# Table of Contents

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

I. Getting Started ............................................................................................................ 3  
   2. Command-Line Options ............................................................................................ 4  
      BaseX Standalone ..................................................................................................... 4  
      BaseX Server .......................................................................................................... 5  
      BaseX Client ........................................................................................................... 6  
      BaseX HTTP Server ............................................................................................... 8  
      BaseX GUI .............................................................................................................. 10  
      Changelog ............................................................................................................... 10  

II. User Interfaces ........................................................................................................... 11  
   III. Getting Started ...................................................................................................... 11  
      Getting Started ...................................................................................................... 11  
      User Interfaces ....................................................................................................... 11  
      Tutorials and Slides ................................................................................................. 11  

I. Getting Started ............................................................................................................ 12  
   4. Start Scripts ............................................................................................................. 12  
      Main Package ........................................................................................................ 12  
      Windows: basex.bat ............................................................................................ 12  
      Linux/Mac: basex ............................................................................................... 12  
      HTTP Server ........................................................................................................ 13  
      Windows: basexhttp.bat .................................................................................... 13  
      Linux/Mac: basexhttp ....................................................................................... 13  
      Changelog ............................................................................................................. 13  

IV. Startup ....................................................................................................................... 15  
   5. Startup .................................................................................................................... 15  
      Requirements ........................................................................................................ 15  
      BaseX .................................................................................................................... 15  
      Java ....................................................................................................................... 15  
      BaseX GUI ........................................................................................................... 15  
      BaseX Standalone ............................................................................................... 15  
      BaseX Server ...................................................................................................... 16  
      BaseX Client ....................................................................................................... 16  
      BaseX HTTP Server ........................................................................................... 16  
      Changelog ............................................................................................................ 17  

V. Database Server ......................................................................................................... 19  
   6. Database Server ....................................................................................................... 19  
      Startup .................................................................................................................. 19  
      Create a database ................................................................................................. 19  
      Execute a query ................................................................................................... 19  
      Create a new database ......................................................................................... 19  
      Switch the database ............................................................................................ 20  
      Close or delete a database .................................................................................. 20  
      Create a collection .............................................................................................. 20  
      Delete a document ............................................................................................... 20  
      Delete a collection ............................................................................................... 21  
      Get server information ....................................................................................... 21  
      Backup and restore .............................................................................................. 21  
      See also ............................................................................................................... 21  

VI. Graphical User Interface ............................................................................................ 22  
   7. Graphical User Interface .......................................................................................... 22  
      Startup .................................................................................................................. 22  
      Create Database ................................................................................................. 22  
      Realtime Options ................................................................................................. 22  
      Querying ................................................................................................................ 22  
      Visualizations ...................................................................................................... 23  
      What’s Next? ....................................................................................................... 25
## BaseX Documentation

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>14. Options</strong></td>
<td>51</td>
</tr>
<tr>
<td>Global Options</td>
<td>51</td>
</tr>
<tr>
<td>General</td>
<td>51</td>
</tr>
<tr>
<td>Client/Server Architecture</td>
<td>52</td>
</tr>
<tr>
<td>HTTP Options</td>
<td>54</td>
</tr>
<tr>
<td>Create Options</td>
<td>55</td>
</tr>
<tr>
<td>General</td>
<td>55</td>
</tr>
<tr>
<td>Parsing</td>
<td>56</td>
</tr>
<tr>
<td>XML Parsing</td>
<td>57</td>
</tr>
<tr>
<td>Indexing</td>
<td>57</td>
</tr>
<tr>
<td>Full-Text</td>
<td>59</td>
</tr>
<tr>
<td>Query Options</td>
<td>59</td>
</tr>
<tr>
<td>QUERYINFO</td>
<td>59</td>
</tr>
<tr>
<td>XQUERY3</td>
<td>60</td>
</tr>
<tr>
<td>BINDINGS</td>
<td>60</td>
</tr>
<tr>
<td>QUERYPATH</td>
<td>60</td>
</tr>
</tbody>
</table>

| **13. Databases**                                                      | 48   |
| Create Databases                                                      | 48   |
| Access Resources                                                      | 48   |
| XML Documents                                                          | 48   |
| Raw Files                                                              | 49   |
| HTTP Services                                                         | 49   |
| **Updates**                                                            | 41   |
| ADD                                                                    | 41   |
| DELETE                                                                 | 42   |
| RENAME                                                                 | 42   |
| REPLACE                                                                | 42   |
| STORE                                                                  | 42   |
| OPTIMIZE                                                               | 43   |
| FLUSH                                                                  | 43   |
| **Server Administration**                                             | 43   |
| SHOW SESSIONS                                                          | 43   |
| SHOW USERS                                                             | 43   |
| KILL                                                                   | 43   |
| CREATE EVENT                                                          | 44   |
| SHOW EVENTS                                                            | 44   |
| DROP EVENT                                                             | 44   |
| **User Management**                                                    | 44   |
| CREATE USER                                                            | 44   |
| ALTER USER                                                             | 44   |
| DROP USER                                                              | 45   |
| GRANT                                                                  | 45   |
| PASSWORD                                                               | 45   |
| **General Commands**                                                   | 45   |
| RUN                                                                    | 45   |
| EXECUTE                                                                | 46   |
| GET                                                                    | 46   |
| SET                                                                    | 46   |
| INFO                                                                   | 46   |
| HELP                                                                   | 46   |
| **EXIT**                                                               | 47   |
| **Changelog**                                                          | 47   |

| **12. User Management**                                               | 44   |
| **11. General Commands**                                              | 45   |
| **10. Changelog**                                                      | 47   |
| **9. Update Resources**                                               | 49   |
| **8. Export Data**                                                    | 50   |
| **7. Client/Server Architecture**                                     | 52   |
| **6. HTTP Options**                                                   | 54   |
| **5. Full-Text**                                                      | 59   |
| **4. Indexing**                                                       | 57   |
| **3. XML Parsing**                                                    | 57   |
| **2. Parsing**                                                        | 56   |
| **1. General**                                                        | 55   |
| **REPO INSTALL**                                                      | 41   |
| **REPO LIST**                                                         | 41   |
| **REPO DELETE**                                                       | 41   |
| **Updates**                                                            | 41   |
| **ADD**                                                                | 41   |
| **DELETE**                                                            | 42   |
| **RENAME**                                                            | 42   |
| **REPLACE**                                                           | 42   |
| **STORE**                                                             | 42   |
| **OPTIMIZE**                                                          | 43   |
| **FLUSH**                                                             | 43   |
| **Server Administration**                                             | 43   |
| **SHOW SESSIONS**                                                     | 43   |
| **SHOW USERS**                                                        | 43   |
| **KILL**                                                              | 43   |
| **CREATE EVENT**                                                      | 44   |
| **SHOW EVENTS**                                                       | 44   |
| **DROP EVENT**                                                        | 44   |
| **User Management**                                                   | 44   |
| **CREATE USER**                                                       | 44   |
| **ALTER USER**                                                        | 44   |
| **DROP USER**                                                          | 45   |
| **GRANT**                                                             | 45   |
| **PASSWORD**                                                           | 45   |
| **General Commands**                                                  | 45   |
| **RUN**                                                               | 45   |
| **EXECUTE**                                                           | 46   |
| **GET**                                                               | 46   |
| **SET**                                                               | 46   |
| **INFO**                                                              | 46   |
| **HELP**                                                              | 46   |
| **EXIT**                                                              | 47   |
| **Changelog**                                                         | 47   |
| **13. Databases**                                                     | 48   |
| **Create Databases**                                                  | 48   |
| **Access Resources**                                                  | 48   |
| **XML Documents**                                                     | 48   |
| **Raw Files**                                                         | 49   |
| **HTTP Services**                                                     | 49   |
| **Update Resources**                                                  | 49   |
| **Export Data**                                                       | 50   |
| **Changelog**                                                         | 50   |

