuery functions. The following examples show how this feature can be utilized:

• FruitsExample.java demonstrates how Java classes can be imported as XQuery modules.
• FruitsModule.java is a simple demo module called by FruitsExample.
• ModuleDemo.java is a simple XQuery demo module that demonstrates how XQuery items can be processed from Java. It is derived from the QueryModule class.
• QueryModule.java is located in the BaseX core. Java query modules can extend this class to get access to the current query context and enrich functions with properties ()

XQJ API

The implementation of the BaseX XQJ API (closed-source) has been written by Charles Foster. It uses the client/server architecture. The basex-examples repository contains various examples on how to use XQJ.
Client API

- BaseXClient.java provides an implementation of the Server Protocol.
- Example.java demonstrates how commands can be executed on a server.
- QueryExample.java shows how queries can be executed in an iterative manner.
- QueryBindExample.java shows how external variables can be bound to XQuery expressions.
- CreateExample.java shows how new databases can be created.
- AddExample.java shows how documents can be added to databases, and how existing documents can be replaced.
- BinaryExample.java shows how binary resource can be added to and retrieved from the database.

REST API

- RESTGet.java presents the HTTP GET method.
- RESTPost.java presents the HTTP POST method.
- RESTPut.java presents the HTTP PUT method.
- RESTAll.java runs all examples at one go.

XML:DB API (deprecated)

Note that the XML:DB API does not talk to the server and can thus only be used in embedded mode.

- XMLDBCreate.java creates a collection using XML:DB.
- XMLDBQuery.java runs a query using XML:DB.
- XMLDBInsert.java inserts a document into a database using XML:DB.
Chapter 95. Docker

Read this entry online in the BaseX Wiki.

This page is part of the Developer Section.

The BaseX server is also available as automated build `basex/basexhttp` on the Docker Hub, providing both release and nightly builds. All images are automatically rebuilt if Docker provides updated base images.

**Running a BaseX Container**

To start a BaseX container based on the latest development release publishing the BaseX server and HTTP ports 1984 and 8984 and bind-mounting your user's `BaseXData` directory, run

```
docker run -ti \
    --name basexhttp \
    --publish 1984:1984 \
    --publish 8984:8984 \
    --volume ~/BaseXData:/srv/BaseXData \
    basex/basexhttp:latest
```

By passing any other BaseX executable, you can also for example run a BaseX client connecting to the linked BaseX server for management operations on the BaseX command line:

```
docker run -ti \
    --link basexhttp:basexhttp \
    basex/basexhttp:latest basexclient -nbasexhttp
```

BaseX is run under the `basex` user with fixed user ID 1984. The user's home directory is `/srv`. Several ports are exposed:

<table>
<thead>
<tr>
<th>Port</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1984</td>
<td>Server port</td>
</tr>
<tr>
<td>8984</td>
<td>HTTP port</td>
</tr>
<tr>
<td>8985</td>
<td>HTTP stop port</td>
</tr>
</tbody>
</table>

Leaving BaseX' defaults but --publishing them under another external port is recommended if you want to change the ports.

**Administration with DBA**

If you prefer the DBA web interface, this can also be linked against your server container:

```
docker run -d \
    --name basex-dba \
    --publix 18984:8984 \
    --link basexhttp:basexhttp \
    basex/dba
```

The DBA is now available as `http://localhost:8984/dba` (adjust host name and port as needed). When logging in, connect to the linked container `basexhttp` by entering `basexhttp:1984` into the Address field.

**Security Considerations**

The Docker image ships the unchanged default credentials. Especially if you publish the server port 1984 or link a public DBA instance against the container, make sure to change the default credentials by adding a custom `.basex` file.
When publishing ports, consider which interfaces to bind to, paying special attention to the server port 1984 and a possibly linked DBA web interface.

A common use case will be linking a well-researched and mature reverse proxy link nginx against the application container. Goals are to reduce exposure of BaseX and Jetty, adding TLS-encryption, serve static resources like images and perform URL rewrites as needed. For database administration and ad-hoc queries, an instance of the DBA will be linked with restricted access. Port `1984` is not published, but only linked to the DBA; the same applies to the HTTP port 8984 if you use a reverse proxy. If you need to access the command line, you can always `docker exec` into the container and run `baseclient`.

**Running your own Application in a Docker Image**

If you want to add your own application, create an image FROM basex/basexhttp:[tag] with [tag] being the BaseX version you're developing against. Usually, you will add your application code to /srv/BaseXWeb and modules to /srv/BaseXModule. BaseXData is persisted as a volume, which means it cannot be preinitialized in the application image.

**Example: DBA**

An example for creating your own Docker image based on basex/basexhttp is the DBA application. A Dockerfile was added to the source code's root directory. The very simple file contains only few statements:

```bash
FROM basex/basexhttp:latest
MAINTAINER BaseX Team <basex-talk@mailman.uni-konstanz.de>
COPY . /srv/BaseXWeb
```

For general production usage, you should choose a fixed version instead of the development branch behind `latest`, so your application does not suddenly break because of unnoticed API changes. The most relevant part happens in the `COPY` statement, which adds the file's contents to the BaseXWeb directory. That's already it -- you're ready to run.

**Advanced Usage**

**BaseX Configuration**

If you need to adjust BaseX' configuration to tune the [Option option] defaults, add a `.baseX` file to /srv:

```bash
COPY .baseX /srv
```

Option not defined in the `.baseX` file with be automatically set to BaseX' default values.

**Jetty Configuration**

If you need to change the embedded web server’s configuration, you can always `COPY` a WEB-INF folder containing the required files and overwrite the predefined configuration.

**Java Runtime Parameters**

Larger applications and databases might require adjusted JRE parameters like increasing the memory limit. You can change those by setting the `BASEX_JVM` environment variable:

```bash
ENV BASEX_JVM="-Xmx2048m"
```

**Installing Debian Packages**

The basex/basexhttp Docker image is build own the official Maven Docker image, which again derives from Debian. You can add arbitrary Debian packages. Make sure to switch to the root user's context before
installing packages and back to the basex user afterwards. As common in the Docker environment, you need to fetch the package catalog using `apt-get update` before installing packages and should clean up afterwards to keep the image small. This example installs some libraries required for image manipulation and adds them to the `$CLASSPATH`:

```
USER root
RUN apt-get update && \
    DEBIAN_FRONTEND=noninteractive apt-get install --no-install-recommends -y \
    libbatik-java libxmlgraphics-commons-java libcommons-codec-java && \
    apt-get clean && rm -rf /var/lib/apt/lists/* /tmp/* /var/tmp/*

USER basex
ENV CLASSPATH="/usr/share/java/xml-commons-external.jar:/usr/share/java/\nxmlgraphics-commons.jar:/usr/share/java/batik.jar:usr/share/java/commons-codec.jar" \
```
Part X. Advanced User's Guide
Chapter 96. Advanced User's Guide

Read this entry online in the BaseX Wiki.

This page is one of the Main Sections of the documentation. It contains details on the BaseX storage and the Server architecture, and presents some more GUI features.