<p>| <strong>14. Options</strong>                                                       | 51   |
| <strong>Global Options</strong>                                                    | 51   |
| <strong>General</strong>                                                           | 51   |
| <strong>Client/Server Architecture</strong>                                       | 52   |
| <strong>HTTP Options</strong>                                                      | 54   |
| <strong>Create Options</strong>                                                   | 55   |
| <strong>General</strong>                                                           | 55   |
| <strong>Parsing</strong>                                                           | 56   |
| <strong>XML Parsing</strong>                                                       | 57   |
| <strong>Indexing</strong>                                                          | 57   |
| <strong>Full-Text</strong>                                                         | 59   |
| <strong>Query Options</strong>                                                     | 59   |
| <strong>QUERYINFO</strong>                                                         | 59   |
| <strong>XQUERY3</strong>                                                            | 60   |
| <strong>BINDINGS</strong>                                                          | 60   |
| <strong>QUERYPATH</strong>                                                         | 60   |</p>
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>19.</td>
<td>Full-Text: Japanese</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>Lexical Analysis</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>Parsing</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>Token Processing</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>Stemming</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>Wildcards</td>
<td>83</td>
</tr>
<tr>
<td>20.</td>
<td>Higher-Order Functions</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>Function Items</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>Function Types</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>Higher-Order Functions</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>Higher-Order Functions on Sequences</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>Folds</td>
<td>87</td>
</tr>
<tr>
<td>21.</td>
<td>Java Bindings</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>Namespace Declarations</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>Module Imports</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>Context-Awareness</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>Changelog</td>
<td>93</td>
</tr>
<tr>
<td>22.</td>
<td>Module Library</td>
<td>94</td>
</tr>
<tr>
<td>23.</td>
<td>Repository</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>Importing Modules</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>Commands</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>Packaging</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>EXPath Packaging</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>URI Rewriting</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Changelog</td>
<td>101</td>
</tr>
<tr>
<td>24.</td>
<td>Serialization</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>Parameters</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>Changelog</td>
<td>104</td>
</tr>
<tr>
<td>25.</td>
<td>XQuery</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>XQuery 3.0</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>Module Library</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>Repository</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>Java Bindings</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>Updates</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>Serialization</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>Errors</td>
<td>106</td>
</tr>
<tr>
<td>26.</td>
<td>XQuery 3.0</td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>Enhanced FLWOR Expressions</td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>Simple Map Operator</td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>Group By</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td>Try/Catch</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td>Switch</td>
<td>109</td>
</tr>
<tr>
<td></td>
<td>Function Items</td>
<td>109</td>
</tr>
<tr>
<td></td>
<td>Expanded QNames</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>Namespace Constructors</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>String Concatenations</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>External Variables</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>Serialization</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td>Context Item</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td>Annotations</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td>Functions</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td>Changelog</td>
<td>113</td>
</tr>
<tr>
<td>27.</td>
<td>XQuery Errors</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td>BaseX Errors</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td>Static Errors</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td>Type Errors</td>
<td>116</td>
</tr>
<tr>
<td>Section</td>
<td>Pages</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>Dynamic Errors</td>
<td>117</td>
<td></td>
</tr>
<tr>
<td>Functions Errors</td>
<td>117</td>
<td></td>
</tr>
<tr>
<td>Serialization Errors</td>
<td>119</td>
<td></td>
</tr>
<tr>
<td>Update Errors</td>
<td>119</td>
<td></td>
</tr>
<tr>
<td>Full-Text Errors</td>
<td>121</td>
<td></td>
</tr>
<tr>
<td>VI. XQuery Modules</td>
<td>126</td>
<td></td>
</tr>
<tr>
<td>28. XQuery Update</td>
<td>122</td>
<td></td>
</tr>
<tr>
<td>Features</td>
<td>122</td>
<td></td>
</tr>
<tr>
<td>Updating Expressions</td>
<td>122</td>
<td></td>
</tr>
<tr>
<td>Non-Updating Expressions</td>
<td>123</td>
<td></td>
</tr>
<tr>
<td>Functions</td>
<td>124</td>
<td></td>
</tr>
<tr>
<td>Concepts</td>
<td>124</td>
<td></td>
</tr>
<tr>
<td>Effects</td>
<td>125</td>
<td></td>
</tr>
<tr>
<td>Error Messages</td>
<td>125</td>
<td></td>
</tr>
<tr>
<td>29. Admin Module</td>
<td>127</td>
<td></td>
</tr>
<tr>
<td>Conventions</td>
<td>127</td>
<td></td>
</tr>
<tr>
<td>Functions</td>
<td>127</td>
<td></td>
</tr>
<tr>
<td>admin:users</td>
<td>127</td>
<td></td>
</tr>
<tr>
<td>admin:sessions</td>
<td>127</td>
<td></td>
</tr>
<tr>
<td>admin:logs</td>
<td>127</td>
<td></td>
</tr>
<tr>
<td>Changelog</td>
<td>127</td>
<td></td>
</tr>
<tr>
<td>30. Archive Module</td>
<td>128</td>
<td></td>
</tr>
<tr>
<td>Conventions</td>
<td>128</td>
<td></td>
</tr>
<tr>
<td>Functions</td>
<td>128</td>
<td></td>
</tr>
<tr>
<td>archive:create</td>
<td>128</td>
<td></td>
</tr>
<tr>
<td>archive:entries</td>
<td>129</td>
<td></td>
</tr>
<tr>
<td>archive:options</td>
<td>129</td>
<td></td>
</tr>
<tr>
<td>archive:extract-text</td>
<td>130</td>
<td></td>
</tr>
<tr>
<td>archive:extract-binary</td>
<td>130</td>
<td></td>
</tr>
<tr>
<td>archive:update</td>
<td>130</td>
<td></td>
</tr>
<tr>
<td>archive:delete</td>
<td>131</td>
<td></td>
</tr>
<tr>
<td>archive:write</td>
<td>131</td>
<td></td>
</tr>
<tr>
<td>Errors</td>
<td>131</td>
<td></td>
</tr>
<tr>
<td>Changelog</td>
<td>132</td>
<td></td>
</tr>
<tr>
<td>31. CSV Module</td>
<td>133</td>
<td></td>
</tr>
<tr>
<td>Conventions</td>
<td>133</td>
<td></td>
</tr>
<tr>
<td>Rules</td>
<td>133</td>
<td></td>
</tr>
<tr>
<td>Functions</td>
<td>133</td>
<td></td>
</tr>
<tr>
<td>csv:parse</td>
<td>133</td>
<td></td>
</tr>
<tr>
<td>csv:serialize</td>
<td>134</td>
<td></td>
</tr>
<tr>
<td>Errors</td>
<td>134</td>
<td></td>
</tr>
<tr>
<td>Changelog</td>
<td>134</td>
<td></td>
</tr>
<tr>
<td>32. Client Module</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td>Conventions</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td>Functions</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td>client:connect</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td>client:execute</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td>client:info</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td>client:query</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td>client:close</td>
<td>136</td>
<td></td>
</tr>
<tr>
<td>Errors</td>
<td>136</td>
<td></td>
</tr>
<tr>
<td>Changelog</td>
<td>136</td>
<td></td>
</tr>
<tr>
<td>33. Conversion Module</td>
<td>137</td>
<td></td>
</tr>
<tr>
<td>Conventions</td>
<td>137</td>
<td></td>
</tr>
<tr>
<td>Strings</td>
<td>137</td>
<td></td>
</tr>
<tr>
<td>convert:binary-to-string</td>
<td>137</td>
<td></td>
</tr>
<tr>
<td>convert:string-to-base64</td>
<td>137</td>
<td></td>
</tr>
<tr>
<td>convert:string-to-hex</td>
<td>137</td>
<td></td>
</tr>
<tr>
<td>Topic</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>Binary Data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>convert:bytes-to-base64</td>
<td>138</td>
<td></td>
</tr>
<tr>
<td>convert:bytes-to-hex</td>
<td>138</td>
<td></td>
</tr>
<tr>
<td>convert:binary-to-bytes</td>
<td>138</td>
<td></td>
</tr>
<tr>
<td>Numbers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>convert:integer-to-base</td>
<td>138</td>
<td></td>
</tr>
<tr>
<td>convert:integer-from-base</td>
<td>138</td>
<td></td>
</tr>
<tr>
<td>Dates and Durations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>convert:integer-to-datetime</td>
<td>139</td>
<td></td>
</tr>
<tr>
<td>convert:dateTime-to-integer</td>
<td>139</td>
<td></td>
</tr>
<tr>
<td>convert:integer-to-dayTime</td>
<td>139</td>
<td></td>
</tr>
<tr>
<td>convert:dayTime-to-integer</td>
<td>139</td>
<td></td>
</tr>
<tr>
<td>Errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Changelog</td>
<td>139</td>
<td></td>
</tr>
<tr>
<td>34. Cryptographic Module</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conventions</td>
<td>141</td>
<td></td>
</tr>
<tr>
<td>Message Authentication</td>
<td>141</td>
<td></td>
</tr>
<tr>
<td>crypto:hmac</td>
<td>141</td>
<td></td>
</tr>
<tr>
<td>Encryption &amp; Decryption</td>
<td></td>
<td></td>
</tr>
<tr>
<td>crypto:encrypt</td>
<td>142</td>
<td></td>
</tr>
<tr>
<td>crypto:decrypt</td>
<td>142</td>
<td></td>
</tr>
<tr>
<td>XML Signatures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>crypto:generate-signature</td>
<td>143</td>
<td></td>
</tr>
<tr>
<td>crypto:validate-signature</td>
<td>144</td>
<td></td>
</tr>
<tr>
<td>Errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Changelog</td>
<td>145</td>
<td></td>
</tr>
<tr>
<td>35. Database Module</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conventions</td>
<td>146</td>
<td></td>
</tr>
<tr>
<td>Database Nodes</td>
<td>146</td>
<td></td>
</tr>
<tr>
<td>General Functions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>db:system</td>
<td>146</td>
<td></td>
</tr>
<tr>
<td>db:info</td>
<td>146</td>
<td></td>
</tr>
<tr>
<td>db:list</td>
<td>146</td>
<td></td>
</tr>
<tr>
<td>db:list-details</td>
<td>146</td>
<td></td>
</tr>
<tr>
<td>db:backups</td>
<td>147</td>
<td></td>
</tr>
<tr>
<td>db:event</td>
<td>147</td>
<td></td>
</tr>
<tr>
<td>Read Operations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>db:open</td>
<td>147</td>
<td></td>
</tr>
<tr>
<td>db:open-pre</td>
<td>147</td>
<td></td>
</tr>
<tr>
<td>db:open-id</td>
<td>147</td>
<td></td>
</tr>
<tr>
<td>db:node-pre</td>
<td>148</td>
<td></td>
</tr>
<tr>
<td>db:node-id</td>
<td>148</td>
<td></td>
</tr>
<tr>
<td>db:retrieve</td>
<td>148</td>
<td></td>
</tr>
<tr>
<td>db:export</td>
<td>148</td>
<td></td>
</tr>
<tr>
<td>Contents</td>
<td></td>
<td></td>
</tr>
<tr>
<td>db:text</td>
<td>149</td>
<td></td>
</tr>
<tr>
<td>db:text-range</td>
<td>149</td>
<td></td>
</tr>
<tr>
<td>db:attribute</td>
<td>149</td>
<td></td>
</tr>
<tr>
<td>db:attribute-range</td>
<td>149</td>
<td></td>
</tr>
<tr>
<td>db:fulltext</td>
<td>150</td>
<td></td>
</tr>
<tr>
<td>Updates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>db:create</td>
<td>150</td>
<td></td>
</tr>
<tr>
<td>db:drop</td>
<td>151</td>
<td></td>
</tr>
<tr>
<td>db:add</td>
<td>151</td>
<td></td>
</tr>
<tr>
<td>db:delete</td>
<td>152</td>
<td></td>
</tr>
<tr>
<td>db:optimize</td>
<td>152</td>
<td></td>
</tr>
<tr>
<td>db:rename</td>
<td>152</td>
<td></td>
</tr>
<tr>
<td>db:replace</td>
<td>152</td>
<td></td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>36. Fetch Module</td>
<td>156</td>
<td></td>
</tr>
<tr>
<td>Conventions</td>
<td>156</td>
<td></td>
</tr>
<tr>
<td>Functions</td>
<td>156</td>
<td></td>
</tr>
<tr>
<td>fetch:binary</td>
<td>156</td>
<td></td>
</tr>
<tr>
<td>fetch:content-type</td>
<td>156</td>
<td></td>
</tr>
<tr>
<td>Errors</td>
<td>157</td>
<td></td>
</tr>
<tr>
<td>Changelog</td>
<td>157</td>
<td></td>
</tr>
<tr>
<td>37. File Module</td>
<td>158</td>
<td></td>
</tr>
<tr>
<td>Conventions</td>
<td>158</td>
<td></td>
</tr>
<tr>
<td>Read Operations</td>
<td>158</td>
<td></td>
</tr>
<tr>
<td>file:list</td>
<td>158</td>
<td></td>
</tr>
<tr>
<td>file:read-binary</td>
<td>158</td>
<td></td>
</tr>
<tr>
<td>file:read-text</td>
<td>158</td>
<td></td>
</tr>
<tr>
<td>file:read-text-lines</td>
<td>159</td>
<td></td>
</tr>
<tr>
<td>Write Operations</td>
<td>159</td>
<td></td>
</tr>
<tr>
<td>file:create-dir</td>
<td>159</td>
<td></td>
</tr>
<tr>
<td>file:create-temp-dir</td>
<td>159</td>
<td></td>
</tr>
<tr>
<td>file:create-temp-file</td>
<td>159</td>
<td></td>
</tr>
<tr>
<td>file:delete</td>
<td>159</td>
<td></td>
</tr>
<tr>
<td>file:write</td>
<td>160</td>
<td></td>
</tr>
<tr>
<td>file:write-binary</td>
<td>160</td>
<td></td>
</tr>
<tr>
<td>file:write-text</td>
<td>160</td>
<td></td>
</tr>
<tr>
<td>file:write-text-lines</td>
<td>160</td>
<td></td>
</tr>
<tr>
<td>file:append</td>
<td>161</td>
<td></td>
</tr>
<tr>
<td>file:append-binary</td>
<td>161</td>
<td></td>
</tr>
<tr>
<td>file:append-text</td>
<td>161</td>
<td></td>
</tr>
<tr>
<td>file:append-text-lines</td>
<td>161</td>
<td></td>
</tr>
<tr>
<td>file:copy</td>
<td>161</td>
<td></td>
</tr>
<tr>
<td>file:move</td>
<td>162</td>
<td></td>
</tr>
<tr>
<td>File Properties</td>
<td>162</td>
<td></td>
</tr>
<tr>
<td>file:exists</td>
<td>162</td>
<td></td>
</tr>
<tr>
<td>file:is-dir</td>
<td>162</td>
<td></td>
</tr>
<tr>
<td>file:is-file</td>
<td>162</td>
<td></td>
</tr>
<tr>
<td>file:last-modified</td>
<td>162</td>
<td></td>
</tr>
<tr>
<td>file:size</td>
<td>162</td>
<td></td>
</tr>
<tr>
<td>Path Functions</td>
<td>162</td>
<td></td>
</tr>
<tr>
<td>file:base-name</td>
<td>162</td>
<td></td>
</tr>
<tr>
<td>file:dir-name</td>
<td>162</td>
<td></td>
</tr>
<tr>
<td>file:path-to-native</td>
<td>163</td>
<td></td>
</tr>
<tr>
<td>file:resolve-path</td>
<td>163</td>
<td></td>
</tr>
<tr>
<td>file:path-to-uri</td>
<td>163</td>
<td></td>
</tr>
<tr>
<td>System Properties</td>
<td>163</td>
<td></td>
</tr>
<tr>
<td>file:dir-separator</td>
<td>163</td>
<td></td>
</tr>
<tr>
<td>file:path-separator</td>
<td>163</td>
<td></td>
</tr>
<tr>
<td>file:line-separator</td>
<td>163</td>
<td></td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>Errors</td>
<td>164</td>
<td></td>
</tr>
<tr>
<td>Changelog</td>
<td>164</td>
<td></td>
</tr>
<tr>
<td>38. Full-Text Module</td>
<td>165</td>
<td></td>
</tr>
<tr>
<td>Conventions</td>
<td>165</td>
<td></td>
</tr>
<tr>
<td>Functions</td>
<td>165</td>
<td></td>
</tr>
<tr>
<td>ft:search</td>
<td>165</td>
<td></td>
</tr>
<tr>
<td>ft:mark</td>
<td>166</td>
<td></td>
</tr>
<tr>
<td>ft:extract</td>
<td>166</td>
<td></td>
</tr>
<tr>
<td>ft:count</td>
<td>167</td>
<td></td>
</tr>
<tr>
<td>ft:score</td>
<td>167</td>
<td></td>
</tr>
<tr>
<td>ft:tokens</td>
<td>167</td>
<td></td>
</tr>
<tr>
<td>ft:tokenize</td>
<td>167</td>
<td></td>
</tr>
<tr>
<td>Errors</td>
<td>167</td>
<td></td>
</tr>
<tr>
<td>Changelog</td>
<td>167</td>
<td></td>
</tr>
<tr>
<td>39. Geo Module</td>
<td>169</td>
<td></td>
</tr>
<tr>
<td>Conventions</td>
<td>169</td>
<td></td>
</tr>
<tr>
<td>General Functions</td>
<td>169</td>
<td></td>
</tr>
<tr>
<td>geo:dimension</td>
<td>169</td>
<td></td>
</tr>
<tr>
<td>geo:geometry-type</td>
<td>169</td>
<td></td>
</tr>
<tr>
<td>geo:srid</td>
<td>170</td>
<td></td>
</tr>
<tr>
<td>geo:envelope</td>
<td>170</td>
<td></td>
</tr>
<tr>
<td>geo:as-text</td>
<td>170</td>
<td></td>
</tr>
<tr>
<td>geo:as-binary</td>
<td>171</td>
<td></td>
</tr>
<tr>
<td>geo:is-simple</td>
<td>171</td>
<td></td>
</tr>
<tr>
<td>geo:boundary</td>
<td>171</td>
<td></td>
</tr>
<tr>
<td>geo:num-geometries</td>
<td>172</td>
<td></td>
</tr>
<tr>
<td>geo:geometry-n</td>
<td>173</td>
<td></td>
</tr>
<tr>
<td>geo:length</td>
<td>174</td>
<td></td>
</tr>
<tr>
<td>geo:num-points</td>
<td>174</td>
<td></td>
</tr>
<tr>
<td>geo:area</td>
<td>175</td>
<td></td>
</tr>
<tr>
<td>geo:centroid</td>
<td>175</td>
<td></td>
</tr>
<tr>
<td>geo:point-on-surface</td>
<td>176</td>
<td></td>
</tr>
<tr>
<td>Spatial Predicate Functions</td>
<td>177</td>
<td></td>
</tr>
<tr>
<td>geo:equals</td>
<td>177</td>
<td></td>
</tr>
<tr>
<td>geo:disjoint</td>
<td>177</td>
<td></td>
</tr>
<tr>
<td>geo:intersects</td>
<td>178</td>
<td></td>
</tr>
<tr>
<td>geo:touches</td>
<td>178</td>
<td></td>
</tr>
<tr>
<td>geo:crosses</td>
<td>179</td>
<td></td>
</tr>
<tr>
<td>geo:within</td>
<td>179</td>
<td></td>
</tr>
<tr>
<td>geo:contains</td>
<td>180</td>
<td></td>
</tr>
<tr>
<td>geo:overlaps</td>
<td>180</td>
<td></td>
</tr>
<tr>
<td>geo:relate</td>
<td>181</td>
<td></td>
</tr>
<tr>
<td>Analysis Functions</td>
<td>182</td>
<td></td>
</tr>
<tr>
<td>geo:distance</td>
<td>182</td>
<td></td>
</tr>
<tr>
<td>geo:buffer</td>
<td>182</td>
<td></td>
</tr>
<tr>
<td>geo:convex-hull</td>
<td>183</td>
<td></td>
</tr>
<tr>
<td>geo:intersection</td>
<td>183</td>
<td></td>
</tr>
<tr>
<td>geo:union</td>
<td>184</td>
<td></td>
</tr>
<tr>
<td>geo:difference</td>
<td>184</td>
<td></td>
</tr>
<tr>
<td>geo:sym-difference</td>
<td>185</td>
<td></td>
</tr>
<tr>
<td>Functions Specific to Geometry Type</td>
<td>185</td>
<td></td>
</tr>
<tr>
<td>geo:x</td>
<td>185</td>
<td></td>
</tr>
<tr>
<td>geo:y</td>
<td>186</td>
<td></td>
</tr>
<tr>
<td>geo:z</td>
<td>186</td>
<td></td>
</tr>
<tr>
<td>geo:start-point</td>
<td>187</td>
<td></td>
</tr>
<tr>
<td>geo:end-point</td>
<td>187</td>
<td></td>
</tr>
<tr>
<td>geo:is-closed</td>
<td>187</td>
<td></td>
</tr>
</tbody>
</table>
40. HTML Module ................................................................. 192
   Conventions ............................................................... 192
   Functions ........................................................................ 192
      html:parser ............................................................... 192
      html:parse ............................................................... 192
   Examples ......................................................................... 192
      Basic Example ............................................................ 192
      Specifying Options .................................................... 193
      Parsing Binary Input .................................................. 193
   Errors ............................................................................ 193
   Changelog .................................................................... 193
41. HTTP Module ................................................................. 194
   Conventions ................................................................. 194
   Functions ........................................................................ 194
      http:send-request ...................................................... 194
   Examples ......................................................................... 194
   Errors ............................................................................ 197
   Changelog .................................................................... 197
42. Hashing Module ............................................................... 198
   Conventions ................................................................. 198
   Functions ........................................................................ 198
      hash:md5 .................................................................. 198
      hash:sha1 .................................................................. 198
      hash:sha256 .............................................................. 198
      hash:hash ................................................................. 198
   Errors ............................................................................ 199
   Changelog .................................................................... 199
43. Higher-Order Functions Module ......................................... 200
   Conventions ................................................................. 200
   Functions ........................................................................ 200
      hof:id ....................................................................... 200
      hof:const ................................................................. 200
      hof:fold-left1 ........................................................... 201
      hof:until ................................................................. 201
      hof:top-k-by ............................................................. 202
      hof:top-k-with ........................................................... 202
   Changelog .................................................................... 202
44. Index Module .................................................................. 203
   Conventions ................................................................. 203
   Functions ........................................................................ 203
      index:facets ............................................................... 203
      index:texts ............................................................... 203
      index:attributes ......................................................... 203
      index:element-names ................................................. 204
      index:attribute-names ................................................. 204
   Changelog .................................................................... 204
45. Inspection Module ........................................................... 205
   Conventions ................................................................. 205
   Reflection ....................................................................... 205
      inspect:functions ....................................................... 205
   Documentation ............................................................ 205
Responses ................................................................. 292
  Custom Responses ........................................ 292
  Forwards and Redirects ..................................... 292
  Output ......................................................... 293
Error Handling ..................................................... 294
  XQuery Errors ............................................... 294
  HTTP Errors .................................................. 295
Functions .......................................................... 295
References .......................................................... 295
Changelog .......................................................... 295
74. WebDAV ......................................................... 297
  Usage ......................................................... 297
  Authorization ............................................... 297
  Locking ...................................................... 297
  WebDAV Clients ............................................. 297
  Changelog .................................................... 297
75. WebDAV: GNOME .............................................. 299
76. WebDAV: KDE .................................................. 301
77. WebDAV: Mac OSX .......................................... 303
78. WebDAV: Windows 7 ......................................... 306
79. WebDAV: Windows XP ....................................... 308

IX. Client APIs ................................................... 312
80. Clients ......................................................... 313
81. Java Examples ............................................... 315
  Local Examples ............................................... 315
    Server Examples ......................................... 315
    XQuery Module Examples ................................ 315
    Client API ................................................. 316
    REST API .................................................. 316
    XML:DB API (deprecated) .................................. 316
82. PHP Example .................................................. 317
  Requirements ............................................... 317
    Setting up ................................................ 317
    Usage ...................................................... 317
83. Query Mode ................................................... 318
  Usage ......................................................... 318
    PHP Example ................................................ 318
    Changelog .................................................. 319
84. Server Protocol ............................................... 320
  Workflow .................................................... 320
    Constructors and Functions ............................... 320
    Example ................................................... 323
    Existing Clients .......................................... 323
    Changelog .................................................. 323
85. Server Protocol: Types ...................................... 325
    XDM Meta Data ............................................ 325
    Type IDs .................................................. 325
86. Standard Mode ............................................... 327
  Usage ......................................................... 327
    Example in PHP ........................................... 327
X. Advanced User’s Guide ........................................ 328
87. Advanced User’s Guide ....................................... 329
  Storage ....................................................... 329
  Use Cases .................................................... 329
88. Backups ......................................................... 330
  GUI Example .................................................. 330
  Console Example ............................................. 330
89. Catalog Resolver ............................................. 331
Chapter 1. Main Page

Read this entry online in the BaseX Wiki.

Welcome to the documentation of BaseX!

BaseX is both a light-weight, high-performance and scalable XML Database and an XQuery 3.0 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 queries in realtime (i.e., with each key click). BaseX is platform-independent and distributed under the free BSD License (find more in Wikipedia).

This documentation is based on BaseX 7.7.2. It can also be downloaded as PDF. Features that have recently been added or changed are highlighted and can be searched for.

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.

Categories: Beginners

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.

Categories: XQuery
**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.

Categories: Developer, HTTP, API

**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.

Categories: Internals

You are invited to contribute to our Wiki: it's easy to get a new account. If you have questions and are looking for direct contact to developers and users, please write to our basex-talk mailing list.
Part I. Getting Started
Chapter 2. Command-Line Options

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.

BaseX 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 [−bcdiLosuvVwxz] [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 initial file or database
−L Append newlines to query results
−o<output> Write output to file
−q<expr> Execute XQuery expression
−s<pars> Set serialization parameter(s)
−u Write updates back to original files
−v/V Show (all) process info
−w Preserve whitespaces from input files
−x Show query execution plan
−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.

<table>
<thead>
<tr>
<th>Flag</th>
<th>Description</th>
<th>Option</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>[input]</td>
<td>Evaluates the specified input:</td>
<td></td>
<td>• &quot;doc('X')//head&quot;* query.xq* commands.bxs</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>
</tr>
<tr>
<td></td>
<td>the file content will be evaluated as Command Script; otherwise, it will</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>be evaluated as XQuery expression.</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>
</tr>
<tr>
<td>−b&lt;pars&gt;</td>
<td>Binds external variables to XQuery expressions. This flag may be specified</td>
<td>BINDINGS</td>
<td>• −b$v=example &quot;$v&quot;* −b(URL)ln=value&quot;declare</td>
</tr>
<tr>
<td></td>
<td>multiple times. Variables names and their</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>value must be enclosed in double quotes.</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.

\[ \text{namespace ns='URL'; $ns:ln} \]

-\c<input> Executes Commands:
- Several commands in the input can be separated by semicolons (\;).
- 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.

-\d Toggles the debugging mode. Debugging information is output to standard error.

-\i<input> Opens a database or XML document specified by the argument. The opened input may be further processed by an XQuery expression.

-\L Separates returned query items by newlines (instead of spaces) and appends a newline to the end of a result.

-\o<file> All command and query output is written to the specified file.

-\q<expr> Executes the specified string as XQuery expression.

-\s<pars> 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 (=).

-\u Modifies original files after evaluating XQuery Update expressions.

-\v Prints process and timing information to the standard output.

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

-\w By default, whitespaces around text nodes are chopped when a database is created. This flag can be specified to preserve whitespaces.

-\x This flags turn on the output of the query execution plan, formatted in XML.

-\z Skips the serialization of XQuery results. This flag is useful if the query is profiled or analyzed.

**BaseX 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 [-cdeinpSz] [stop]
  stop  Stop running server
  -c<cmds>  Execute initial database commands
  -d  Activate debugging mode
  -e<port>  Set event port
  -i  Enter interactive mode
  -n<name>  Set host the server is bound to
  -p<port>  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 an existing 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>DEBUG</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 standard error.</td>
<td></td>
<td>EVENTPORT</td>
<td>1985 -e9998</td>
</tr>
<tr>
<td>-e&lt;num&gt;</td>
<td>Specifies the port on which the server will send events to clients.</td>
<td></td>
<td>SERVERPORT</td>
<td>1984 -p9999</td>
</tr>
<tr>
<td>-i</td>
<td>Starts the interactive console mode, which can be used to enter database commands. This mode is similar to the default standalone and client mode.</td>
<td></td>
<td>SERVERHOST</td>
<td>-p127.0.0.1</td>
</tr>
<tr>
<td>-n&lt;name&gt;</td>
<td>Specifies the host the server will be bound to. Version 7.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-p&lt;num&gt;</td>
<td>Specifies the port on which the server will be addressable.</td>
<td></td>
<td></td>
<td></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></td>
<td>LOG</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.

**BaseX 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 

$ basexclient -h
BaseX [Client]
Usage: basexclient [-bcdiLnopPsUvVwxz] [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 initial file or database
-L Append newlines to query results
-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
-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
-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>• 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.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Otherwise, the input string itself is evaluated as XQuery expression.</td>
</tr>
<tr>
<td>-b&lt;pars&gt;</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 (=). 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.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-c&lt;input&gt;</td>
<td>Executes Commands:</td>
<td>BINDINGS</td>
<td></td>
<td>• Several commands in the input can be separated by semicolons (;).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• If the specified input is a valid file reference or URL, its content will be stored as Command Script.</td>
</tr>
</tbody>
</table>

• "doc('X')//head"* query.xq* commands.bxs
• $v=example "$v"* - b(URL)ln=value"declare namespace ns='URL'; $ns:ln"
• c"list;info"* - ccommands.txt* - c"<info/>"
### Command-Line Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>-d</code></td>
<td>Toggles the debugging mode. Debugging information is output to standard error.</td>
<td>DEBUG</td>
</tr>
<tr>
<td><code>-i&lt;input&gt;</code></td>
<td>Opens a database or XML document specified by the argument. The opened input may be further processed by an XQuery expression.</td>
<td>-iitems.xml &quot;//item&quot;</td>
</tr>
<tr>
<td><code>-L</code></td>
<td>Separates returned query items by newlines (instead of spaces) and appends a newline to the end of a result.</td>
<td></td>
</tr>
<tr>
<td><code>-n&lt;name&gt;</code></td>
<td>Specifies the host name on which the server is running.</td>
<td>HOST localhost -nserver.basex.org</td>
</tr>
<tr>
<td><code>-o&lt;file&gt;</code></td>
<td>All command and query output is written to the specified file.</td>
<td></td>
</tr>
<tr>
<td><code>-p&lt;num&gt;</code></td>
<td>Specifies the port on which the server is running.</td>
<td>PORT 1984 -p9999</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>
<td>PASSWORD -Uadmin -Padmin</td>
</tr>
<tr>
<td><code>-q&lt;expr&gt;</code></td>
<td>Executes the specified string as XQuery expression.</td>
<td>-q&quot;doc('input')//head&quot;</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>
<td>SERIALIZER -smethod=text</td>
</tr>
<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 -Uadmin</td>
</tr>
<tr>
<td><code>-v</code></td>
<td>Prints process and timing information to the standard output.</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>
</tr>
<tr>
<td><code>-w</code></td>
<td>By default, whitespaces around text nodes are chopped when a database is created. This flag can be specified to preserve whitespaces.</td>
<td>CHOP</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>
</tr>
<tr>
<td><code>-z</code></td>
<td>Skips the serialization of XQuery results. This flag is useful if the query is profiled or analyzed.</td>
<td>SERIALIZE</td>
</tr>
</tbody>
</table>

### BaseX HTTP Server

Launch the HTTP server
Command-Line Options

$ 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 [-dehlnPpPRUWz] [stop]
    stop      Stop running server
    -d        Activate debugging mode
    -e<port>  Set event port
    -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 running HTTP server. By default, the database server will be stopped as well, unless -l has been specified.</td>
<td>pom.xml</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-d</td>
<td>Turns on the debugging mode. Debugging information is output to standard error.</td>
<td>DEBUG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-e&lt;num&gt;</td>
<td>Specifies the port on which the server will send events to clients.</td>
<td>EVENTPORT</td>
<td>1985</td>
<td>-e9998</td>
</tr>
<tr>
<td>-h&lt;num&gt;</td>
<td>Specifies the port on which the HTTP server will be addressable.</td>
<td>jetty.xml</td>
<td>8984</td>
<td>-h9999</td>
</tr>
<tr>
<td>-l</td>
<td>Starts the server in local mode, and executes all commands in the embedded database context.</td>
<td>HTTPLOCAL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-n&lt;name&gt;</td>
<td>Specifies the host name on which the server is running.</td>
<td>HOST</td>
<td>localhost</td>
<td>nserver.basex.org</td>
</tr>
<tr>
<td>-p&lt;num&gt;</td>
<td>Specifies the port on which the database server will be addressable.</td>
<td>SERVERPORT</td>
<td>1984</td>
<td>-p9998</td>
</tr>
<tr>
<td>-P&lt;pass&gt;</td>
<td>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.</td>
<td>PASSWORD</td>
<td></td>
<td>-Uadmin -Padmin</td>
</tr>
<tr>
<td>-s&lt;num&gt;</td>
<td>Specifies the port that will be used to stop the HTTP server.</td>
<td>STOPPORT</td>
<td>8983</td>
<td></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>
</tbody>
</table>
Command-Line Options

- **-U<name>** Specifies a user name, which will be used by the HTTP services for opening a new session.
- **-z** Does not generate any log files.

**BaseX 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 7.7
- Added: `-n` binds a host to the BaseX Server.

Version 7.5
- Added: detection of Command Scripts.

Version 7.3
- Updated: all options are now evaluated in the given order
- Updated: `-i` creates main-memory representations for specified sources
- 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 3. 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.

Getting Started

• 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

Tutorials and Slides

• XMLPrague 2013, Slides and Examples

• Neven Jovanović, BaseX Adventures

• W3 Schools, XQuery Tutorial

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
Chapter 4. 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 from the etc directory of our GitHub basex and basex-api repositories.

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

```batch
@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

```bash
#!/bin/bash

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

# 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: basehttp.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: basehttp

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

# 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
REM 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.
Chapter 5. 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. The application can be run and used in various ways:

- as a standalone application, using the Graphical User Interface or the Command-Line Interface
- as a client/server application
- as a Web Application, either standalone or based on the client/server architecture

Important: If you plan to perform concurrent read and write operations on a single database, the client/server architecture is the right choice. 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.

Requirements

BaseX

Please download the latest BaseX version from our homepage. The official releases include the BaseX JAR file, libraries and optional Start Scripts. If you do not use an installer, we recommend to manually add the project’s bin directory to your path environment; this way, you will be able to run BaseX from everywhere in your shell/terminal.

Java

A Runtime Environment of Java 1.6 (JRE) is needed to run BaseX. BaseX is platform independent and runs on any system that provides a Java Virtual Machine. BaseX has been tested on Windows (2000, XP, Vista, 7), Max OS X (10.x), Linux(SuSE xxx, Debian, Ubuntu) and OpenBSD (4.x).

BaseX GUI

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.

The GUI 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.

BaseX 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.

The standalone version 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.

Note that the standalone mode does *not* interact with the client/server architecture.

**BaseX 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.

**BaseX 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.

Please check out the article on the Database Server for more details.

**BaseX HTTP Server**

The HTTP Server gives access to the REST, RESTXQ and WebDAV Services of BaseX. By default, it starts an instance of the Jetty Web Server, which by default listens to the port 8984, and the BaseX Server, which listens to 1984.

To run the HTTP Server, you need to download one of the full distributions of BaseX (exe, zip, war), as the JAR version does not include any additionally required libraries. It can then 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.

• You can also deploy BaseX as a [Web Application](#).

• For Maven users: type in `mvn jetty:run` in the main directory of the `basex-api` project, and press `Ctrl + c` to shut down the process (see [Web Application: Maven](#) for more details).

## Changelog

Version 7.0

• Updated: BaseXJAXRX has been replaced with BaseXHTTP
Part II. User Interfaces
Chapter 6. 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 of all, please 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).

Create a database

• To create a database you need an XML document, e.g. factbook.xml.
• Save this document to the directory you are working in.
• In the client terminal, type in:

> CREATE DB factbook factbook.xml

  factbook is the name of the database

  factbook.xml is the xml file, which is used to create the database

If everything works you see the following lines:

Database 'factbook' created in 950.83 ms.

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 have used the Windows Installer, 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

Create a new database

Now we will create another database from the xmark.xml document.

• Create the new database, named 'xmark'.

> CREATE DB xmark xmark.xml
• Set the new database xmark as the context:
  > OPEN xmark

• Now you can easily execute queries on your new database:
  > XQUERY //people/person/name

Switch the database

• You can explicitly query the factbook database with the doc(...) function, no matter what the current context is.
  > XQUERY doc("factbook")//country

• Otherwise, to set factbook as the current context, execute the following:
  > OPEN factbook

• To list the current context, type:
  > SHOW DATABASES

That yields the following lines:

1 opened database(s):
  - factbook (1x)

Close or delete a database

• To close the current context database, please type:
  > CLOSE

• Use the DROP DB command to delete the xmark database:
  > DROP DB xmark

Create a collection

What is a collection? With BaseX you can group documents into one logical collection. A collection is a database that contains two or more documents. Collections accept any type of XML documents, regardless of their structure.

Let's add the xmark.xml document to the factbook database to create a collection. The name of the original factbook database remains.

• First make sure factbook is opened:
  > OPEN factbook

• Now add the xmark.xml document:
  > ADD xmark.xml

Delete a document

• Deleting a document from a collection is easy:
  > DELETE xmark.xml

Make sure that the collection, which contains the xmark.xml document, is opened.
Delete a collection

Deleting a collection is the same as deleting a database.

• To delete the collection factbook, type:
  > DROP DB factbook

Get server information

Several commands help to explore the state of a server. For a complete list, please visit the Commands Section.

• To see all databases on the server, type:
  > LIST

• To see which database is currently opened:
  > SHOW DATABASES

• To see the general information of the opened database, type:
  > INFO

• To see the users in BaseX, type:
  > SHOW USERS

Backup and restore

• To backup your database, type:
  > CREATE BACKUP factbook

• To restore your database, type:
  > RESTORE factbook

Where is the backup-file stored?

The backup-file is stored in the database directory. The file is named `factbook-timestamp.zip` (db_name-timestamp.zip). To restore the database the file with the newest timestamp is taken.

See also

Standalone Mode, GUI, Getting Started, Advanced Usage
Chapter 7. 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 the introductory video presents some of the interactive features that the BaseX GUI provides.

Startup

First of all, please 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 specified in the above 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.

Querying

Keyword Search

The Keyword Search can be executed in the Search mode in the combo box of the main window. This option allows for a simple, keyword-based search in the opened database.

The following syntax is supported:

<table>
<thead>
<tr>
<th>Query</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>world</td>
<td>Find tags and texts containing world</td>
</tr>
<tr>
<td>=world</td>
<td>Find exact matching text nodes</td>
</tr>
<tr>
<td>~world</td>
<td>Find text nodes similar to world</td>
</tr>
<tr>
<td>@world</td>
<td>Find attributes and attribute values</td>
</tr>
<tr>
<td>@=world</td>
<td>Find exact attribute values</td>
</tr>
<tr>
<td>&quot;united world&quot;</td>
<td>Find tags and texts containing the phrase &quot;united world&quot;</td>
</tr>
</tbody>
</table>

XPath/XQuery

Apart from the basic search facilities, BaseX offers far more sophisticated processing options to query your documents. Below are some examples you might give a try. This guide is far from being a comprehensive XQuery reference, but might point you in the right direction.
To execute the following queries, enter them in the XQuery Panel and press ENTER or click on the START button.

XPath provides an easy facility to query your documents in a navigational manner. It is the basic tool of all node-related operations that you encounter when using XQuery. We will start with a trivial example and extend it to our needs.

**Example: Find Countries**

```
//country
```

tells BaseX to look for all country elements in the document. The query is introduced by two slashes `//`, which trigger the traversal of all document nodes. The queries `//country` and `/descendant::country` will return the same results.

**Example: Find Cities in Switzerland**

The following query uses a **predicate** `...` to filter all country nodes which have a `name` child, the string value of which is "Switzerland":

```
//country[name = "Switzerland"]
```

To return all cities of the resulting element node, the query can be extended by a trailing `//city` path:

```
//country[name = "Switzerland"]//city
```

**Text Editor**

The text editor can be used to type in XQuery expressions, Command Scripts, XML documents, or any other text files. Query files and XML documents can be started by clicking on the green triangle. They will automatically parsed with each key click, and errors will be highlighted. Various keyboard shortcuts are available to speed up editing and debugging.

**Visualizations**

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

- **Text View Text**

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

- **Map View Map**

  This visualization represents 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**.
Graphical User Interface

Tree View **Tree**

This visualization 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**.

Folder View **Folder**

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

Scatterplot View **Plot**

This visualization 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**

This visualization 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.