Storage

- **Configuration**: BaseX start files and directories
- **Backups**: Backup and restore databases
- **Catalog Resolver**: Information on entity resolving
- **Storage Layout**: How data is stored in the database files

Use Cases

- **Statistics**: Exemplary statistics on databases created with BaseX
- **Twitter**: Storing live tweets in BaseX

Server and Query Architecture

- **User Management**: User management in the client/server environment
- **Transaction Management**: Insight into the BaseX transaction management
- **Logging**: Description of the server logs
Chapter 97. Configuration

This article is part of the Advanced User's Guide. It gives some more insight into the configuration of BaseX.

Configuration Files

BaseX maintains some configuration files, which are stored in the project’s Home Directory:

- .basex contains all options that are relevant for running the server or standalone versions of BaseX.
- .basexgui defines all options relevant to the BaseX GUI.
- .basexhistory contains commands that have been typed in most recently.
- .basexhome can be created by a user to mark a folder as home directory. Its contents do not matter, so it is usually empty.

Note that:

- Depending on your OS and configuration, files and folders with a ‘.’ prefix may be hidden.
- In the Web Application context, options can be defined in the web.xml file.

Home Directory

As BaseX is distributed in different flavors, and may be started from different locations, it dynamically determines its home directory:

- First, the system property org.basex.path is checked. If it contains a value, it is chosen as directory path.
- If not, the current user directory (defined by the system property user.dir) is chosen if the .basex or .basexhome file is found in this directory.
- Otherwise, the files are searched in the application directory (the folder in which the BaseX code is located).
- In all other cases, the user’s home directory (defined in user.home) is chosen.

If BaseX is used in an embedded environment (such as a servlet in a Web Application), it may not immediately be clear what directory was chosen. In this case, you can run the XQuery expression Q{org.basex.util.Prop}USERHOME() to find out.

Database Directory

A database in BaseX consists of several files, which are located in a directory named by the name of the database. If the user’s home directory has been chosen as base directory, the database directories will be planed in a BaseXData directory. Otherwise, the directory will be named data.

The database path can be changed as follows:

- GUI: Choose Options → Preferences and choose a new database path.
- General: edit the DBPATH option in the .basex configuration file

Note: Existing databases will not be automatically moved to the new destination.

Log Files

Log files are stored in text format in a .logs sub-directory of the database folder (see Logging for more information).
Changelog

Version 8.0

• Updated: `.basexpem` is obsolete. Users are now stored in `users.xml` in the database directory (see User Management for more information).

Version 7.7

• Updated: the `.basehome` file marks a folder as home directory.
Chapter 98. Backups

Read this entry online in the BaseX Wiki.

This page is part of the Advanced User's Guide. The following two paragraphs demonstrate how to create a backup and restore the database within BaseX.

GUI Example

1. Start the BaseX GUI and create a new database in Database → New... with your XML document.
2. Go to Database → Manage... and create a backup of your database. The backup will be created in the database directory.
3. Go to Database → Add... and add another document.
4. Go to Database → Manage... and restore your database. The database will be restored from the latest backup of to the database found in the database directory.

Console Example

1. Start the BaseX Standalone client from a console.
2. Create a new database via the CREATE DB command.
3. Use the CREATE BACKUP command to back up your database.
4. Add a new document via ADD: ADD AS newdoc.xml <newdoc/>
5. Use the RESTORE command to restore the original database.
6. Type in XQUERY / to see the restored database contents.

The same commands can be used with a BaseX client connected to a remote Database Server.
Chapter 99. Catalog Resolver

Read this entry online in the BaseX Wiki.

This article is part of the Advanced User's Guide. It clarifies how to deal with external DTD declarations when parsing XML data.

Overview

XML documents often rely on Document Type Definitions (DTDs). While parsing a document with BaseX, entities can be resolved with respect to that particular DTD. By default, the DTD is only used for entity resolution.

XHTML, for example, defines its doctype via the following line:

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
```

Fetching `xhtml1-strict.dtd` obviously involves network traffic. When dealing with single files, this may seem tolerable, but importing large collections benefits from caching these resources. Depending on the remote server, you will experience significant speed improvements when caching DTDs locally.

XML Entity and URI Resolvers

BaseX comes with a default URI resolver that is usable out of the box.

To enable entity resolving you have to provide a valid XML Catalog file, so that the parser knows where to look for mirrored DTDs.

A simple working example for XHTML might look like this:

```
<?xml version="1.0"?>
<catalog prefer="system" xmlns:urn:oasis:names:tc:entity:xmlns:xml:catalog>
  <rewriteSystem systemIdStartString="http://www.w3.org/TR/xhtml1/DTD/" rewritePrefix="file:///path/to/dtds/" />
</catalog>
```

This rewrites all systemIds starting with: `http://www.w3.org/TR/xhtml1/DTD/` to `file:///path/to/dtds/`.

The XHTML DTD `xhtml1-strict.dtd` and all its linked resources will now be loaded from the specified path.

GUI Mode

When running BaseX in GUI mode, simply provide the path to your XML Catalog file in the Parsing Tab of the Database Creation Dialog.

Console & Server Mode

To enable Entity Resolving in Console Mode, specify the following options:

- `SET CATFILE [path]`

Now entity resolving is active for the current session. All subsequent `ADD` commands will use the catalog file to resolve entities.

The paths to your catalog file and the actual DTDs are either absolute or relative to the current working directory. When using BaseX in Client-Server-Mode, this is relative to the server's working directory.
Please Note

Entity resolving only works if the internal XML parser is switched off (which is the default case). If you use the internal parser, you can manually specify whether you want to parse DTDs and entities or not.

Using other Resolvers

There might be some cases when you do not want to use the built-in resolver that Java provides by default (via com.sun.org.apache.xml.internal.resolver.*).

BaseX offers support for the Apache-maintained XML Commons Resolver, available for download here.

To use it add resolver.jar to the classpath when starting BaseX:

```
java -cp basex.jar:resolver.jar org.basex.BaseXServer
```

More Information

- Wikipedia on Document Type Definitions
- Apache XML Commons Article on Entity Resolving
- XML Entity and URI Resolvers, Sun
Chapter 100. Storage Layout

This article is part of the Advanced User's Guide. It presents some low-level details on how data is stored in the database files.

Data Types

The following data types are used for specifying the storage layout:

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>Example (native → hex integers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Num</td>
<td>Compressed integer (1-5 bytes), specified in Num.java</td>
<td>15 → 0F; 511 → 41 FF</td>
</tr>
<tr>
<td>Token</td>
<td>Length (Num) and bytes of UTF8 byte representation</td>
<td>Hello → 05 48 65 6c 6c 6f</td>
</tr>
<tr>
<td>Double</td>
<td>Number, stored as token</td>
<td>123 → 03 31 32 33</td>
</tr>
<tr>
<td>Boolean</td>
<td>Boolean (1 byte, 00 or 01)</td>
<td>true → 01</td>
</tr>
<tr>
<td>Nums, Tokens, Doubles</td>
<td>Arrays of values, introduced with the number of entries</td>
<td>1,2 → 02 01 31 01 32</td>
</tr>
<tr>
<td>TokenSet</td>
<td>Key array (Tokens), next/bucket/size arrays (3x Nums)</td>
<td></td>
</tr>
</tbody>
</table>

Database Files

The following tables illustrate the layout of the BaseX database files. All files are suffixed with .basex.

Meta Data, Name/Path/Doc Indexes: inf

<table>
<thead>
<tr>
<th>Description</th>
<th>Format</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Meta Data</td>
<td>1. Key/value pairs, in no particular order (Token / Token): • Examples: FNAME, TIME, SIZE, ... • PERM → Number of users (Num), and name/password/permission values for each user (Token / Token / Num). Empty key as finalizer</td>
<td>DiskData() MetaData() Users()</td>
</tr>
<tr>
<td>2. Main memory indexes</td>
<td>1. Key/value pairs, in no particular order (Token / Token): • TAGS → Element Name Index • ATTS → Attribute Name Index • PATH → Path Index • NS → Namespaces • DOCS → Document Index. Empty key as finalizer</td>
<td>DiskData()</td>
</tr>
<tr>
<td>2 a) Name IndexElement/attribute names</td>
<td>1. Token set, storing all names (TokenSet). One StatsKey instance per entry. 2.1. Content kind (Num). 2.1.1. Number: min/max (Doubles). 2.1.2. Category: number of entries (Num), entries (Tokens). 2.2. Number of entries (Num). 2.3. Leaf flag (Boolean).</td>
<td>Names() TokenSet.read() StatsKey()</td>
</tr>
</tbody>
</table>
### Storage Layout

<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
</table>
| 2 b) Path Index | **1.** Flag for path definition (*Boolean*, always *true*; legacy)  
**2.** PathNode: **2.1.** Name reference (*Num*)  
**2.2.** Node kind (*Num*)  
**2.3.** Number of occurrences (*Num*)  
**2.4.** Number of children (*Num*)  
**2.5.** *Double*; legacy, can be reused or discarded  
**2.6.** Recursive generation of child nodes (→ 2) |
| 2 c) Namespaces | **1.** Token set, storing prefixes (*TokenSet*)  
**2.** Token set, storing URIs (*TokenSet*)  
**3.** NSNode: **3.1.** pre value (*Num*)  
**3.2.** References to prefix/URI pairs (*Nums*)  
**3.3.** Number of children (*Num*)  
**3.4.** Recursive generation of child nodes (→ 3) |
| 2 d) Document Index | Array of integers, representing the distances between all document pre values (*Nums*) |

---

**Node Table: tbl, tbli**
- **tbl**: Main database table, stored in blocks.
- **tbli**: Database directory, organizing the database blocks.

Some more information on the node storage is available.

**Texts: txt, atv**
- **txt**: Heap file for text values (document names, string values of texts, comments and processing instructions)
- **atv**: Heap file for attribute values.

**Value Indexes: txtl, txtr, atvl, atvr**

**Text Index:**
- **txtl**: Heap file with ID lists.
- **txtr**: Index file with references to ID lists.

The **Attribute Index** is contained in the files **atvl** and **atvr**, the **Token Index** in **tokl** and **tokr**. All have the same layout.

For a more detailed discussion and examples of these file formats please see Index File Structure.

**Full-Text Fuzzy Index: ftxx, ftxy, ftxz**

...may soon be reimplemented.
Chapter 101. Node Storage

Read this entry online in the BaseX Wiki.

This article describes the Storage Layout of the main database table.

Node Table

BaseX stores all XML nodes in a flat table. The node table of a database can be displayed via the INFO STORAGE command:

```
$ basex -c:"create db db <xml>HiThere</xml>" -c:"info storage"
PRE DIS SIZ ATS ID NS KIND CONTENT
-----------------------------
0 1 3 1 0 0 DOC db.xml
1 1 2 1 1 0 ELEM xml
2 1 1 1 2 0 TEXT HiThere
```

PRE Value

The pre value of a node represents the order in which the XML nodes are visited by a SAX parser. It is actually not stored in the database; instead, it is implicitly given by the table position. As a result, it will change whenever a node with a smaller pre values is added to or deleted from a database.

ID Value

Each database node has a persistent id value, which remains valid after update operations, and which is referenced by the value indexes. As long as no updates are performed on a database, the pre and id values are identical. The values will remain to be identical if new nodes are exclusively added to the end of the database. If nodes are deleted or inserted somewhere else, the values will diverge, as shown in the next example:

```
$ basex -c:"create db db <xml>HiThere</xml>" -q:"insert node <b/> before /xml" -c:"info storage"
PRE DIS SIZ ATS ID NS KIND CONTENT
-----------------------------
0 1 4 1 0 0 DOC db.xml
1 1 1 1 3 0 ELEM b
2 1 1 1 0 ELEM xml
3 1 1 2 0 TEXT HiThere
```

The db:node-pre and db:node-id functions can be called to retrieve the pre and id values of a node, and db:open-pre and db:open-id can be used to go back and retrieve the original node. By default, id lookups are expensive. If the UPDINDEX option is turned on, an additional index will be maintained to speed up the process.

Block Storage

BaseX logically splits the tbl.base file into blocks with length 4096 bytes, i.e. each block can have max 256 records each with length 16 bytes. The records within a block are sorted by their pre value (which, therefore, can be implicitly determined and need not be saved).

For each block BaseX stores in a separate file (tbli.base) the smallest pre value within that block (and since the records are sorted, that will be the pre value of the first record stored in the block). These will be referred as fpre from now on. The physical address of each block is stored in tbli.base, too.

Since these two maps will not grow excessively large, but are accessed resp. changed on each read resp. write operation, they are kept in main memory and flushed to disk on closing the database.
A newly created database with 256 + 10 records will occupy the first two blocks with physical addresses 0 and 4096. The corresponding fpre's will be 0 and 256.

If a record with pre = 12 is to be inserted, it needs to be stored in the first block, which is, however, full. In this case, a new block with physical address 8192 will be allocated, the records with pre values from 12 to 255 will be copied to the new block, the new record will be stored in the old block at pre = 12, and the two maps will look like this:

\[
\text{fpre}'s = 0, 13, 257 \\
\text{addr}'s = 0, 8192, 4096
\]

Basically, the old records remain in the first block, but they will not be read, since the fpre's array says that only 13 records are stored in the first block. This causes redundant storage of the records with old pres from 13 to 255.

Additionally to these two maps (fpre's and addr's), BaseX maintains a bit map (which is also stored in \texttt{tbli.basex}) which reflects which physical blocks are free and which not, so that when a new block is needed, an already free one will be reused.
Chapter 102. User Management

Read this entry online in the BaseX Wiki.

This article is part of the Advanced User's Guide. The user management defines which permissions are required by a user to perform a database command or XQuery expression.

Permissions are mostly relevant in the client/server architecture, as the Standalone Mode and the GUI is run with admin permissions. There are a few exceptions such as the xquery:eval function: its execution scope can also be limited by specifying a permission.

With Version 8.0, the user management has been revised:

- Permissions can now be manually edited, as they are stored as XML.
- The permission file has been moved from the home directory to the database directory. It was renamed from .baseperm to users.xml.
- Local permissions are now defined for database glob patterns instead of single databases. Both local and global permissions are stored in the same file.
- A new User Module is available, which allows user management via XQuery.
- MD5 password hashing has been discarded, as many md5 hashes can be easily uncovered with rainbow tables.
- Salted sha256 hashes are now used for authentication (the current timestamp will be used as salt).
- Additionally, digest hashes are used in the client/server architecture and the Language Bindings, and in the HTTP Context if the AUTHMETHOD is set to Digest.

Warning: As the available md5 hashes cannot automatically be converted to the new format, existing credentials will be ignored, and you will need to recreate your user data. Moreover, we will incrementally provide new Language Bindings, which will be based on the digest hashes.

Passwords in commands and XQuery functions are now specified in plain text. The rationale behind this is:

- Transmission of passwords as md5 has not been safe anyway, as indicated above.
- Different hash values can now be created from the original password.

You will be safe if you follow basic security measures: ensure that your password will not end up in your bash history, avoid sending passwords via ordinary REST requests, etc.

Rules

In the permission hierarchy below, the existing permissions are illustrated. A higher permission includes all lower permissions. For example, all users who have the write permission assigned will also be able to execute commands requiring read permission.

Local permissions are applied to databases. They have a higher precedence and override global permissions.
Permissions hierarchy

User names must follow the valid names constraints, and the database patterns must follow the Glob Syntax.

### Commands

Admin permissions are required to execute all of the following commands:

**Create user 'test' (password will be entered on command line). By default, the user will have no permissions ('none'):**