Explorer View **Explorer**

With this visualization you can explore the contents of your database via drop-down menus, search fields and double sliders.

Info View **Info**

This view is helpful for analyzing the query plans of your XQuery expressions. It also displays
information on the compilation and evaluation of queries.

What's Next?

Various tutorials on XPath are available in the internet. We invite you to e.g. have a look at the XQuery Tutorial at W3Schools.
Chapter 8. Shortcuts

Read this entry online in the BaseX Wiki.

This page is part of the Getting Started Section. It gives you an overview of the hotkeys available in the GUI of BaseX.

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>

<table>
<thead>
<tr>
<th>Description</th>
<th>Win/Linux</th>
<th>Mac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Browse back/forward</td>
<td>Alt ←/#Backspace</td>
<td># ←/→</td>
</tr>
<tr>
<td>Browse one level up</td>
<td>Alt ↑</td>
<td># ↑</td>
</tr>
<tr>
<td>Browse to the root node</td>
<td>Alt Home</td>
<td># Home</td>
</tr>
</tbody>
</table>

Editor Shortcuts

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

Processing

<table>
<thead>
<tr>
<th>Description</th>
<th>Win/Linux</th>
<th>Mac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Execute query</td>
<td>Ctrl Enter</td>
<td># Enter</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>Jump to highlighted error</td>
<td>Ctrl .</td>
<td># .</td>
</tr>
</tbody>
</table>

Finding

<table>
<thead>
<tr>
<th>Description</th>
<th>Win/Linux</th>
<th>Mac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Find text</td>
<td>Ctrl F</td>
<td># F</td>
</tr>
<tr>
<td>Find next instance of text</td>
<td>Ctrl GF3</td>
<td># G# F3</td>
</tr>
<tr>
<td>Find previous instance of text</td>
<td>Ctrl Shift GShift F3</td>
<td># Shift G# Shift F3</td>
</tr>
<tr>
<td>Go to line</td>
<td>Ctrl L</td>
<td># L</td>
</tr>
</tbody>
</table>

Navigating

<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>Description</td>
<td>Win/Linux</td>
<td>Mac</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>-----------</td>
<td>-----------</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>

### 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</td>
<td># X</td>
</tr>
<tr>
<td>Copy selection to clipboard</td>
<td>Ctrl C</td>
<td># C</td>
</tr>
<tr>
<td>Paste from clipboard</td>
<td>Ctrl V</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>

### Menu Shortcuts

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

### 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 O</td>
<td># O</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 W</td>
<td># W</td>
</tr>
<tr>
<td>Exit application</td>
<td>Ctrl Q</td>
<td># Q</td>
</tr>
</tbody>
</table>

### Query

<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 query info view</td>
<td>Ctrl I</td>
<td># I</td>
</tr>
<tr>
<td>Create new editor tab</td>
<td>Ctrl T</td>
<td># T</td>
</tr>
<tr>
<td>Open existing text file</td>
<td>Ctrl R</td>
<td># R</td>
</tr>
<tr>
<td>Save text file</td>
<td>Ctrl S</td>
<td># S</td>
</tr>
<tr>
<td>Save copy of text file</td>
<td>Ctrl Shift S</td>
<td># Shift S</td>
</tr>
<tr>
<td>Close editor tab</td>
<td>Ctrl W</td>
<td># W</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>

### View

<table>
<thead>
<tr>
<th>Description</th>
<th>Win/Linux</th>
<th>Mac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toggle text view</td>
<td>Ctrl 1</td>
<td># 1</td>
</tr>
<tr>
<td>Toggle map view</td>
<td>Ctrl 2</td>
<td># 2</td>
</tr>
<tr>
<td>Toggle tree view</td>
<td>Ctrl 3</td>
<td># 3</td>
</tr>
<tr>
<td>Toggle folder view</td>
<td>Ctrl 4</td>
<td># 4</td>
</tr>
<tr>
<td>Toggle plot view</td>
<td>Ctrl 5</td>
<td># 5</td>
</tr>
<tr>
<td>Toggle table view</td>
<td>Ctrl 5</td>
<td># 5</td>
</tr>
<tr>
<td>Toggle explorer view</td>
<td>Ctrl 7</td>
<td># 7</td>
</tr>
</tbody>
</table>

### Nodes

<table>
<thead>
<tr>
<th>Description</th>
<th>Win/Linux</th>
<th>Mac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copy path to current node</td>
<td>Ctrl Shift C</td>
<td># Shift C</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 P</td>
<td># , (comma)</td>
</tr>
</tbody>
</table>

### Changelog

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 9. Standalone Mode

Read this entry online in the BaseX Wiki.

This page is part of the Getting Started Section. BaseX offers a standalone (embedded) 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).

Working with the BaseX Console

After the BaseX Console has been started, the HELP command can be used to to list all database commands. Multiple commands can be separated by semicolons.

To evaluate commands without entering the console mode, you can use the -c option on the command line:

```
basex -Vc "CREATE DB input <example/>; XQUERY /"
```

Database 'input' created in 124.95 ms.
<example/>

Query: /
Compiling:
Result: root()
Parsing: 0.42 ms
Compiling: 9.3 ms
Evaluating: 0.35 ms
Printing: 5.53 ms
Total Time: 15.62 ms

Hit(s): 1 Item
Updated: 0 Items
Printed: 10 Bytes

Query executed in 15.62 ms.

All available command-line options can be found here.

See also

GUI, Database Server, Getting Started
Chapter 10. Web Application

Read this entry online in the BaseX Wiki.

BaseX provides access to stored database resources and to the XQuery engine via REST, RESTXQ and WebDAV services. This article describes different ways of deploying and configuring these services. The services can be deployed in three different ways:

• as standalone application by running the BaseX HTTP Server,

• as web servlets in a J2EE Servlet Container, and

• for development purposes, 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. 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 located at `http://localhost:8080/manager/html/`).

Configuration

All database options can be specified in the `web.xml` file by prefixing the key 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>User name. By default, the admin user is specified. 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>Login data. By default, the admin password is specified. 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, the servlets will work in client/server mode, and a database server instance will be started along with the web server, which can also be addressed from other BaseX clients. If the flag is set to true, all servlets will communicate with a local database context which is not accessible from outside.</td>
</tr>
<tr>
<td>RESTXQPATH</td>
<td>.</td>
<td>RESTXQ directory. By default, all RESTXQ modules are located in the standard web application directory.</td>
</tr>
</tbody>
</table>

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>
```
How to set these options in the web.xml of the BaseX web application is specific to the servlet container. For example, in Jetty it is 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.

Since Version 7.7, different credentials can be assigned to each HTTP service by specifying local init parameters. In the following example, the global credentials are overwritten and reset 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/>
    </init-param>
    <init-param>
        <param-name>org.basex.password</param-name>
        <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>Version 7.7: http://[host]:[port]/[servlet_context_path]/static Before: 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>Version 7.7: http://[host]:[port]/[servlet_context_path]/restxq</td>
<td>Create XQuery web services and applications.</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 or webdav://[host]:[port]/[servlet_context_path]/webdav (depending on client)</td>
<td>Access databases via the filesystem.</td>
</tr>
</tbody>
</table>
### Maven

Checkout the sources of `basex` and `basex-api` via [Maven](https://maven.apache.org) or [Eclipse](https://eclipse.org). Execute `mvn install` in the `basex` 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](#section). The Jetty stop port can be changed in the [Maven Jetty Plugin](https://maven.apache.org/plugins/plugin-index.html#plugintype-jetty) session in the `pom.xml` file.

### User Management

If the HTTP server is started with no pre-defined credentials, users and passwords can be sent via [HTTP basic authentication](https://tools.ietf.org/html/rfc7235) with each HTTP request. Login data can be stored server-side in the `web.xml` file, or specified as command-line arguments.

For multi-user access, or a changed `admin` password, you may place the `.basexperr` configuration file in the server root. More details are found in the [User Management](#section) article.

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 `USERNAME:PASSWORD@URL`. An example:

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

### Changelog

**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
Part III. General Info
Chapter 11. 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 12. 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 e.g. be executed from the Command Line, Scripts, 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

Database commands in both the string and XML syntax can be placed in a text file and stored on disk. The default extension for BaseX command scripts is .bxs. If the path to a command script is passed on to BaseX, it will automatically be recognized and evaluated as such.

String Syntax

Multiple commands can be written in a single line and separated by semicolons, or stored as command script. Lines starting with # are interpreted as comments and are skipped. The following script creates a database, adds two documents to it and performs a query:

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

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.

This is why database commands can also be specified in XML. Multiple commands can be enclosed by a
<commands/> root element:

```
<commands>
    <create-db name='test'/>
    <add>input.xml</add>
    <add path='embedded.xml'><root>embedded</root></add>
    <xquery>count(//text())</xquery>
</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

Updated with Version 7.7:
Database, user and event 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

Before Version 7.7, names were restricted to letters, numbers, underscores and dashes, matching the regular expression `[-_a-zA-Z0-9]{1,128}`.

**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 the database `[name]` with an optional `[input]` and opens it. 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
- several additional Create Options can be set to influence the creation process.

**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 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]
```

**XML Syntax**

```
<open name='...'/>
```

**Permission**

`READ`

**Summary**

Opens the database specified by `[name]`.
<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><strong>CHECK</strong></td>
<td>CHECK [input]</td>
<td><code>&lt;check input='...'/&gt;</code></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 specified database does not exist, is currently being updated by another process or cannot be opened for some other reason.</td>
</tr>
<tr>
<td><strong>CLOSE</strong></td>
<td>CLOSE</td>
<td><code>&lt;close/&gt;</code></td>
<td>READ</td>
<td>Closes the currently opened database.</td>
<td>The command fails if the addressed database could neither be opened nor created.</td>
</tr>
<tr>
<td><strong>EXPORT</strong></td>
<td>EXPORT [path]</td>
<td><code>&lt;export path='...'/&gt;</code></td>
<td>CREATE</td>
<td>Exports all documents in the database to the specified [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><strong>CREATE INDEX</strong></td>
<td>CREATE INDEX [TEXT</td>
<td>ATTRIBUTE</td>
<td>FULLTEXT]</td>
<td>`&lt;create-index type='text</td>
<td>attribute</td>
</tr>
<tr>
<td><strong>DROP INDEX</strong></td>
<td>DROP INDEX [TEXT</td>
<td>ATTRIBUTE</td>
<td>FULLTEXT]</td>
<td>`&lt;drop-index type='text</td>
<td>attribute</td>
</tr>
</tbody>
</table>
## 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. Please note that Java 7 is required to handle ZIP files larger than 4 GB.

**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-2011-04-01-12-27-28.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.

**Errors**

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 ([TAG|ATTNAME|PATH|TEXT|ATTRIBUTE|FULLTEXT])

#### XML Syntax

- `<info-index type='tag|attname|path|text|attribute|fulltext'/>`

#### Permission

- **READ**

#### Summary

Shows information on the existing index structures. The output can be optionally limited to the specified index.

#### Errors

The command fails if no database is opened, or if the specified index is unknown.

### INFO STORAGE

#### Syntax

INFO STORAGE [start end] | [query]

#### XML Syntax

- `<info-storage>{{query}}</info-storage>`
### Commands

<table>
<thead>
<tr>
<th>Permission</th>
<th>Summary</th>
<th>Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>READ</strong></td>
<td>Shows the internal main table of the currently opened database. An integer range or a query may be specified as argument.</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>XML Syntax</th>
<th>Permission</th>
<th>Summary</th>
<th>Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST(<a href="%5E%5Bpath%5D">^name</a>)</td>
<td>&lt;list (name='...' (path='...'))/&gt;</td>
<td><strong>NONE</strong></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].</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>XML Syntax</th>
<th>Permission</th>
<th>Summary</th>
<th>Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>XQUERY [query]</td>
<td>&lt;xquery&gt;[query]&lt;/xquery&gt;</td>
<td><em>depends on query</em></td>
<td>Runs the specified [query] and prints the result.</td>
<td>The command fails if the specified query is invalid.</td>
</tr>
</tbody>
</table>

**RETRIEVE**

<table>
<thead>
<tr>
<th>Syntax</th>
<th>XML Syntax</th>
<th>Permission</th>
<th>Summary</th>
<th>Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>RETRIEVE [path]</td>
<td>&lt;retrieve path='...'/&gt;</td>
<td><strong>READ</strong></td>
<td>Retrieves a raw file from the opened database at the specified [path].</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>XML Syntax</th>
<th>Permission</th>
<th>Summary</th>
<th>Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIND [query]</td>
<td>&lt;find&gt;[query]&lt;/find&gt;</td>
<td><strong>READ</strong></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>
<td>The command fails if no database is opened.</td>
</tr>
</tbody>
</table>
### Commands

#### CS

<table>
<thead>
<tr>
<th>Syntax</th>
<th>CS [query]</th>
</tr>
</thead>
<tbody>
<tr>
<td>XML Syntax</td>
<td><code>&lt;cs&gt;[query]&lt;/cs&gt;</code></td>
</tr>
<tr>
<td>Permission</td>
<td>depends on query</td>
</tr>
<tr>
<td>Summary</td>
<td>Evaluates the specified [query] and declares the resulting nodes as new context set. In subsequent queries, the context set will be available via the context item expression of XQuery (.)</td>
</tr>
<tr>
<td>Errors</td>
<td>The command fails if no database is opened, if the specified query is invalid or if it does not return nodes of the currently opened database.</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><code>&lt;repo-install path='...'/&gt;</code></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><code>&lt;repo-list/&gt;</code></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

<table>
<thead>
<tr>
<th>Syntax</th>
<th>REPO DELETE [name]</th>
</tr>
</thead>
<tbody>
<tr>
<td>XML Syntax</td>
<td><code>&lt;repo-delete name='...'/&gt;</code></td>
</tr>
<tr>
<td>Permission</td>
<td>CREATE</td>
</tr>
<tr>
<td>Summary</td>
<td>Deletes the package with name [name], optionally followed by a version.</td>
</tr>
<tr>
<td>Errors</td>
<td>The command fails if the package to be deleted participates in a dependency.</td>
</tr>
</tbody>
</table>

### Updates

#### ADD

<table>
<thead>
<tr>
<th>Syntax</th>
<th>ADD (TO [path]) [input]</th>
</tr>
</thead>
<tbody>
<tr>
<td>XML Syntax</td>
<td><code>&lt;add (path='...')&gt;[input]&lt;/add&gt;</code></td>
</tr>
</tbody>
</table>
## Commands

<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><strong>ADD</strong></td>
<td>ADD [path] [input]</td>
<td>&lt;add path='...'/&gt;</td>
<td>WRITE</td>
<td>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.</td>
<td>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.</td>
</tr>
<tr>
<td><strong>DELETE</strong></td>
<td>DELETE [path]</td>
<td>&lt;delete path='...'/&gt;</td>
<td>WRITE</td>
<td>Deletes all documents from the currently opened database that start with the specified [path].</td>
<td>The command fails if no database is opened.</td>
</tr>
<tr>
<td><strong>RENAME</strong></td>
<td>RENAME [path] [newpath]</td>
<td>&lt;rename path='...' newpath='...'/&gt;</td>
<td>WRITE</td>
<td>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.</td>
<td>The command fails if no database is opened, or if the target path is empty.</td>
</tr>
<tr>
<td><strong>REPLACE</strong></td>
<td>REPLACE [path] [input]</td>
<td>&lt;replace path='...'&gt;[input]&lt;/replace&gt;</td>
<td>WRITE</td>
<td>Replaces a document in the currently opened database, addressed by [path], with the file or XML string specified by [input], or adds it as a new document.</td>
<td>The command fails if no database is opened, if the specified path points to a database directory, or if the input is not found.</td>
</tr>
<tr>
<td><strong>STORE</strong></td>
<td>STORE (TO [path]) [input]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**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.
**Commands**

**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 the ALL flag is specified, the internal database structures are completely rebuilt; this often leads to a reduction of the total database size.

**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

**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, local users are shown.

**Errors** The command fails if the optional database could not be opened.

#### KILL

**Syntax** KILL [target]
# Commands

**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.

## CREATE EVENT

**Syntax** `CREATE EVENT [NAME]`

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

**Permission** `ADMIN`

**Summary** Creates the specified event.

**Errors** The command fails if event already exists.

## SHOW EVENTS

**Syntax** `SHOW EVENTS`

**XML Syntax** `<show-events/>`

**Permission** `ADMIN`

**Summary** Shows all events that have been registered in the database.

## DROP EVENT

**Syntax** `DROP EVENT [NAME]`

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

**Permission** `ADMIN`

**Summary** Drops the specified event.

**Errors** The command fails if the event doesn’t exist.

## 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]. [name] must be a valid user name. The password must be a valid MD5 hash value. If no password is specified in the console mode, it is requested via standard input.

**Errors** The command fails if the specified user already exists, or if the password is no valid MD5 hash value.

## ALTER USER

**Syntax** `ALTER USER [name] ([password])`

**XML Syntax** `<alter-user name='...'>([password])</alter-user>`

**Permission** `ADMIN`
### PASSWORD

**Syntax**