```shell
> CREATE USER test
```

**Change password of user 'test' to '71x343sd#':**

```shell
> ALTER PASSWORD test 71x343sd#
```

**Grant local write permissions to user 'test':**

```shell
> GRANT write ON unit* TO test
```

Note: Local permissions overwrite global permissions. As a consequence, the 'test' user will only be allowed to access (i.e., read and write) database starting with the letters 'unit'. If no local permissions are set, the global rights are inherited.

**Show global permissions:**

```shell
> SHOW USERS
```

**Show detailed information about user 'test' via XQuery:**

```shell
> XQUERY user:list-details()[@name = 'test']
```

**Drop of user 'test' via XQuery:**

```shell
> XQUERY user:drop('test')
```

### Changelog

Revised in Version 8.0.
Chapter 103. Transaction Management

Read this entry online in the BaseX Wiki.

This article is part of the Advanced User's Guide. The BaseX client-server architecture offers ACID-safe transactions, with multiple readers and writers. Here is some more information about the transaction management.

Introduction

In a nutshell, a transaction is equal to a command or query. So each command or query sent to the server becomes a transaction.

Incoming requests are parsed and checked for errors on the server. If the command or query is not correct, the request will not be executed, and the user will receive an error message. Otherwise the request becomes a transaction and gets into the transaction monitor.

Please note that:

- Locks cannot be synchronized across BaseX instances that run in different JVMs. If concurrent write operations are to be performed, we generally recommend working with the client/server or the HTTP architecture.

- An unexpected abort of the server during a transaction, caused by a hardware failure or power cut, may lead to an inconsistent database state if a transaction was active at shutdown time. So it is advisable to use the BACKUP command to regularly backup your database. If the worst case occurs, you can try the INSPECT command to check if your database has obvious inconsistencies, and use RESTORE to restore the last backed up version of the database.

XQuery Update

Many update operations are triggered by XQuery Update expressions. When executing an updating query, all update operations of the query are stored in a pending update list. They will be executed all at once, so the database is updated atomically. If any of the update sub-operations is erroneous, the overall transaction will be aborted.

Concurrency Control

BaseX provides support for multiple read and single write operations (using preclaiming and starvation-free two phase locking). This means that:

- Read transactions are executed in parallel.

- If an updating transaction comes in, it will be queued and executed after all previous read transaction have been executed.

- Subsequent operations (read or write) will be queued until the updating transaction has completed.

Each database has its own queue: An update on database A will not block operations on database B. This is under the premise that it can be statically determined, i.e., before the transaction is evaluated) which databases will be accessed by a transaction (see below). The number of maximum parallel transactions can be adjusted with the PARALLEL option.

External Side Effects

Access to external resources (files on hard disk, HTTP requests, ...) is not controlled by the transaction monitor of BaseX unless specified by the user.
**XQuery Locking Options**

Custom locks can be acquired by setting the BaseX-specific XQuery options `query:read-lock` and `query:write-lock`. Multiple option declarations may occur in the prolog of a query, but multiple values can also be separated with commas in a single declaration. These locks are in another namespace than the database names: the lock value `factbook` will not lock a database named `factbook`.

These option declarations will put read locks on `foo`, `bar` and `batz` and a write lock on `quix`:

```xquery
declare option query:read-lock "foo,bar";
declare option query:read-lock "batz";
declare option query:write-lock "quix";
```

**Java Modules**

Locks can also be acquired on Java functions which are imported and invoked from an XQuery expression. It is advisable to explicitly lock Java code whenever it performs sensitive read and write operations.

**Limitations**

**Commands**

Database locking works with all commands unless the glob syntax is used, such as in the following command call:

- `DROP DB new*:` drop all databases starting with "new"

**XQuery**

Deciding which databases will be accessed by a complex XQuery expression is a non-trivial task. Database detection works for the following types of queries:

- `//item`, read-locking of the database opened by a client
- `doc('factbook')`, read-locking of "factbook"
- `collection('db/path/to/docs')`, read-locking of "db"
- `fn:sum(1 to 100)`, locking nothing at all
- `delete nodes doc('test')//*[string-length(local-name(.)) > 5]`, write-locking of "test"

All databases will be locked by queries of the following kind:

- `for $db in ('db1', 'db2') return doc($db)`
- `doc(doc('test')/reference/text())`
- `let $db := 'test' return insert nodes <test/> into doc($db)`

You can consult the query info output (which you find in the Info View of the GUI or which you can turn on by setting `QUERYINFO` to `true`) to find out which databases have been locked by a query.

**GUI**

Database locking is currently disabled if the BaseX GUI is used.

**Process Locking**

In order to enable locking on global (process) level, the option `GLOBALLOCK` can be set to `true`. This can e.g. be done by editing your `.basex` file (see Options for more details). If process locking is active, a process that performs write operations will queue all other operations.
File-System Locks

Update Operations

During a database update, a locking file `upd.basex` will reside in that database directory. If the update fails for some unexpected reason, or if the process is killed ungracefully, this file will not be deleted. In this case, the database cannot be opened anymore, and the message "Database ... is being updated, or update was not completed" will be shown instead.

If the locking file is manually removed, you may be able to reopen the database, but you should be aware that database may have got corrupt due to the interrupted update process, and you should revert to the most recent database backup.

Database Locks

To avoid database corruptions that are caused by accidental write operations from different JVMs, a shared lock is requested on the database table file (`tbl.basex`) whenever a database is opened. If an update operation is triggered, and if no exclusive lock can be acquired, it will be rejected with the message "Database ... is opened by another process.”.

Please note that you cannot 100% rely on this mechanism, as it is not possible to synchronize operations across different JVMs. You will be safe when using the client/server or HTTP architecture.

Changelog

Version 7.8

• Added: Locks can also be acquired on Java functions.

Version 7.6

• Added: database locking introduced, replacing process locking.

Version 7.2.1

• Updated: pin files replaced with shared/exclusive filesystem locking.

Version 7.2

• Added: pin files to mark open databases.

Version 7.1

• Added: update lock files.
Chapter 104. Logging

Read this entry online in the BaseX Wiki.

This article is part of the Advanced User's Guide. It describes how client operations are logged by the server. The server logs can e.g. be used to get an overview of all processes executed on your server, trace any errors or compile performance statistics.

The server logs are written in plain text. In your Database Directory, you can find a folder named .logs in which all log files are stored with the according date. Note that, depending on your OS and configuration, files and folders beinning with a . may be hidden.

Some more notes on the logging facility:

• HTTP requests are included in the log files.

• Logging can be turned on/off via the LOG option.

• The maximum length of logging messages can be changed via LOGMSGMAXLEN.

• The Admin Module provides access to the log files from XQuery.

Format

Example 1

<table>
<thead>
<tr>
<th>Time</th>
<th>IP Address</th>
<th>User</th>
<th>Action</th>
<th>Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>01:18:12.892</td>
<td>SERVER</td>
<td>admin</td>
<td>OK</td>
<td>Server was started (port: 1984)</td>
</tr>
<tr>
<td>01:18:15.436</td>
<td>127.0.0.1:4722</td>
<td>jack</td>
<td>REQUEST</td>
<td>XQUERY for $i in 1 to 5 return random:double()</td>
</tr>
<tr>
<td>01:18:15.446</td>
<td>127.0.0.1:4722</td>
<td>jack</td>
<td>OK</td>
<td>Query executed in 2.38 ms.</td>
</tr>
<tr>
<td>01:18:15.447</td>
<td>127.0.0.1:4722</td>
<td>jack</td>
<td>REQUEST</td>
<td>EXIT</td>
</tr>
<tr>
<td>01:18:15.447</td>
<td>127.0.0.1:4722</td>
<td>jack</td>
<td>OK</td>
<td>0.39 ms</td>
</tr>
</tbody>
</table>

A server has been started and a user jack has connected to the server to perform a query and exit properly.

Example 2

<table>
<thead>
<tr>
<th>Time</th>
<th>IP Address</th>
<th>User</th>
<th>Action</th>
<th>Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>01:23:33.251</td>
<td>127.0.0.1:4736</td>
<td>john</td>
<td>OK</td>
<td>QUERY[0] 'hi' 0.44 ms</td>
</tr>
<tr>
<td>01:23:33.337</td>
<td>127.0.0.1:4736</td>
<td>john</td>
<td>OK</td>
<td>ITER[0] 1.14 ms</td>
</tr>
<tr>
<td>01:23:33.338</td>
<td>127.0.0.1:4736</td>
<td>john</td>
<td>OK</td>
<td>INFO[0] 0.36 ms</td>
</tr>
<tr>
<td>01:23:33.339</td>
<td>127.0.0.1:4736</td>
<td>john</td>
<td>OK</td>
<td>CLOSE[0] 0.21 ms</td>
</tr>
<tr>
<td>01:23:33.359</td>
<td>127.0.0.1:4736</td>
<td>john</td>
<td>REQUEST</td>
<td>EXIT</td>
</tr>
<tr>
<td>01:23:33.359</td>
<td>127.0.0.1:4736</td>
<td>john</td>
<td>OK</td>
<td>0.14 ms</td>
</tr>
</tbody>
</table>

A user john has performed an iterative query, using one of the client APIs.

Example 3

<table>
<thead>
<tr>
<th>Time</th>
<th>IP Address</th>
<th>User</th>
<th>Action</th>
<th>Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>01:31:51.892</td>
<td>127.0.0.1:4803</td>
<td>admin</td>
<td>200</td>
<td>4.43 ms</td>
</tr>
</tbody>
</table>

An admin user has accessed the factbook database via REST.
Part XI. Use Cases
Chapter 105. Statistics

Read this entry online in the BaseX Wiki.

This article is part of the Advanced User's Guide. It lists statistics on various XML instances that have been created with BaseX, with the value and full-text indexes turned off. The URLs to the original sources, if available or public, are listed below.

### Databases

- **FileSize** is the original size of the input documents
- **#Files** indicates the number of stored XML documents
- **DbSize** is the size of the resulting database (excluding the value index structures)
- **#Nodes** represents the number of XML nodes (elements, attributes, texts, etc.) stored in the database
- **#Attr** indicates the maximum number of attributes stored for a single element
- **#ENames** and **#ANames** reflect the number of distinct element and attribute names
- **#URIs** represent the number of distinct namespace URIs
- **Height** indicates the maximum level depth of the stored nodes

If a fixed database limit is reached, documents can be distributed in several database instances, which can then accessed from a single XQuery expression.

<table>
<thead>
<tr>
<th>Instances</th>
<th>FileSize</th>
<th>#Files</th>
<th>DbSize</th>
<th>#Nodes</th>
<th>#Attr</th>
<th>#ENames</th>
<th>#ANames</th>
<th>#URIs</th>
<th>Height</th>
</tr>
</thead>
<tbody>
<tr>
<td>RuWikiHist</td>
<td>421 GiB</td>
<td>1</td>
<td>416 GiB</td>
<td>324'848'508</td>
<td>21</td>
<td>6</td>
<td>2</td>
<td>6</td>
<td>no limit</td>
</tr>
<tr>
<td>ZhWikiHist</td>
<td>26 GiB</td>
<td>1</td>
<td>120 GiB</td>
<td>179'199'663</td>
<td>21</td>
<td>6</td>
<td>2</td>
<td>6</td>
<td>no limit</td>
</tr>
<tr>
<td>EnWiktionary</td>
<td>29 GiB</td>
<td>1</td>
<td>75 GiB</td>
<td>134'380'393</td>
<td>21</td>
<td>6</td>
<td>2</td>
<td>6</td>
<td>2^8</td>
</tr>
<tr>
<td>XMark</td>
<td>55 GiB</td>
<td>1</td>
<td>64 GiB</td>
<td>1'615'071'328</td>
<td>74</td>
<td>9</td>
<td>0</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>EnWikiMedia</td>
<td>64 GiB</td>
<td>1</td>
<td>52 GiB</td>
<td>401'456'348</td>
<td>21</td>
<td>6</td>
<td>2</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>MedLine</td>
<td>38 GiB</td>
<td>379</td>
<td>36 GiB</td>
<td>1'623'764'254</td>
<td>84</td>
<td>6</td>
<td>0</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>iProClass</td>
<td>36 GiB</td>
<td>1</td>
<td>37 GiB</td>
<td>1'631'218'984</td>
<td>245</td>
<td>4</td>
<td>2</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Inex2009</td>
<td>31 GiB</td>
<td>2'666'500</td>
<td>34 GiB</td>
<td>1'336'110'639</td>
<td>28'034</td>
<td>451</td>
<td>1</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>CoPhIR</td>
<td>29 GiB</td>
<td>10'000'000</td>
<td>31 GiB</td>
<td>1'104'623'370</td>
<td>42</td>
<td>42</td>
<td>0</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>EnWikipedia</td>
<td>26 GiB</td>
<td>1</td>
<td>25 GiB</td>
<td>198'546'743</td>
<td>24</td>
<td>21</td>
<td>2</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>XMark</td>
<td>22 GiB</td>
<td>1</td>
<td>26 GiB</td>
<td>645'997'962</td>
<td>74</td>
<td>9</td>
<td>0</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>InterPro</td>
<td>14 GiB</td>
<td>1</td>
<td>19 GiB</td>
<td>860'304'235</td>
<td>7</td>
<td>15</td>
<td>0</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Genome1</td>
<td>13 GiB</td>
<td>1</td>
<td>13 GiB</td>
<td>432'628'1032</td>
<td>26</td>
<td>101</td>
<td>2</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>NewYorkTimes</td>
<td>64 GiB</td>
<td>1</td>
<td>1855'659</td>
<td>13 GiB</td>
<td>280'407'005</td>
<td>41</td>
<td>33</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>TrEMBL</td>
<td>11 GiB</td>
<td>1</td>
<td>14 GiB</td>
<td>589'650'535</td>
<td>47</td>
<td>30</td>
<td>2</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>XMark</td>
<td>11 GiB</td>
<td>1</td>
<td>13 GiB</td>
<td>323'083'402</td>
<td>74</td>
<td>9</td>
<td>0</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>IntAct</td>
<td>7973 MiB</td>
<td>25'624</td>
<td>6717 MiB</td>
<td>297'478'397</td>
<td>64</td>
<td>22</td>
<td>2</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Freebase</td>
<td>7366 MiB</td>
<td>10 GiB</td>
<td>443'627'994</td>
<td>81</td>
<td>283</td>
<td>1</td>
<td>93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDMX</td>
<td>6356 MiB</td>
<td>1</td>
<td>8028 MiB</td>
<td>395'871'872</td>
<td>22</td>
<td>6</td>
<td>3</td>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

408
<table>
<thead>
<tr>
<th>Dataset</th>
<th>Size</th>
<th>Size</th>
<th>Size</th>
<th>Size</th>
<th>Size</th>
<th>Size</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>OpenStreet</td>
<td>MiB</td>
<td>MiB</td>
<td>MiB</td>
<td>MiB</td>
<td>MiB</td>
<td>MiB</td>
<td>MiB</td>
</tr>
<tr>
<td>SwissProt</td>
<td>4604</td>
<td>5422</td>
<td>241'274'408</td>
<td>70</td>
<td>39</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>EURLex</td>
<td>4815</td>
<td>5532</td>
<td>167'328'032</td>
<td>186</td>
<td>46</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>Wikicorporus</td>
<td>649'238</td>
<td>4432</td>
<td>157'948'561</td>
<td>1'257</td>
<td>2'687</td>
<td>2</td>
<td>50</td>
</tr>
<tr>
<td>EnWiki</td>
<td>6679</td>
<td>3537</td>
<td>98'433'1941</td>
<td>11</td>
<td>2</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>CoPhIR</td>
<td>2695</td>
<td>1'000'000</td>
<td>101'638'857</td>
<td>42</td>
<td>42</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>MeSH</td>
<td>2091</td>
<td>2410</td>
<td>104'845'819</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>FreeDB</td>
<td>1723</td>
<td>2462</td>
<td>102'901'512</td>
<td>7</td>
<td>3</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>XMark</td>
<td>1134</td>
<td>1303</td>
<td>32'298'982</td>
<td>74</td>
<td>9</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>DeepFS</td>
<td>810</td>
<td>850</td>
<td>44'821'506</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>LibraryUK</td>