```
PASSWORD ([password])
```

**XML Syntax**

```
<password>([password])</password>
```

**Permission**

*NONE*

**Summary**

Changes the [password] of the current user. The password must be a valid MD5 hash value. If no password is specified in the console mode, it is requested via standard input.

**Errors**

The command fails if the password is no valid MD5 hash value.

### RUN

**Syntax**

```
RUN [file]
```

**XML Syntax**

```
<run file='...'/>
```

**Permission**

*depends on input*

**Summary**

Evaluates the contents of [file] as XQuery expression. If the file ends with the suffix .bxs, the file content will be evaluated as command script. This command can be used to run several commands in a single transaction.

---

**Summary**

Alters the [password] of the user specified by [name]. The password must be a valid MD5 hash value. If no password is specified in the console mode, it is requested via standard input.

**Errors**

The command fails if the specified user does not exist, or if the password is no valid MD5 hash value.

### DROP USER

**Syntax**

```
DROP USER [name] (ON [database]):
```

**XML Syntax**

```
<drop-user name='...' (database='...')/>
```

**Permission**

*ADMIN*

**Summary**

Drops the user with the specified [name]. If a [database] is specified, the user is only dropped locally. The Glob Syntax can be used to address more than one database or user.

**Errors**

The command fails if admin is specified as user name, if the specified user does not exist or is logged in, or if the optional database could not be opened for modification.

### GRANT

**Syntax**

```
GRANT [NONE|READ|WRITE|CREATE|ADMIN] (ON [database]) TO [user]
```

**XML Syntax**

```
<grant name='...' permission='none|read|write|create|admin' (database='...')/>
```

**Permission**

*ADMIN*

**Summary**

Grants the specified permission to the specified [user]. If a [database] is specified, the permissions are only granted locally. The Glob Syntax can be used to address more than one database or user.

**Errors**

The command fails if admin is specified as user name, if the specified user does not exist, or if the optional database could not be opened for modification.

**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*.

### General Commands
### Commands

<table>
<thead>
<tr>
<th>Errors</th>
<th>The command fails if the specified file does not exist, or if the retrieved input is invalid. It will be canceled as soon as one of the executed commands fails.</th>
</tr>
</thead>
</table>
| Examples | • RUN query.xq will evaluated the specified file as XQuery expression  
• RUN commands.bxs will evaluated the specified file as command script |

**EXECUTE**

<table>
<thead>
<tr>
<th>Syntax</th>
<th>EXECUTE [input]</th>
</tr>
</thead>
<tbody>
<tr>
<td>XML Syntax</td>
<td><code>&lt;execute&gt;[input]&lt;/execute&gt;</code></td>
</tr>
<tr>
<td>Permission</td>
<td>depends on input</td>
</tr>
<tr>
<td>Summary</td>
<td>Evaluates the specified [input] as command script. This command can be used to run several commands in a single transaction.</td>
</tr>
<tr>
<td>Errors</td>
<td>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.</td>
</tr>
</tbody>
</table>
| Examples | • 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**

<table>
<thead>
<tr>
<th>Syntax</th>
<th>GET [option]</th>
</tr>
</thead>
<tbody>
<tr>
<td>XML Syntax</td>
<td><code>&lt;get option='...'/&gt;</code></td>
</tr>
<tr>
<td>Permission</td>
<td>NONE</td>
</tr>
<tr>
<td>Summary</td>
<td>Returns the current value of the Option specified via [option]. Global options can only be requested by users with ADMIN permissions.</td>
</tr>
<tr>
<td>Errors</td>
<td>The command fails if the specified option is unknown.</td>
</tr>
</tbody>
</table>

**SET**

<table>
<thead>
<tr>
<th>Syntax</th>
<th>SET [option] ([value])</th>
</tr>
</thead>
<tbody>
<tr>
<td>XML Syntax</td>
<td><code>&lt;set option='...'&gt;([value])&lt;/set&gt;</code></td>
</tr>
<tr>
<td>Permission</td>
<td>NONE</td>
</tr>
<tr>
<td>Summary</td>
<td>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.</td>
</tr>
<tr>
<td>Errors</td>
<td>The command fails if the specified option is unknown or if the specified value is invalid.</td>
</tr>
</tbody>
</table>

**INFO**

<table>
<thead>
<tr>
<th>Syntax</th>
<th>INFO</th>
</tr>
</thead>
<tbody>
<tr>
<td>XML Syntax</td>
<td><code>&lt;info/&gt;</code></td>
</tr>
<tr>
<td>Permission</td>
<td>READ</td>
</tr>
<tr>
<td>Summary</td>
<td>Shows global information.</td>
</tr>
</tbody>
</table>

**HELP**

<table>
<thead>
<tr>
<th>Syntax</th>
<th>HELP ([command])</th>
</tr>
</thead>
<tbody>
<tr>
<td>XML Syntax</td>
<td><code>&lt;help&gt;([command])&lt;/help&gt;</code></td>
</tr>
</tbody>
</table>
Commands

<table>
<thead>
<tr>
<th>Permission</th>
<th>NONE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>If [command] is specified, information on the specific command is printed; otherwise, all commands are listed.</td>
</tr>
<tr>
<td>Errors</td>
<td>The command fails if the specified command is unknown.</td>
</tr>
</tbody>
</table>

**EXIT**

<table>
<thead>
<tr>
<th>Syntax</th>
<th>EXIT</th>
</tr>
</thead>
<tbody>
<tr>
<td>XML Syntax</td>
<td>&lt;exit/&gt;</td>
</tr>
<tr>
<td>Permission</td>
<td>NONE</td>
</tr>
<tr>
<td>Summary</td>
<td>Exits the console mode.</td>
</tr>
</tbody>
</table>

**Changelog**

Version 7.7
- 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 13. Databases

Read this entry online in the BaseX Wiki.

This page is part of the Getting Started Section.

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.

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 and from other locations:

<table>
<thead>
<tr>
<th>Function</th>
<th>Example</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>db:open()</td>
<td>db:open(&quot;db&quot;, &quot;path/to/docs&quot;)</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>collection(&quot;db/path/to/docs&quot;)</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 currently opened database will be returned.</td>
</tr>
<tr>
<td>fn:doc()</td>
<td>doc(&quot;db/path/to/doc.xml&quot;)</td>
<td>Returns the document at the location path/to/docs in the database db. An error is raised if the specified addresses does not address exactly one document.</td>
</tr>
</tbody>
</table>

The fn:document-uri() and fn:base-uri() functions return URIs that can be reused as arguments for the fn:doc() and fn:collection() functions. As a result of this, as an example, the following query will always return 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.
- collection("/path/to/docs") : returns a main-memory collection with all XML documents found at the addressed file path.

### Raw Files

- XQuery: db:retrieve("dbname", "path/to/docs") returns raw files in their Base64 representation. By choosing "method=raw" as Serialization Option, the data is returned in its original byte representation:

```xml
declare option output:method "raw";
db:retrieve('multimedia', 'sample.avi')
```

- Commands: RETRIEVE returns raw files without modifications.

### 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.
- Resource can be replaced with other ones 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 (introduced with Version 7.7) 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 to create, add, replace, and delete XML documents:

```xml
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

**Changelog**

Version 7.2.1

- Updated: `fn:document-uri()` and `fn:base-uri()` now return strings that can be reused with `fn:doc()` or `fn:collection()` to reopen the original document.
Chapter 14. 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 is used to store options to disk:

- Global options can only be set via the configuration file or system properties (see below).
- Local options can also be specified in the configuration file after the # Local Options comment.

Various ways exist to access and change options:

- The current value of an option can be requested with the GET and changed with the SET command. 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 existing value will be inverted.
- Initial values for 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 needs to be prefixed with org.basex.. An example:

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

  CHOP: false

- Options can also be set in the prolog of an XQuery expression. 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';
  ...

- Options can also be applied locally by using pragmas:

  (# db:chop false #) { parse-xml('<xml> hi </xml>') }

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

Global Options

General

DBPATH

<table>
<thead>
<tr>
<th>Signature</th>
<th>DBPATH [path]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>(home)/BaseXData or (home)/data</td>
</tr>
<tr>
<td>Summary</td>
<td>Points to the directory in which all databases are located.</td>
</tr>
</tbody>
</table>

REPOPATH

| Signature   | REPOPATH [path] |
## Options

<table>
<thead>
<tr>
<th>Default</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>(home)/BaseXRepo</td>
<td>Points to the Repository, in which all XQuery modules are located.</td>
</tr>
</tbody>
</table>

### DEBUG

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

### LANG

<table>
<thead>
<tr>
<th>Signature</th>
<th>Default</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<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>
</tbody>
</table>

### LANGKEY

<table>
<thead>
<tr>
<th>Signature</th>
<th>Default</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<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>
</tbody>
</table>

### GLOBALLOCK

<table>
<thead>
<tr>
<th>Signature</th>
<th>Default</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<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

### HOST

<table>
<thead>
<tr>
<th>Signature</th>
<th>Default</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<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>
</tbody>
</table>

### PORT

<table>
<thead>
<tr>
<th>Signature</th>
<th>Default</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<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>
</tbody>
</table>

### SERVERPORT

<table>
<thead>
<tr>
<th>Signature</th>
<th>Default</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>SERVERPORT [port]</td>
<td>1984</td>
<td></td>
</tr>
</tbody>
</table>
**Options**

**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`.

**EVENTPORT**

**Signature**
EVENTPORT [port]

**Default**
1985

**Summary**
This port is used by the client to listen for server events. It will only be bound if a client attaches itself to a database event. This option can also be changed when running the server on command line via `-e`.

**USER**

**Signature**
USER [name]

**Default**
empty

**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**

**Signature**
PASSWORD [password]

**Default**
empty

**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.

**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**

**Signature**
PROXYHOST [host]

**Default**
empty

**Summary**
This is the host name of a proxy server.

**PROXYPOR**

**Signature**
PROXYPOR [port]
### Options

<table>
<thead>
<tr>
<th>Default</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>This is the port number of a proxy server.</td>
</tr>
</tbody>
</table>

### NONPROXYHOSTS

<table>
<thead>
<tr>
<th>Signature</th>
<th>NONPROXYHOSTS [hosts]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>empty</td>
</tr>
<tr>
<td>Summary</td>
<td>This is a list of hosts that should be directly accessed.</td>
</tr>
</tbody>
</table>

### TIMEOUT

<table>
<thead>
<tr>
<th>Signature</th>
<th>TIMEOUT [seconds]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>30</td>
</tr>
<tr>
<td>Summary</td>
<td>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 ADMIN operations.</td>
</tr>
</tbody>
</table>

### KEEPALIVE

<table>
<thead>
<tr>
<th>Signature</th>
<th>KEEPALIVE [seconds]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>600</td>
</tr>
<tr>
<td>Summary</td>
<td>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.</td>
</tr>
</tbody>
</table>

### PARALLEL

<table>
<thead>
<tr>
<th>Signature</th>
<th>PARALLEL [number]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>8</td>
</tr>
<tr>
<td>Summary</td>
<td>Denotes the maximum allowed number of parallel read transactions.</td>
</tr>
</tbody>
</table>

### LOG

<table>
<thead>
<tr>
<th>Signature</th>
<th>LOG [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>true</td>
</tr>
<tr>
<td>Summary</td>
<td>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.</td>
</tr>
</tbody>
</table>

### LOGMSGMAXLEN

<table>
<thead>
<tr>
<th>Signature</th>
<th>LOGMSGMAXLEN [length]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>1000</td>
</tr>
<tr>
<td>Summary</td>
<td>Specifies the maximum length of a single log message.</td>
</tr>
</tbody>
</table>

### HTTP Options

If BaseX is run as Web Application, the HTTP options are either determined by the web server, or specified in the webapp/WEB-INF directory and the jetty.xml and web.xml configuration files.
Options

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.

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.

HTTPLOCAL

**Signature**  HTTPLOCAL [boolean]
**Default**  false
**Summary**  By default, a database server instance will be opened along with the web server. If the flag is set to true, all commands will be executed in an embedded database context. If BaseX is run as Web Application, and if the flag is false, the server will be started as soon as the first HTTP service is called.

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

Create Options

General

MAINMEM

**Signature**  MAINMEM [boolean]
**Default**  false
**Summary**  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.

ADDCACHE

*Introduced with Version 7.7:*

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

### Parsing

#### CREATEFILTER

<table>
<thead>
<tr>
<th>Signature</th>
<th><code>CREATEFILTER [filter]</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td><code>*.xml</code></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>

#### ADDARCHIVES

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

#### SKIPCORRUPT

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

#### ADDRAW

| Signature         | `ADDRAW [boolean]`                                                  |
|-------------------|                                                                     |
| Default           | `false`                                                             |
| Summary           | 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). |

#### PARSER

| Signature         | `PARSER [type]`                                                     |
|-------------------|                                                                     |
| Default           | `XML`                                                               |
| Summary           | 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. |

#### PARSEROPT

| Signature         | `PARSEROPT [options]`                                               |
|-------------------|                                                                     |
| Default           | `empty`                                                             |
| Summary           | Defines parser-specific options; see Parsers for more information. |

#### HTMLOPT

| Signature         | `HTMLOPT [options]`                                                 |
|-------------------|                                                                     |
### Options

<table>
<thead>
<tr>
<th>Default</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>empty</td>
<td>Allows to specify TagSoup options for HTML parsing; see HTML Parser for more information.</td>
</tr>
</tbody>
</table>

#### XML Parsing

**CHOP**

<table>
<thead>
<tr>
<th>Signature</th>
<th>CHOP [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>true</td>
</tr>
<tr>
<td>Summary</td>
<td>Chops all leading and trailing whitespaces from text nodes while building a database, and discards empty text nodes. By default, this option is set to <code>true</code>, as it often reduces the database size by up to 50%. It can also be turned off on command line via <code>-w</code>.</td>
</tr>
</tbody>
</table>

**INTPARSE**

<table>
<thead>
<tr>
<th>Signature</th>
<th>INTPARSE [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>true</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>

**CATFILE**

<table>
<thead>
<tr>
<th>Signature</th>
<th>CATFILE [path]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>empty</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 current index and full-text index options will be stored in a new database, and take effect if indexes are rebuilt via the `OPTIMIZE`.

**TEXTINDEX**

<table>
<thead>
<tr>
<th>Signature</th>
<th>TEXTINDEX [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>true</td>
</tr>
<tr>
<td>Summary</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 Indexes for more details.</td>
</tr>
</tbody>
</table>

**ATTRINDEX**

<table>
<thead>
<tr>
<th>Signature</th>
<th>ATTRINDEX [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>true</td>
</tr>
</tbody>
</table>
## Options

### Summary
Creates an attribute index whenever a new database is created. An attribute index speeds up queries with equality comparisons on attribute values; see Indexes for more details.