<td>560</td>
<td>918</td>
<td>46'401'941</td>
<td>23</td>
<td>3</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Twitter</td>
<td>736</td>
<td>1'177'495</td>
<td>767'815</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Organizational</td>
<td>633</td>
<td>724</td>
<td>33'112'3923</td>
<td>38</td>
<td>9</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>DBLP</td>
<td>694</td>
<td>944</td>
<td>36'878'1814</td>
<td>35</td>
<td>6</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Feeds</td>
<td>692</td>
<td>604</td>
<td>5'933'713</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>MedLine</td>
<td>449</td>
<td>273</td>
<td>14'512'851</td>
<td>111</td>
<td>5</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>MedLineDB</td>
<td>195</td>
<td>1040'1'847</td>
<td>66</td>
<td>8</td>
<td>0</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>ZDNET</td>
<td>130</td>
<td>133</td>
<td>3'060'186</td>
<td>40</td>
<td>90</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>JMNEdict</td>
<td>124</td>
<td>171</td>
<td>8'592'666</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>XMark</td>
<td>111</td>
<td>130</td>
<td>3'221'926</td>
<td>74</td>
<td>9</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Freshmeat</td>
<td>105</td>
<td>86</td>
<td>3'832'028</td>
<td>58</td>
<td>1</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>DeepFS</td>
<td>83</td>
<td>93</td>
<td>4'842'638</td>
<td>4</td>
<td>6</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>Treebank</td>
<td>82</td>
<td>92</td>
<td>3'829'513</td>
<td>250</td>
<td>1</td>
<td>0</td>
<td>37</td>
</tr>
<tr>
<td>DBLP2</td>
<td>80</td>
<td>102</td>
<td>4'044'649</td>
<td>35</td>
<td>6</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>DDI</td>
<td>76</td>
<td>39</td>
<td>2'070'157</td>
<td>104</td>
<td>16</td>
<td>21</td>
<td>11</td>
</tr>
<tr>
<td>Alfred</td>
<td>75</td>
<td>68</td>
<td>3'784'285</td>
<td>60</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>University</td>
<td>56</td>
<td>66</td>
<td>3'468'606</td>
<td>28</td>
<td>4</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>MediaUK</td>
<td>38</td>
<td>45</td>
<td>1'619'443</td>
<td>21</td>
<td>3</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>HCIBIB2</td>
<td>32</td>
<td>33</td>
<td>61'702'3</td>
<td>39</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Nasa</td>
<td>24</td>
<td>25</td>
<td>845'805</td>
<td>61</td>
<td>8</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>MovieDB</td>
<td>16</td>
<td>19</td>
<td>868'980</td>
<td>7</td>
<td>8</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>KanjiDic2</td>
<td>13</td>
<td>18</td>
<td>917'833</td>
<td>27</td>
<td>10</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>XMark</td>
<td>11</td>
<td>13</td>
<td>324'274</td>
<td>74</td>
<td>9</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Shakespear</td>
<td>771</td>
<td>9854</td>
<td>327'170</td>
<td>59</td>
<td>0</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>TreeOfLife</td>
<td>5425</td>
<td>7106</td>
<td>363'560</td>
<td>4</td>
<td>7</td>
<td>0</td>
<td>243</td>
</tr>
<tr>
<td>Thesaurus</td>
<td>4288</td>
<td>4088</td>
<td>201'798</td>
<td>33</td>
<td>9</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>MusicXML</td>
<td>3155</td>
<td>2942</td>
<td>171'400</td>
<td>179</td>
<td>56</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>BibDBPub</td>
<td>2292</td>
<td>2359</td>
<td>80'178</td>
<td>54</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Factbook</td>
<td>1743</td>
<td>1560</td>
<td>77'315</td>
<td>23</td>
<td>32</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>XMark</td>
<td>1134</td>
<td>1334</td>
<td>33'056</td>
<td>74</td>
<td>9</td>
<td>0</td>
<td>13</td>
</tr>
</tbody>
</table>
## Sources

<table>
<thead>
<tr>
<th>Instances</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>AirBase</td>
<td><a href="http://air-climate.eionet.europa.eu/databases/airbase/airbasexml">http://air-climate.eionet.europa.eu/databases/airbase/airbasexml</a></td>
</tr>
<tr>
<td>Alfred</td>
<td><a href="http://alfred.med.yale.edu/alfred/alfredWithDescription.zip">http://alfred.med.yale.edu/alfred/alfredWithDescription.zip</a></td>
</tr>
<tr>
<td>BibDBPub</td>
<td><a href="http://inex.is.informatik.uni-duisburg.de/2005/">http://inex.is.informatik.uni-duisburg.de/2005/</a></td>
</tr>
<tr>
<td>CoPhIR</td>
<td><a href="http://cophir.isti.cnr.it/">http://cophir.isti.cnr.it/</a></td>
</tr>
<tr>
<td>DBLP</td>
<td><a href="http://dblp.uni-trier.de/xml">http://dblp.uni-trier.de/xml</a></td>
</tr>
<tr>
<td>DBLP2</td>
<td><a href="http://inex.is.informatik.uni-duisburg.de/2005/">http://inex.is.informatik.uni-duisburg.de/2005/</a></td>
</tr>
<tr>
<td>DDI</td>
<td><a href="http://tools.ddialliance.org/">http://tools.ddialliance.org/</a></td>
</tr>
<tr>
<td>EnWikiRDF</td>
<td><a href="http://www.xml-benchmark.org/generated">http://www.xml-benchmark.org/generated</a> with xmlgen</td>
</tr>
<tr>
<td>EnWiktionary</td>
<td><a href="http://dumps.wikimedia.org/enwiktionary/latest/enwiktionary-latest-pages-meta-history.xml.7z">http://dumps.wikimedia.org/enwiktionary/latest/enwiktionary-latest-pages-meta-history.xml.7z</a></td>
</tr>
<tr>
<td>EURLex</td>
<td><a href="http://www.epsiplatform.eu/">http://www.epsiplatform.eu/</a></td>
</tr>
<tr>
<td>Factbook</td>
<td><a href="http://www.cs.washington.edu/research/xmldatasets/www/repository.html">http://www.cs.washington.edu/research/xmldatasets/www/repository.html</a></td>
</tr>
<tr>
<td>Freebase</td>
<td><a href="http://download.freebase.com/wex">http://download.freebase.com/wex</a></td>
</tr>
<tr>
<td>FreeDB</td>
<td><a href="http://www.xmldatabases.org/radio/xmlDatabases/projects/FreeDBtoXML">http://www.xmldatabases.org/radio/xmlDatabases/projects/FreeDBtoXML</a></td>
</tr>
<tr>
<td>Freshmeat</td>
<td><a href="http://freshmeat.net/articles/freshmeat-xml-rpc-api-available">http://freshmeat.net/articles/freshmeat-xml-rpc-api-available</a></td>
</tr>
<tr>
<td>HCIBIB2</td>
<td><a href="http://inex.is.informatik.uni-duisburg.de/2005/">http://inex.is.informatik.uni-duisburg.de/2005/</a></td>
</tr>
<tr>
<td>InterPro</td>
<td>ftp://ftp.bio.net/biomirror/interpro/match_complete.xml.gz</td>
</tr>
<tr>
<td>iProClass</td>
<td>ftp://ftp.pir.georgetown.edu/pir_databases/i proclass/i proclass.xml.gz</td>
</tr>
<tr>
<td>KanjiDic2</td>
<td><a href="http://www.csse.monash.edu.au/~jwb/kanjidic2">http://www.csse.monash.edu.au/~jwb/kanjidic2</a></td>
</tr>
<tr>
<td>MovieDB</td>
<td><a href="http://eagereyes.org/InfoVisContest2007Data.html">http://eagereyes.org/InfoVisContest2007Data.html</a></td>
</tr>
<tr>
<td>MusicXML</td>
<td><a href="http://www.recordare.com/xml/samples.html">http://www.recordare.com/xml/samples.html</a></td>
</tr>
<tr>
<td>Dataset</td>
<td>URL</td>
</tr>
<tr>
<td>-----------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>OpenStreetMap</td>
<td><a href="http://dumpwiki.openstreetmap.org/osmwiki-latest-files.tar.gz">http://dumpwiki.openstreetmap.org/osmwiki-latest-files.tar.gz</a></td>
</tr>
<tr>
<td>Organizations</td>
<td><a href="http://www.data.gov/raw/1358">http://www.data.gov/raw/1358</a></td>
</tr>
<tr>
<td>RuWikiHist</td>
<td><a href="http://dumps.wikimedia.org/ruwiki/latest/ruwiki-latest-pages-meta-history.xml.7z">http://dumps.wikimedia.org/ruwiki/latest/ruwiki-latest-pages-meta-history.xml.7z</a></td>
</tr>
<tr>
<td>SDMX</td>
<td><a href="http://www.metadatatechnology.com/">http://www.metadatatechnology.com/</a></td>
</tr>
<tr>
<td>Shakespeare</td>
<td><a href="http://www.cafeconleche.org/examples/shakespeare">http://www.cafeconleche.org/examples/shakespeare</a></td>
</tr>
<tr>
<td>Thesaurus</td>
<td><a href="http://www.drz.de/BELIT/thesaurus">http://www.drz.de/BELIT/thesaurus</a></td>
</tr>
<tr>
<td>Treebank</td>
<td><a href="http://www.cs.washington.edu/research/xmldatasets">http://www.cs.washington.edu/research/xmldatasets</a></td>
</tr>
<tr>
<td>TreeOfLife</td>
<td><a href="http://tolweb.org/data/tolskeletaldump.xml">http://tolweb.org/data/tolskeletaldump.xml</a></td>
</tr>
<tr>
<td>Wikicorpus</td>
<td><a href="http://www-connex.lip6.fr/~denoyer/wikipediaXML">http://www-connex.lip6.fr/~denoyer/wikipediaXML</a></td>
</tr>
<tr>
<td>XMark</td>
<td><a href="http://www.xml-benchmark.org/generated">http://www.xml-benchmark.org/generated</a> with xmlgen</td>
</tr>
<tr>
<td>ZDNET</td>
<td><a href="http://inex.is.informatik.uni-duisburg.de/2005/">http://inex.is.informatik.uni-duisburg.de/2005/</a></td>
</tr>
<tr>
<td>ZhWikiHist</td>
<td><a href="http://dumps.wikimedia.org/zhwiki/latest/zhwiki-latest-pages-meta-history.xml.7z">http://dumps.wikimedia.org/zhwiki/latest/zhwiki-latest-pages-meta-history.xml.7z</a></td>
</tr>
<tr>
<td>LibraryUKN</td>
<td>generated from university library data</td>
</tr>
<tr>
<td>MediaUKN</td>
<td>generated from university library data</td>
</tr>
<tr>
<td>DeepFS</td>
<td>generated from filesystem structure</td>
</tr>
<tr>
<td>University</td>
<td>generated from students test data</td>
</tr>
<tr>
<td>Feeds</td>
<td>compiled from news feeds</td>
</tr>
<tr>
<td>Twitter</td>
<td>compiled from Twitter feeds</td>
</tr>
</tbody>
</table>
Chapter 106. Twitter

Read this entry online in the BaseX Wiki.

This article is part of the Advanced User's Guide. It is about the usage of BaseX for processing and storing the live data stream of Twitter. We illustrate some statistics about the Twitter data and the performance of BaseX.

As Twitter attracts more and more users (over 140 million active users in 2012) and is generating large amounts of data (over 340 millions of short messages ('tweets') daily), it became a really exciting data source for all kind of analytics. Twitter provides the developer community with a set of APIs for retrieving the data about its users and their communication, including the Streaming API for data-intensive applications, the Search API for querying and filtering the messaging content, and the REST API for accessing the core primitives of the Twitter platform.

BaseX as Twitter Storage

For retrieving the Twitter stream we connect with the Streaming API to the endpoint of Twitter and receive a never ending tweet stream. As Twitter delivers the tweets as JSON objects the objects has to be converted into XML fragments. For this purpose the parse function of the XQuery JSON Module is used. In the examples section both versions are shown (tweet as JSON and tweet as XML). For storing the tweets including the meta-data, we use the standard insert function of XQuery Update.

Twitter’s Streaming Data

Each tweet object in the data stream contains the tweet message itself and over 60 data fields (for further information see the fields description). The following section shows the amount of data, that is delivered by the Twitter Streaming API to the connected endpoints with the 10% gardenhose access per hour on the 6th of the months February, March, April and May. It is the pure public live stream without any filtering applied.

Statistics

<table>
<thead>
<tr>
<th>Day</th>
<th>Description</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mon, 6-Feb-2012</td>
<td>Total tweets</td>
<td>30.824.976</td>
</tr>
<tr>
<td></td>
<td>Average tweets per hour</td>
<td>1.284.374</td>
</tr>
<tr>
<td></td>
<td>Average tweets per minute</td>
<td>21.406</td>
</tr>
<tr>
<td></td>
<td>Average tweets per second</td>
<td>356</td>
</tr>
<tr>
<td>Tue, 6-Mar-2012</td>
<td>Total tweets</td>
<td>31.823.776</td>
</tr>
<tr>
<td></td>
<td>Average tweets per hour</td>
<td>1.325.990</td>
</tr>
<tr>
<td></td>
<td>Average tweets per minute</td>
<td>22.099</td>
</tr>
<tr>
<td></td>
<td>Average tweets per second</td>
<td>368</td>
</tr>
<tr>
<td>Fri, 6-Apr-2012</td>
<td>Total tweets</td>
<td>34.638.976</td>
</tr>
<tr>
<td></td>
<td>Average tweets per hour</td>
<td>1.443.290</td>
</tr>
<tr>
<td>Date</td>
<td>Category</td>
<td>Value</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Sun, 6-May-2012</td>
<td>Total tweets</td>
<td>35,982,976</td>
</tr>
<tr>
<td></td>
<td>Average tweets per hour</td>
<td>1,499,290</td>
</tr>
<tr>
<td></td>
<td>Average tweets per minute</td>
<td>24,988</td>
</tr>
<tr>
<td></td>
<td>Average tweets per second</td>
<td>416</td>
</tr>
</tbody>
</table>

**Example Tweet (JSON)**