### FTINDEX

<table>
<thead>
<tr>
<th>Signature</th>
<th>FTINDEX [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>false</td>
</tr>
<tr>
<td>Summary</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 Indexes for more details.</td>
</tr>
</tbody>
</table>

### MAXLEN

<table>
<thead>
<tr>
<th>Signature</th>
<th>MAXLEN [int]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>96</td>
</tr>
<tr>
<td>Summary</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>
</tbody>
</table>

### MAXCATS

<table>
<thead>
<tr>
<th>Signature</th>
<th>MAXCATS [int]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>100</td>
</tr>
<tr>
<td>Summary</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>
</tbody>
</table>

### UPDINDEX

<table>
<thead>
<tr>
<th>Signature</th>
<th>UPDINDEX [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>false</td>
</tr>
</tbody>
</table>
| Summary   | If turned on, incremental indexing will be activated:  
  - with each update, the text and attribute value indexes will be updated as well.  
  - The value of this option will be assigned once to a new database, and cannot be changed after that.  
  - The advantage of incremental indexes is that the value index structures will always be up-to-date.  
  - The downside is that updates will take longer. The article on Index Structures includes additional details. |

### INDEXSPLITSIZE

*Introduced with Version 7.7:*

<table>
<thead>
<tr>
<th>Signature</th>
<th>INDEXSPLITSIZE [num]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>0</td>
</tr>
<tr>
<td>Summary</td>
<td>This option affects the construction of new text and attribute indexes. It specifies the number of index build operations that are performed before writing partial index data to disk. By default, if the value is set to 0, some dynamic split heuristics are applied.</td>
</tr>
</tbody>
</table>

### FTINDEXSPLITSIZE

*Introduced with Version 7.7:*
**Options**

<table>
<thead>
<tr>
<th><strong>Signature</strong></th>
<th><strong>FTINDEXSPLITSIZE [num]</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Default</strong></td>
<td>0</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>This option affects the construction of new full-text indexes. It specifies the number of index build operations that are performed before writing partial index data to disk. By default, if the value is set to 0, some dynamic split heuristics are applied.</td>
</tr>
</tbody>
</table>

**Full-Text**

**STEMMING**

<table>
<thead>
<tr>
<th><strong>Signature</strong></th>
<th><strong>STEMMING [boolean]</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Default</strong></td>
<td>false</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>A new full-text index will stem all tokens and speed up queries on stemmed tokens. The same stemming normalization will be applied to all query tokens that are checked against tokens in this index.</td>
</tr>
</tbody>
</table>

**CASESENS**

<table>
<thead>
<tr>
<th><strong>Signature</strong></th>
<th><strong>CASESENS [boolean]</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Default</strong></td>
<td>false</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>A new full-text index will preserve the case of all tokens. The same case normalization will be applied to all query tokens that are checked against tokens in this index.</td>
</tr>
</tbody>
</table>

**DIACRITICS**

<table>
<thead>
<tr>
<th><strong>Signature</strong></th>
<th><strong>DIACRITICS [boolean]</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Default</strong></td>
<td>false</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>A new full-text index will preserve the diacritics of all tokens. The same diacritics normalization will be applied to all query tokens that are checked against tokens in this index.</td>
</tr>
</tbody>
</table>

**LANGUAGE**

<table>
<thead>
<tr>
<th><strong>Signature</strong></th>
<th><strong>LANGUAGE [lang]</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Default</strong></td>
<td>en</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>A new full-text index will use the given language to normalize all tokens. This option is mainly important if tokens are to be stemmed, or if the tokenization of a language differs from Western languages.</td>
</tr>
</tbody>
</table>

**STOPWORDS**

<table>
<thead>
<tr>
<th><strong>Signature</strong></th>
<th><strong>STOPWORDS [path]</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Default</strong></td>
<td>empty</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>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. A standard stopword list for English texts is provided in the directory etc/stopwords.txt in the official releases.</td>
</tr>
</tbody>
</table>

**Query Options**

**QUERYINFO**

<table>
<thead>
<tr>
<th><strong>Signature</strong></th>
<th><strong>QUERYINFO [boolean]</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Default</strong></td>
<td>false</td>
</tr>
<tr>
<td>Summary</td>
<td>Prints more information on internal query rewritings, optimizations, and performance. By default, this info is shown in the <strong>Info View</strong> in the GUI. It can also be activated on <strong>command line</strong> via <code>-V</code>.</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>

**XQUERY3**

<table>
<thead>
<tr>
<th>Signature</th>
<th>XQUERY3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>true</td>
</tr>
<tr>
<td>Summary</td>
<td>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.</td>
</tr>
</tbody>
</table>

**BINDINGS**

<table>
<thead>
<tr>
<th>Signature</th>
<th>BINDINGS [vars]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>empty</td>
</tr>
<tr>
<td>Summary</td>
<td>Contains external variables to be bound to a query:</td>
</tr>
<tr>
<td></td>
<td>• Variable names and values are separated by equality signs, and multiple variables are delimited by commas.</td>
</tr>
<tr>
<td></td>
<td>• Variables may optionally be introduced with a leading dollar sign.</td>
</tr>
<tr>
<td></td>
<td>• Commas that occur in the value itself are encoded by duplication.</td>
</tr>
<tr>
<td></td>
<td>• If a variable uses a namespace different to the default namespace, it can be specified with the Clark Notation or Expanded QName Notation.</td>
</tr>
<tr>
<td></td>
<td>• This option can also be used on <strong>command line</strong> with the flag -b.</td>
</tr>
<tr>
<td>Examples</td>
<td>$a=1,$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 $a$. <code>{URI}a=x</code> binds the value x to the variable $a$ with the namespace URI.</td>
</tr>
</tbody>
</table>

**QUERYPATH**

<table>
<thead>
<tr>
<th>Signature</th>
<th>QUERYPATH [path]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>empty</td>
</tr>
<tr>
<td>Summary</td>
<td>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.</td>
</tr>
</tbody>
</table>

**CACHEQUERY**

<table>
<thead>
<tr>
<th>Signature</th>
<th>CACHEQUERY [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>false</td>
</tr>
<tr>
<td>Summary</td>
<td>Caches the query results before returning them to the client. This option may be set to true if the whole result is needed for further operations (such as is e.g. the case in the GUI of BaseX).</td>
</tr>
</tbody>
</table>

**FORCECREATE**

<table>
<thead>
<tr>
<th>Signature</th>
<th>FORCECREATE [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>false</td>
</tr>
<tr>
<td>Summary</td>
<td>By activating this option, the XQuery <code>doc()</code> and <code>collection()</code> functions will create database instances for the addressed input files.</td>
</tr>
</tbody>
</table>

**CHECKSTRINGS**

*Introduced with Version 7.7:*
### Options

<table>
<thead>
<tr>
<th>Signature</th>
<th>CHECKSTRINGS [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>true</td>
</tr>
</tbody>
</table>
| 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

<table>
<thead>
<tr>
<th>Signature</th>
<th>LSERROR [error]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>0</td>
</tr>
<tr>
<td>Summary</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>

### RUNS

<table>
<thead>
<tr>
<th>Signature</th>
<th>RUNS [num]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>1</td>
</tr>
<tr>
<td>Summary</td>
<td>Specify number of runs a query is executed by the XQUERY command. The result is only serialized once, and the measured times are averages of all runs.</td>
</tr>
</tbody>
</table>

### Serialization Options

#### SERIALIZER

<table>
<thead>
<tr>
<th>Signature</th>
<th>SERIALIZE [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>true</td>
</tr>
<tr>
<td>Summary</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>

<table>
<thead>
<tr>
<th>Signature</th>
<th>SERIALIZER [params]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>empty</td>
</tr>
</tbody>
</table>
| Summary         | Contains parameters for serializing query results:  
  - Keys and values are separated by equality signs.  
  - Multiple parameters are delimited by commas.  
  - The option can also be used on command line with the flag -s.  
  Example: `encoding=US-ASCII,omit-xml-declaration=no` sets the encoding to US-ASCII and prints the XML declaration. |

### EXPORTER

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

**XMLPLAN**

<table>
<thead>
<tr>
<th>Signature</th>
<th>XMLPLAN [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Default</strong></td>
<td>false</td>
</tr>
<tr>
<td><strong>Summary</strong></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><strong>Default</strong></td>
<td>true</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>Creates the query plan before or after the compilation step. Query plans might change due to optimizations.</td>
</tr>
</tbody>
</table>

**DOTPLAN**

<table>
<thead>
<tr>
<th>Signature</th>
<th>DOTPLAN [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Default</strong></td>
<td>false</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>Visualizes the execution plan of an XQuery expression with <em>dotty</em> and saves its dot file in the query directory.</td>
</tr>
</tbody>
</table>

**DOTCOMPACT**

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

**DOTDISPLAY**

<table>
<thead>
<tr>
<th>Signature</th>
<th>DOTDISPLAY [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Default</strong></td>
<td>true</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>Visualizes the dot representation after the query execution.</td>
</tr>
</tbody>
</table>

**DOTTY**

<table>
<thead>
<tr>
<th>Signature</th>
<th>DOTTY [path]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Default</strong></td>
<td><em>dotty</em></td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>Location of the <em>dotty</em> executable.</td>
</tr>
</tbody>
</table>

**Other Options**

**AUTOFLUSH**

<table>
<thead>
<tr>
<th>Signature</th>
<th>AUTOFLUSH [boolean]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Default</strong></td>
<td>true</td>
</tr>
</tbody>
</table>
Options

Summary | 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.

**WRITEBACK**

**Signature** | WRITEBACK [boolean]
---|---
**Default** | false

**Summary** | Updates on XML nodes are written back to the input files. No backups of your original files will be created if this option is turned on. This option can also be activated on command line via −u.

**MAXSTAT**

**Signature** | MAXSTAT [num]
---|---
**Default** | 30

**Summary** | Specifies the maximum number of index occurrences printed by the INFO INDEX command.

**Changelog**

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
• Added: USER, PASSWORD, LOG, LOGMSGMAXLEN, WEBPATH, RESTXQPATH HTTPLOCAL, CREATEONLY, STRIPNS
• Removed: HTTPATH; 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, PROXYPORT, NONPROXYHOSTS, HTMLOPT
• Updated: TIMEOUT: ignore timeout for admin users

Version 7.1

• Added: ADDRAW, MAXLEN, MAXCATS, UPDINDEX
Options

- Updated: BINDINGS

Version 7.0

- Added: SERVERHOST, KEEPALIVE, AUTOFLUSH, QUERYPATH
Chapter 15. Parsers

Read this entry online in the BaseX Wiki.

This article is part of the Getting Started Section. It presents different parsers for importing various data source into BaseX databases. For export see Serialization.

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() or db:replace() function can be used as well to add new XML documents to the database. The following example query uses the internal parser and adds all files to the database DB that are found in the directory 2Bimported:

```
declare option db:intparse "yes";
for $file in file:list("2Bimported")
return db:add('DB', $file)
```

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 this 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:

<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 HTMLOPT method=xml,nons=true,ncdata=true,nodefaults=true,nobogons=true,nocolons=true,ignorable=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 in the query prolog:
declare option db:parser "html";
declare option db:htmlopt "html=false";
doc("index.html")

**JSON Parser**

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 PARSEOPT 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.
- **Removed in Version 7.7.2**: **XML format**: Choose the XML format. Possible: verbose, simple.

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

**XQuery**

Since Version 7.7.2, the CSV Module provides a function for converting HTML to XML documents.
Documents can also be converted by specifying the parser in the query prolog. 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:

```
declare option db:parser "csv";
declare option db:parseropt "header=yes";
for $file in file:list("2Bimported", false(), "*.csv")
return db:add('DB', $file)
```

**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:

```
SET PARSER text
SET PARSEROPT lines=yes
SET CREATEFILTER *
```

**XQuery**

Similar to the other formats the text parser can be specified in the prolog of an XQuery expression:

```
declare option db:parser "text";
for $file in file:list("2Bimported", true(), "*.txt")
return db:add('DB', $file)
```

**Changelog**

Version 7.7.2

- Removed: CSV option "format".

Version 7.3

- Updated: SEPARATOR option of CSV parser may now contain arbitrary code points.

Version 7.2

- Updated: Enhanced support for TagSoup options.
Part IV. Integration
Chapter 16. 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 OSX

- 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.base6.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.
Usage

The query execution works as follows:

1. Create a new XQuery Project with File → New → XQuery Project.
3. Edit your XQuery Module and execute it with Run.
4. The results are displayed in the Console window of Eclipse.
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

• Start the WebDAV service first, which will allow you to access all resources via the client/server architecture

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 also directly open single 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".
Integrating oXygen

Perform Queries

Preparations

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

2. Start a BaseX Server instance

Data Source

1. Start oXygen and go to Options → Preferences → Data Sources

2. Add a new Data Source with the New button

3. Enter "BaseX-XQJ" as connection name and choose XQuery API for Java (XQJ) as type

4. Add the following JAR files above with the Add button: xqj-api-1.0.jar, xqj2-0.1.0.jar and basex-xqj-1.2.3.jar (the version names of the JAR file may differ)

5. Now press OK, and your Data Source is ready for use

Connection

1. Now press New in the Connection Panel below.

2. Enter Name "BaseX" and select "BaseX-XQJ" in the Data Source box.

3. Enter the following connection details (or modify them when necessary):
   - Port: 1984
   - serverName: localhost
   - user: admin
Integrating oXygen

- password: admin

4. Now press OK, and your connection is ready.

**Usage**

The query execution works as follows:

1. Configure a new transformation scenario in Window → Show View → Transformation Scenarios.
2. Choose the XQuery Transformation tree entry.
3. Press the plus sign to add a new scenario.
4. Enter a Name and an optional XML and XQuery URL (e.g. your query document/file).
5. Choose "BaseX” as Transformer from the combo box.
6. Press OK, and your scenario is ready. Now you can start the transformation, e.g. by clicking on the red Play button.
7. The results should immediately occur in the result panel.
Part V. Query Features
Chapter 18. Full-Text

Read this entry online in the BaseX Wiki.

This article is part of the XQuery Portal. It summarizes the full-text and language-specific features of BaseX.

Full-text retrieval is an essential query feature for working with XML documents, and BaseX was the first query processor that fully supported the W3C XQuery Full Text 1.0 Recommendation. This page lists some singularities and extensions of the BaseX implementation.

Features

Options

The available full-text index can handle various combinations of the match options defined in the XQuery Full Text Recommendation. By default, most options are disabled. The GUI dialogs for creating new databases or displaying the database properties contain a tab for choosing between all available options. On the command-line, the SET command can be used to activate full-text indexing or creating a full-text index for existing databases:

- SET FTINDEX true; CREATE DB input.xml
- CREATE INDEX fulltext

The following indexing options are available:

- **Language**: see below for more details (SET LANGUAGE EN).
- **Stemming**: tokens are stemmed with the Porter Stemmer before being indexed (SET STEMMING true).
- **Case Sensitive**: tokens are indexed in case-sensitive mode (SET CASESENS true).
- **Diacritics**: diacritics are indexed as well (SET DIACRITICS true).
- **Stopword List**: a stop word list can be defined to reduce the number of indexed tokens (SET STOPWORDS [filename]).

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.
BaseX offers an internal scoring model which can be extended to different application scenarios. The score of a full-text result is calculated by taking the number of found terms and their frequency in a single text node into account. Terms will be ranked higher if they are found in short text nodes.

**Thesaurus**

BaseX supports full-text queries using thesauri, but it does not provide a default thesaurus. This is why query 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 fuzzy querying. The XQFT grammar was enhanced by the FTMatchOption using fuzzy to allow for approximate searches in full texts. By default, the standard full-text index already supports the efficient execution of fuzzy searches.

**Document 'doc.xml':**

```xml
<doc>
  <a>house</a>
  <a>hous</a>
  <a>haus</a>
</doc>
```

**Command:** CREATE DB doc.xml; CREATE INDEX fullext

**Query:**

```
//a[text() contains text 'house' using fuzzy]
```

**Result:**

```
<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”.

**Performance**

**Index Processing**

BaseX offers different evaluation strategies for XQFT queries, the choice of which depends on the input data and the existence of a full text index. The query compiler tries to optimize and speed up queries by applying a full text index structure whenever possible and useful. Three evaluation strategies are available: the standard sequential
Full-Text

database scan, a full-text index based evaluation and a hybrid one, combining both strategies (see XQuery Full Text implementation in BaseX). Query optimization and selection of the most efficient evaluation strategy is done in a full-fledged automatic manner. The output of the query optimizer indicates which evaluation plan is chosen for a specific query. It can be inspected by activating verbose querying (Command: SET VERBOSE ON) or opening the Query Info in the GUI. The message

Applying full-text index

suggests that the full-text index is applied to speed up query evaluation. A second message

Removing path with no index results

indicates that the index does not yield any results for the specified term and is thus skipped. If index optimizations are missing, it sometimes helps to give the compiler a second chance and try different rewritings of the same query.

FTAnd

The internal XQuery Full Text data model is pretty complex and may consume more main memory as would initially guess. If you plan to combine search terms via ft:and, we recommend you to resort to an alternative, memory-saving representation:

(: representation via "ftand" :)  
"A B" contains text "A" ftand "B" ftor "C" ftor "D"

(: memory saving representation :)  
"A B" contains text { "A", "B" } all ftor { "C", "D" } all

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:

<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 → 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 completely 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:

(: Structure is ignored; no highlighting: :)  
ft:mark('//p[. contains text 'real']')

(: Single text nodes are addressed; results will be highlighted: :)  
ft:mark('//p[.//text() contains text 'real']')

Note that BaseX does not support the ignore option (without content) of the W3C XQuery Full Text 1.0 Recommendation. This means that it is not possible to ignore descendant element content, such as footnotes or other material that does not belong to the same logical text flow. Here is an example document:
The ignore option would enable you to search for the string “illustrative purposes”:

//p[. contains text 'illustrative purposes' without content note]

For more examples, see XQuery and XPath Full Text 1.0 Use Cases.

As BaseX does not support the ignore option, it raises error FTST0007 when it encounters without content in a full-text contains expression.

**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**

*Introduced with Version 7.7:*

Another XQuery feature related to natural language processing are Collations. 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. Other collations are completely implementation-defined. In BaseX, the following namespace syntax is supported to specify collations:

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. For example, in German, “Ä” and “Ä” are considered primary differences, “Ä” and “ä” are secondary differences, “Ä” and “À” and “#x308;” are tertiary differences, and “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>

**Examples**

If a default collation is specified, it applies to all collation-dependent string operations in the query. The following expression yields true:

```
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!:

```
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:

\[
\begin{align*}
distinct\text{-}values\left(\text{"a", "á", "à"}, \text{"?lang=it-IT;strength=primary"}\right), \\
index\text{-}of\left(\text{"a", "á", "à"}, \text{"a"}, \text{"?lang=it-IT;strength=primary"}\right)
\end{align*}
\]

Changelog

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 19. Full-Text: Japanese

Read this entry online in the BaseX Wiki.

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.