```json
{
  "contributors": null,
  "text": "Using BaseX for storing the Twitter Stream",
  "geo": null,
  "retweeted": false,
  "in_reply_to_screen_name": null,
  "possibly_sensitive": false,
  "truncated": false,
  "entities": {
    "urls": [],
    "hashtags": [],
    "user_mentions": []
  },
  "in_reply_to_status_id_str": null,
  "id": 1984009055807*****,
  "in_reply_to_user_id_str": null,
  "source": "<a href="http://twitterfeed.com" rel="nofollow">twitterfeed</a>",
  "favorited": false,
  "in_reply_to_status_id": null,
  "retweeted_count": 0,
  "created_at": "Fri May 04 13:17:16 +0000 2012",
  "in_reply_to_user_id": null,
  "possibly_sensitive_editable": true,
  "id_str": "1984009055807*****",
  "place": null,
  "user": {
    "location": "",
    "default_profile": true,
    "statuses_count": 9096,
    "profile_background_tile": false,
    "lang": "en",
    "profile_link_color": "0084B4",
    "id": 5024566*,
    "following": null,
    "protected": false,
    "favourites_count": 0,
    "profile_text_color": "333333",
    "contributors_enabled": false,
    "verified": false,
    "description": "http:////\basex.org",
    "profile_sidebar_border_color": "CODEED",
    "name": "BaseX",
    "profile_background_color": "CODEED",
    "created_at": "Sat Feb 25 04:05:30 +0000 2012",
    "default_profile_image": true,
    "followers_count": 860,
  }
}```
Example Tweet (XML)

```xml
<json booleans="retweeted possibly_sensitive truncated favorited possibly_sensitive_editable default_profile profile_background_image_tile protected contributors_enabled verified default_profile_image geo_enabled profile_use_background_image show_all_inline_media is_translator"
    numbers="id retweet_count statuses_count favourites_count followers_count friends_count listed_count"
    nulls="contributors geo in_reply_to_screen_name in_reply_to_status_id_str in_reply_to_user_id_str"
    in_reply_to_status_id in_reply_to_user_id place following follow_request_sent utc_offset time_zone notifications coordinates"
    arrays="urls indices hashtags user_mentions"
    objects="json entities user">
    <contributors/>
    <text>Using BaseX for storing the Twitter Stream</text>
    <geo/>
    <retweeted>false</retweeted>
    <in_reply_to_screen_name/>
    <possibly_sensitive>false</possibly_sensitive>
    <truncated>false</truncated>
    <entities/>
    <urls/>
    <user_mentions/>
    <in_reply_to_status_id_str/>
    <id>1984009055807*****</id>
    <in_reply_to_user_id_str/>
    <source><a href="http://twitterfeed.com" rel="nofollow">twitterfeed</a></source>
    <favorited>false</favorited>
    <in_reply_to_status_id/>
    <tweet_count>0</tweet_count>
    <created_at>Fri May 04 13:17:16 +0000 2012</created_at>
    <in_reply_to_status_id/>
    <possibly_sensitive_editable>true</possibly_sensitive_editable>
    <id_str>1984009055807*****</id_str>
</json>
```

Twitter
The test show the time BaseX needs to insert large amounts of real tweets into a database. We can derive that BaseX scales very well and can keep up with the incoming amount of tweets in the stream. Some lower values can occur, cause the size of the tweets differ according to the meta-data contained in the tweet object. Note: The AUTOFLUSH option is set to FALSE (default: SET AUTOFLUSH TRUE)

System Setup: Mac OS X 10.6.8, 3.2 GHz Intel Core i3, 8 GB 1333 MHz DDR3 RAM BaseX Version: BaseX 7.3 beta

Insert with XQuery Update

These tests show the performance of BaseX performing inserts with XQuery Update as single updates per tweet or bulk updates with different amount of tweets. The initial database just contained a root node `<tweets/>` and all incoming tweets are inserted after converting from JSON to XML into the root node. The time needed for the inserts includes the conversion time.
## Single Updates

<table>
<thead>
<tr>
<th>Amount of tweets</th>
<th>Time in seconds</th>
<th>Time in minutes</th>
<th>Database Size (without indexes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.000.000</td>
<td>492.26346</td>
<td>8.2</td>
<td>3396 MB</td>
</tr>
<tr>
<td>2.000.000</td>
<td>461.87326</td>
<td>7.6</td>
<td>6997 MB</td>
</tr>
<tr>
<td>3.000.000</td>
<td>470.7054</td>
<td>7.8</td>
<td>10452 MB</td>
</tr>
</tbody>
</table>