```
##    ##,##,*,*,########,###,##,##,##
##    ###,*,*,*,####,###,#,#,#
#      ##,##,*,*,*,*,#,##,##
#      ##,###,##,*,*,*,#,##,##
#      ##,###,##,*,*,*,#,##,##
#      ##,###,##,*,*,*,#,##,##
#      ##,###,##,*,*,*,#,##,##
#      ##,###,##,*,*,*,#,##,##
```

The element of the decomposed part is called "Surface", the content analysis is called "Morpheme". The Morpheme component is built as follows:

```
(POS, subtyping POS 1, subtyping POS 2, subtyping POS 3, inflections, use type, prototype, reading, pronunciation)
```

Of these, the surface is used as a token. Also, The contents of analysis of a morpheme are used in indexing and stemming.

 Parsing

During indexing and parsing, the input strings are split into single tokens. In order to reduce the index size and speed up search, the following word classes have been intentionally excluded:

- Mark
- 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 stemming using language 'ja'
'#######' contains text '###' using stemming 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 '## ftand '.##' using wildcards using language 'ja'
Chapter 20. Higher-Order Functions

Read this entry online in the BaseX Wiki.

This page talks about higher-order functions introduced with XQuery 3.0. The BaseX-specific hof module containing some more very useful functions can be found at Higher-Order Functions Module.

Version 7.7: In the upcoming version of the XQuery Functions and Operators specification, some functions will be modified! Function arguments are now placed last in the function signature. Details are found below.

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, making 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:

```
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

```
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:

```
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:

```
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. `fn:for-each` and `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.

**fn:for-each**

*Updated with Version 7.7:* the function has been renamed, and the arguments have been swapped.

| Signatures | `fn:for-each($seq as item()* , $fun as function(item()) as item()*) as item() *`  
| Old signature: `fn:map($fun as function(item()) as item)* , $seq as item() *` as item() * |
| Summary | Applies the function item $fun to every element of the sequence $seq and returns all of the results as a sequence. |
| Examples |  
• Squaring all numbers from 1 to 10:  

```
fn:for-each(l to 10, math:pow(? , 2))
```

*Result:* 1 4 9 16 25 36 49 64 81 100  

• Applying a list of functions to a string:  

```
let $fs := (  
fn:upper-case#1 ,
fn:substring(?, 4) ,
fn:string-length#1
)
return fn:for-each($fs , function($f) { $f('foobar') })
```

*Result:* FOOBAR bar 6

**XQuery 1.0**

```
declare function local:for-each(  
$seq as item()* ,
$fun as function(item()) as item()*) as item()* {  
for $s in $seq  
return $fun($s)
}
```

**fn:filter**

*Updated with Version 7.7:* the arguments have been swapped.

| Signatures | `fn:filter($seq as item()* , $pred as function(item()) as xs:boolean) as item() * fn:filter($pred as function(item()) as xs:boolean , $seq as item()*) as item() *` |
| Summary | Applies the boolean predicate $pred to all elements of the sequence $seq, returning those for which it returns true(). |
Examples

- All even integers until 10:

\[
\text{fn:filter(1 to 10, function($x) { $x mod 2 eq 0 })}
\]

Result: 2 4 6 8 10

- Strings that start with an upper-case letter:

\[
\text{let $first-upper := function($str) { }
||\text{let $first := fn:substring($str, 1, 1)}
||\text{return $first eq fn:upper-case($first) }
||\text{result := fn:filter(('FooBar', 'foo', 'BAR'), $first-upper)}
\]

Result: FooBar BAR

- Inefficient prime number generator:

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

Result: 2 3 5 7 11 13 17 19

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

\[
\text{declare function local:filter($seq, $pred) { }
||\text{for-each($seq, function($x) { }
||||\text{if($pred($x)) then $x else ()}
||\text{})}
\]

XQuery 1.0

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

fn:for-each-pair

Updated with Version 7.7: the function has been renamed, and the arguments have been swapped.

Signatures

\[
\text{fn:for-each-pair($seq1 as item()*, $seq2 as item()*, $fun as function(item(), item()) as item()*) as item()*  Old signature: fn:map-pairs($fun as function(item(), item()) as item()*, $seq1 as item()*, $seq2 as item()*) as item()*)}
\]

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:

\[
\text{fn:for-each-pair(}
\text{  fn:for-each(1 to 10, function($x) { $x mod 2 })},
\text{  (1, 1, 1, 1, 1),}
\text{  function($a, $b) { $a + $b })}
\]
Higher-Order Functions

Result: 2 1 2 1 2

• Line numbering:

```xquery
let $number-lines := function($str) {
    fn:string-join(
        fn:for-each-pair(
            1 to 1000,
            tokenize($str, '\r?\n|\r'),
            concat(?,' ':'?
        ),
        '
    )
    return $number-lines('hello world,
how are you?')
}
```

Result:

```
1: hello world,
2: how are you?
```

• Checking if a sequence is sorted:

```xquery
let $is-sorted := function($seq) {
    every $b in
        fn:for-each-pair(
            $seq,
            fn:tail($seq),
            function($a, $b) { $a le $b }
        )
    satisfies $b
    return ($is-sorted(1 to 10),
        $is-sorted((1, 2, 42, 4, 5))
    )
}
```

Result: true false

### XQuery 1.0

```xquery
declare function local:for-each-pair(
    $seq1 as item()*,
    $seq2 as item()*,
    $fun as function(item(), item()) as item()*
) as item()* {
    for $pos in 1 to min((count($seq1), count($seq2)))
    return $fun($seq1[$pos], $seq2[$pos])
}
```

### 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 with 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**

*Updated with Version 7.7:* the $seq$ and $fun$ arguments have been swapped.

<table>
<thead>
<tr>
<th>Signatures</th>
<th>fn:fold-left($seq as item()<em>, $seed as item()</em>, $fun as function(item()<em>, item()) as item()</em>) as item()*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Old signature</td>
<td>fn:fold-left($fun as function(item()<em>, item()) as item()</em>, $seed as item()<em>, $seq as item()</em>) as item()*</td>
</tr>
</tbody>
</table>

**Summary**
The left fold traverses the sequence from the left. The query `fn:fold-left(1 to 5, 0, $f)` for example would be evaluated as:

```xquery
$f($f($f($f(0, 1), 2), 3), 4), 5)
```

**Examples**

- **Product of a sequence of integers:**

```xquery
let $product := fn:fold-left(? , 1,
  function($result, $i) { $result * $i })
return $product(1 to 5)
```

*Result:* 120

- **Illustrating the evaluation order:**

```xquery
fn:fold-left(1 to 5, '$seed',
  concat('$f(', ?, ', ', ?, ', ')')
)
```

*Result:* `$f($f($f($f($seed, 1), 2), 3), 4), 5)`

- **Building a decimal number from digits:**

```xquery
let $from-digits := fold-left(? , 0,
  function($n, $d) { 10 * $n + $d })
return ($from-digits(1 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:
Higher-Order Functions

```
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
    )
};
```

**fn:fold-right**

*Updated with Version 7.7:* the $seq and $fun arguments have been swapped.

**Signatures**

```
fn:fold-right($seq as item()*, $seed as item()*, $fun as function(item(), item()*) as item()*)
```

**Old signature:**

```
fn:fold-right($fun as function(item(), item()*) as item()*, $seed as item()*, $seq as item()*) as item()*
```

**Summary**

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

```
$f(1, $f(2, $f(3, $f(4, $f(5, 0))))))
```

**Examples**

- **Product of a sequence of integers:**

  ```
  let $product := fn:fold-right(?,
      function($i, $result) { $result * $i }
  )
  return $product(1 to 5)
  ```

  **Result:** 120

- **Illustrating the evaluation order:**

  ```
  fn:fold-right(1 to 5, '$seed',
      concat('$f(', ?, ' ', ?, ' ', ?, ' '))
  )
  ```

  **Result:** $f(1, $f(2, $f(3, $f(4, $f(5, $seed))))))

- **Reversing a sequence of items:**

  ```
  let $reverse := fn:fold-right(?,
      function($item, $rev) {
        $rev, $item
      }
  )
  return $reverse(1 to 10)
  ```

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

**XQuery 1.0**

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

Note that the order of the arguments of $fun are inverted compared to that in fn:fold-left(...).
Chapter 21. Java Bindings

This article is part of the XQuery Portal. It demonstrates two ways to invoke Java code from XQuery and an extension to make Java code aware of the current context.

The Java Binding feature is an extensibility mechanism which enables developers to directly access Java variables and execute code from XQuery. Java classes are identified by namespaces. The namespace URI must simply contain the fully qualified class name. The URI can optionally be prefixed with the string java: to enforce that the addressed code is written in Java.

If the addressed Java code is not found in the classpath, it first needs to be installed in the Repository.

Namespace Declarations

Java classes can be declared via namespaces. The namespace can then be used to call static functions contained in that class. Variables are represented as function with 0 parameters.

The following example uses Java’s Math class to return the cosine of an angle by calling the static method cos(), and the value of π by addressing the static variable via PI():

```xquery
declare namespace math = "java:java.lang.Math";
math:cos(xs:double(0)), math:PI()
```

The new Expanded QName notation of XQuery 3.0 can be applied as well to directly specify a namespace URI instead of the prefix:

```xquery
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 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. The java: prefix is omitted in the URI:

```xquery
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)
)
```

Function names with dashes will be rewritten to Java’s camel case notation:

```xquery
XQuery: get-contents($x as xs:string)
Java : getContents(String x)
```

Since Version 7.7, strings with invalid XML characters will be rejected by default. 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:

```xquery
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 often 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.

An example (the boolean values returned by set:add() are ignored):

```xquery
import module namespace set = "java.util.HashSet";
let $loop :=
    for $i in 1 to 10000
       return set:add($i)
return set:size()
```

Advantages of this approach are:

- imported code can be executed faster than instances created at runtime via new().
- the work on class instances ensures that queries run in parallel will not cause any concurrency issues (provided that the class contains no static variables or functions).

A drawback is that no arguments can be passed on to the class constructor. As a consequence, the addressed class must provide a constructor with no arguments.

**Context-Awareness**

Java classes can be coupled more closely to the BaseX core library. If an instantiated class inherits the abstract `QueryModule` class of BaseX, it will get access to the `context` variable, which is an instance of the `QueryContext` class. It provides access to all static and dynamic properties of the current 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(<Permission>) 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 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';
<context>{
   context:function-namespace()
}</context>,
<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.*;

/**
 * This example is inherited from the `@link QueryModule` class.
 */
public class ContextModule extends QueryModule {
  /**
   * Returns the default function namespace.
   * @return default function namespace
   */
  @Requires(Permission.NONE)
  @Deterministic
  @ContextDependent
  public Str functionNamespace() {
    return Str.get(context.sc.nsFunc);
  }

  /**
   * 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());
    }
  }
}
```

The result will look as follows:

```xml
<context>http://www.w3.org/2005/xpath-functions</context>
<to-int>42</to-int>
```

Please visit the XQuery 3.0 specification if you want to get more insight into function properties.

**Changelog**

Version 7.2.1

- Added: import of Java modules, context awareness
Chapter 22. Module Library

Read this entry online in the BaseX Wiki.

This article is part of the XQuery Portal.

Beside the standard XQuery Functions, BaseX offers additional function modules, which are listed in the following table. The prefixes of all of these modules 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 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 archive ZIP archives.</td>
<td>archive</td>
<td><a href="http://basex.org/modules/archive">http://basex.org/modules/archive</a></td>
</tr>
<tr>
<td>Client</td>
<td>Executing commands and client 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, convert 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, crypto based on the EXPath Cryptographic module.</td>
<td>crypto</td>
<td><a href="http://expat.org/ns/crypto">http://expat.org/ns/crypto</a></td>
</tr>
<tr>
<td>CSV</td>
<td>Functions for processing CSV input. Version 7.7.2</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 db 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 fetch 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://expat.org/ns/file">http://expat.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>Geo</td>
<td>Functions for processing geospatial data.</td>
<td>geo</td>
<td><a href="http://expat.org/ns/geo">http://expat.org/ns/geo</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>HTTP</td>
<td>Sending HTTP requests, built on the EXPath HTTP module.</td>
<td>http</td>
<td><a href="http://expat.org/ns/http-client">http://expat.org/ns/http-client</a></td>
</tr>
<tr>
<td>Index</td>
<td>Functions for requesting details on database indexes.</td>
<td>index</td>
<td><a href="http://basex.org/modules/index">http://basex.org/modules/index</a></td>
</tr>
<tr>
<td>Module</td>
<td>Description</td>
<td>Prefix</td>
<td>Namespace URI</td>
</tr>
<tr>
<td>-------------</td>
<td>--------------------------------------------------</td>
<td>----------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>Inspection</td>
<td>Functions for extracting internal module information. Version 7.7</td>
<td>inspect</td>
<td><a href="http://basex.org/modules/inspect">http://basex.org/modules/inspect</a></td>
</tr>
<tr>
<td>JSON</td>
<td>Parsing and serializing JSON documents.</td>
<td>json</td>
<td><a href="http://basex.org/modules/json">http://basex.org/modules/json</a></td>
</tr>
<tr>
<td>Map</td>
<td>Functions for handling maps (key/value pairs).</td>
<td>map</td>
<td><a href="http://www.w3.org/2005/xpath-functions/map">http://www.w3.org/2005/xpath-functions/map</a></td>
</tr>
<tr>
<td>Math</td>
<td>Mathematical operations, extending the W3C Working Draft.</td>
<td>math</td>
<td><a href="http://www.w3.org/2005/xpath-functions/math">http://www.w3.org/2005/xpath-functions/math</a></td>
</tr>
<tr>
<td>Output</td>
<td>Functions for simplifying formatted output.</td>
<td>out</td>
<td><a href="http://basex.org/modules/out">http://basex.org/modules/out</a></td>
</tr>
<tr>
<td>Process</td>
<td>Executing system commands from XQuery.</td>
<td>proc</td>
<td><a href="http://basex.org/modules/proc">http://basex.org/modules/proc</a></td>
</tr>
<tr>
<td>Profiling</td>
<td>Functions for profiling code snippets.</td>
<td>prof</td>
<td><a href="http://basex.org/modules/prof">http://basex.org/modules/prof</a></td>
</tr>
<tr>
<td>Random</td>
<td>Functions for creating random numbers.</td>
<td>random</td>
<td><a href="http://basex.org/modules/random">http://basex.org/modules/random</a></td>
</tr>
<tr>
<td>Repository</td>
<td>Installing, deleting and listing packages.</td>
<td>repo</td>
<td><a href="http://basex.org/modules/repo">http://basex.org/modules/repo</a></td>
</tr>
<tr>
<td>SQL</td>
<td>JDBC bridge to access relational databases.</td>
<td>sql</td>
<td><a href="http://basex.org/modules/sql">http://basex.org/modules/sql</a></td>
</tr>
<tr>
<td>Streaming</td>
<td>Functions for handling streamable items. Version 7.7</td>
<td>stream</td>
<td><a href="http://basex.org/modules/stream">http://basex.org/modules/stream</a></td>
</tr>
<tr>
<td>Unit</td>
<td>Unit testing framework. Version 7.7</td>
<td>unit</td>
<td><a href="http://basex.org/modules/unit">http://basex.org/modules/unit</a></td>
</tr>
<tr>
<td>Validation</td>
<td>Validating documents against DTDs or XML Schema files.</td>
<td>validate</td>
<td><a href="http://basex.org/modules/validate">http://basex.org/modules/validate</a></td>
</tr>
<tr>
<td>XQuery</td>
<td>Evaluates new XQuery expressions at runtime.</td>
<td>xquery</td>
<td><a href="http://basex.org/modules/xquery">http://basex.org/modules/xquery</a></td>
</tr>
<tr>
<td>XSLT</td>
<td>Stylesheet transformations, based on Java’s and Saxon’s XSLT processor.</td>
<td>xslt</td>
<td><a href="http://basex.org/modules/xslt">http://basex.org/modules/xslt</a></td>
</tr>
<tr>
<td>ZIP</td>
<td>ZIP functionality, based on the EXPath ZIP module (soon obsolete).</td>
<td>zip</td>
<td><a href="http://expath.org/ns/zip">http://expath.org/ns/zip</a></td>
</tr>
</tbody>
</table>

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>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>RESTXQ</td>
<td>Helper functions for the rest RESTXQ API. Version 7.7</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 session server-side HTTP Sessions.</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 sessions all server-side HTTP Sessions.</td>
<td><a href="http://basex.org/modules/sessions">http://basex.org/modules/sessions</a></td>
</tr>
</tbody>
</table>
Chapter 23. 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.

Importing 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:

import module namespace m = 'http://basex.org/modules/Hello' at 'HelloWorld.xqm';
m:hello("Universe")

Library Module HelloWorld.xqm (in the same directory):

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 modules is placed in the repository, there is no need to specify a location. The following example calls a function from the FunctX module:

import module namespace functx = 'http://www.functx.com';
functx:capitalize-first('test')

Commands

BaseX provides three commands for interaction with the package repository: REPO INSTALL, REPO DELETE, and REPO LIST. Packages can also be managed from within XQuery, using the Repository Module.

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

Modules can also be installed, deleted and listed from within XQuery via the Repository Module.

**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.xqm` subdirectory of the repository):

REPO INSTALL http://files.basex.org/modules/org/basex/modules/Hello/HelloWorld.xqm

Importing the repository module:

```xquery
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 XQuery Library is packaged as follows:

```
expath-pkg.xml
functx/
    functx.xql
    functx.xsl
```

**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 modules are looked up in the repository, their URIs are rewritten to a local file path. The URI transformation has been inspired by Zorba:

1. In the URI authority, the order of all substrings separated by dots is reversed.
2. Dots in the authority and the path are replaced by slashes. If no path exists, a single slash is appended.
3. 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

Changelog

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 24. 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 3.0 document. In BaseX, they can be specified in several ways:

- by including them in the prolog of the XQuery expression,
- by specifying them in the XQuery functions file:write() or fn:serialize(),
- by using the -s flag of the BaseX command-line clients,
- by setting the SERIALIZER option before running a query,
- by setting the EXPORTER option before exporting a database, or
- by 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>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>method</td>
<td>Specifies the serialization method:</td>
<td>xml, xhtml, html, text</td>
<td>xml</td>
<td>method=xml</td>
</tr>
<tr>
<td></td>
<td>• xml, xhtml, html, and text are adopted from the official specification.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• json is specific to BaseX and can be used to output XML nodes as JSON objects (see the JSON Module for more details).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Version 7.7.2: csv is BaseX-specific and can be used to output XML nodes as CSV data (see the CSV Module for more details).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• raw is BaseX-specific, too: Binary data types are output in their raw form, i.e., without modifications. For all other types, the items’ string values are returned. No indentation takes place, and no characters are encoded via entities.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Version 7.7.2: jsonml is deprecated.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>version</td>
<td>Specifies the version of the serialization method.</td>
<td>xml/ xhtml:1.0, 1.1 html: 4.0, 4.01, 5.0</td>
<td>1.0</td>
<td>version=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>
<td>html-version=5.0</td>
</tr>
<tr>
<td>item-separator</td>
<td>Determines a string to be used as item separator.</td>
<td>arbitrary strings, \n, \r\n, \r skipped.</td>
<td>empty</td>
<td>item-separator=&amp;#a;</td>
</tr>
</tbody>
</table>
## Serialization

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
<th>Default</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>encoding</td>
<td>Encoding to be used for outputting the data.</td>
<td>all encodings supported by Java</td>
<td>UTF-8</td>
<td>encoding=US-ASCII</td>
</tr>
<tr>
<td>indent</td>
<td>Adjusts whitespaces to make the output better readable.</td>
<td>yes,no</td>
<td>yes</td>
<td>indent=no</td>
</tr>
<tr>
<td>cdata-section-elements</td>
<td>List of elements to be output as CDATA, separated by whitespaces. Example:</td>
<td></td>
<td></td>
<td>cdata-section-elements=text</td>
</tr>
<tr>
<td>omit-xml-declaration</td>
<td>Omit the XML declaration, which is serialized before the actual query result</td>
<td>yes,no</td>
<td>yes</td>
<td>omit=xml-declaration=no</td>
</tr>
<tr>
<td>standalone</td>
<td>Prints or omits the &quot;standalone&quot; attribute in the XML declaration.</td>
<td>yes,no,omit</td>
<td>omit</td>
<td>standalone=yes</td>
</tr>
<tr>
<td>doctype-system</td>
<td>Introduces the output with a document type declaration and the given system identifier. Example:</td>
<td></td>
<td></td>
<td>doctype-system=entities.dtd</td>
</tr>
<tr>
<td>doctype-public</td>
<td>If doctype-system is specified, adds a public identifier. Example:</td>
<td></td>
<td></td>
<td>doctype-public=-//W3C//DTD HTML 4.01//EN&quot;<a href="http://www.w3.org/TR/html4/strict.dtd%3E">http://www.w3.org/TR/html4/strict.dtd&gt;</a></td>
</tr>
<tr>
<td>undeclare-prefixes</td>
<td>Undeclares prefixes in XML 1.1.</td>
<td>yes,no</td>
<td>no</td>
<td>undeclare-prefixes=yes</td>
</tr>
<tr>
<td>normalization-form</td>
<td>Specifies a normalization form. BaseX supports NFC, none Form C (NFC).</td>
<td></td>
<td>NFC</td>
<td>normalization-form=none</td>
</tr>
<tr>
<td>media-type</td>
<td>Specifies the media type.</td>
<td></td>
<td>application/xml</td>
<td>media-type=text/plain</td>
</tr>
<tr>
<td>use-character-maps</td>
<td>Defines character mappings (not supported).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>byte-order-mark</td>
<td>Prints a byte-order-mark before starting serialization.</td>
<td>yes,no</td>
<td>no</td>
<td>byte-order-mark=yes</td>
</tr>
<tr>
<td>escape-uri-attributes</td>
<td>Escapes URI information in certain HTML attributes. Example:</td>
<td>yes,no</td>
<td>no</td>
<td>escape-uri-attributes=yes, method=html</td>
</tr>
<tr>
<td>include-content-type</td>
<td>Includes a meta content-type element if the result is output as HTML.</td>
<td>yes,no</td>
<td>no</td>
<td>include-content-type=yes, method=html</td>
</tr>
</tbody>
</table>

BaseX provides some additional, implementation-specific serialization parameters (the `csv` and `json` parameters may be subject to change):
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Allowed</th>
<th>Default</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>csv-header</td>
<td>Indicates if the input contains a header row. <em>Version 7.7.2</em></td>
<td>yes, no</td>
<td>no</td>
<td>csv-header=yes</td>
</tr>
<tr>
<td>csv-separator</td>
<td>Defines the character which separates fields in a single row. <em>Version 7.7.2</em></td>
<td></td>
<td></td>
<td>csv-separator=	</td>
</tr>
<tr>
<td>json-spec</td>
<td>Determines the used JSON specification. <em>Version 7.7.2</em> RFC4627, ECMA-262, liberal</td>
<td>RFC4627</td>
<td></td>
<td>json-spec=liberal</td>
</tr>
<tr>
<td>json-unescape</td>
<td>Determines whether escape sequences (marked by a backslash) in the input are expanded. <em>Version 7.7.2</em></td>
<td>yes, no</td>
<td>yes</td>
<td>json-unescape=no</td>
</tr>
<tr>
<td>json-format</td>
<td>Determines the conversion format. <em>Version 7.7.2</em> json, jsonml</td>
<td>json</td>
<td></td>
<td>json-format=jsonml</td>
</tr>
<tr>
<td>format</td>
<td>Turns output formatting on/off, including the conversion of special characters to entities and insertion of item separators.</td>
<td>yes, no</td>
<td>yes</td>
<td>format=no</td>
</tr>
<tr>
<td>tabulator</td>
<td>Uses tab characters (\t) for indenting elements.</td>
<td>yes, no</td>
<td>no</td>
<td>tabulator=yes</td>
</tr>
<tr>
<td>indents</td>
<td>Specifies the number of characters to be indented. positive number</td>
<td>2</td>
<td></td>
<td>indents=1, tabulator=yes</td>
</tr>
<tr>
<td>wrap-prefix, wrap-uri</td>
<td>Specifies a prefix and/or URI for wrapping the query results.</td>
<td></td>
<td></td>
<td>wrap-prefix=rest, wrap-uri=<a href="http://basex.org/rest">http://basex.org/rest</a></td>
</tr>
<tr>
<td>newline</td>
<td>Specifies the type of newline to be used as end-of-line marker. \n, \r\n, \r</td>
<td>system dependent</td>
<td></td>
<td>newline=\r\n</td>
</tr>
<tr>
<td>separator</td>
<td>Determines the string to be used as item separator (deprecated, replaced with item-separator). \n, \r\n, \r, arbitrary strings</td>
<td>single space</td>
<td></td>
<td>separator=\n</td>
</tr>
</tbody>
</table>

**Changelog**

**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 25. XQuery

Read this entry online in the BaseX Wiki.

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**
Features of the upcoming XQuery 3.0 Recommendation.

**Module Library**
Additional functions included in the internal modules.

**Repository**
Install and manage XQuery and Java modules.

**Java Bindings**
Accessing and calling Java code from XQuery.

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

**Updates**
Updating databases and local resources via XQuery Update.

**Serialization**
Serialization parameters supported by BaseX.

**Errors**
Errors raised by XQuery expressions.
Chapter 26. 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 the upcoming XQuery 3.0 and XPath 3.0 Recommendations. All extensions are already available in the latest versions of BaseX.

Enhanced FLWOR Expressions

*Introduced with Version 7.7:*

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:**

```
for $country in db:open('factbook')//country
    where $country/@population > 100000000
    let $name := $country/name[1]
    for $city in $country//city.population > 1000000
        group by $name
    return <country name='{{ $name }}'>{ $city/name }</country>
```

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

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

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

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

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.

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:**

```
(1 to 10) ! element node { . },
(1 to 10) ! FLWOR notation :
for $i in 1 to 10
```
A map operator is defined to be part of a path expression, which may now be mixed of path and map operators. In contrast to the map operator, the results of the map operator will not be made duplicate-free and returned in document order.

**Group By**

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

**Example:**

```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:

```xquery
...
return element { $income } { $ppl }
```

**Result:**

```xml
<challenge>
  <person id="person0">
    <name>Kasidi Trewell</name>
  ...
  </person>
  ...
</challenge>
```

**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

**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**:

```xquery
let $f := function($x, $y) { $x + $y }
return $f(17, 25)
```

**Result:** 42
• Getting the function item of an existing (built-in oder 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:

\[
\text{let } \$f := \text{math:pow}\#2 \\
\text{return } \$f(5, 2)
\]

\text{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.

\[
\text{let } \$f := \text{fn:substring}(?, 1, 3) \\
\text{return } ( \\
\hspace{1em} \$f('foo123'), \\
\hspace{1em} \$f('bar456') \\
) 
\]

\text{Result: } \text{foo bar}

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

\section*{Expanded QNames}

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

\textbf{Examples:}

• \texttt{Q\{http://www.w3.org/2005/xpath-functions/math\}pi()} returns the number \( \pi \)

• \texttt{Q\{java:java.io.FileOutputStream\}new("output.txt")} creates a new Java file output stream

The syntax differed in older versions of the XQuery 3.0 specification, in which the prefixed namespace URI was quoted:

• \texttt{"http://www.w3.org/2005/xpath-functions/math":pi()}

• \texttt{"java:java.io.FileOutputStream":new("output")}

\section*{Namespace Constructors}

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

\texttt{element node \{ namespace pref \{ 'http://url.org/' \} \}}

\section*{String Concatenations}

Two vertical bars \(||\) (also names \textit{pipe characters}) can be used to concatenate strings. This operator is a shortcut for the \textit{fn:concat()} function.

\texttt{''Hello'' \|| ''Universe''}

\section*{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.

\texttt{declare variable \$user external := "admin"; \\
"User:"}, \$user
Serialization

Serialization parameters can now be defined within 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:

```
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:

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

Functions

BaseX supports all functions that have been added in Version 3.0 of the XQuery Functions and Operators Working Draft. The new functions are listed below:

- `math:pi()`, `math:sin()`, and many others (see Math Module)
- `fn:analyze-string()`
- `fn:available-environment-variables()`
- `fn:element-with-id()`
• fn:environment-variable()
• fn:filter()
• fn:fold-left()
• fn:fold-right()
• 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:map()
• fn:map-pairs()
• 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() with 0 arguments
• fn:string-join() with 1 argument
• fn:nodeName() with 0 arguments
• `fn:round()` with 2 arguments
• `fn:data()` with 0 arguments

Changelog

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 27. XQuery Errors

Read this entry online in the BaseX Wiki.

This article is part of the XQuery Portal. It summarizes all error codes that may be thrown by the BaseX XQuery processor.

As the original specifications are rather bulky and meticulous, 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.

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.

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></td>
</tr>
<tr>
<td>BASX0001</td>
<td>The current user has insufficient permissions to execute an expression.</td>
<td>file:delete('file.txt'): Create rights needed.</td>
</tr>
<tr>
<td>BASX0002</td>
<td>The specified database option is unknown.</td>
<td>declare option db:xyz &quot;no&quot;; 1</td>
</tr>
<tr>
<td>BASX0003</td>
<td>Errors related to RESTXQ.</td>
<td>%restxq:GET('x')</td>
</tr>
</tbody>
</table>

Additional, module-specific error codes are listed in the descriptions of the query modules.

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 $i</td>
</tr>
<tr>
<td>XPST0005</td>
<td>An expression will never result, no matter what input will be processed.</td>
<td>doc('input')/..</td>
</tr>
<tr>
<td>XPST0008</td>
<td>A variable or type name is used that has not been defined in the current scope.</td>
<td>$a--- element(*, x)</td>
</tr>
<tr>
<td>XPST0017</td>
<td>The specified function is unknown, or it uses the wrong 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 1 instance of x &quot;test&quot; cast as xs:itr</td>
<td></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>1 instance of x &quot;test&quot; cast as xs:itr</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 }</td>
</tr>
<tr>
<td>XQST0009</td>
<td>The query imports a schema (schema import is not supported by BaseX).</td>
<td></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>Code</td>
<td>Description</td>
<td>Example</td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>XQST0031</td>
<td>The specified XQuery version is not specified.</td>
<td><code>xquery version &quot;9.9&quot;; ()</code></td>
</tr>
<tr>
<td>XQST0032</td>
<td>The base URI was declared more than once.</td>
<td><code>declare base-uri ...</code></td>
</tr>
<tr>
<td>XQST0033</td>
<td>A namespace prefix was declared more than once.</td>
<td><code>declare namespace a=&quot;a&quot;; declare namespace a=&quot;b&quot;;</code></td>
</tr>
<tr>
<td>XQST0034</td>
<td>A function was declared more than once.</td>
<td><code>declare function local:a() { 1 }; declare function local:a() { 2 }; local:a()</code></td>
</tr>
<tr>
<td>XQST0038</td>
<td>The default collation was declared more than once.</td>
<td><code>declare default collation ...</code></td>
</tr>
<tr>
<td>XQST0039</td>
<td>Two or more parameters in a user-defined function have the same name.</td>
<td><code>declare function local:fun($a, $a) { $a * $a }; local:fun(1,2)</code></td>
</tr>
<tr>
<td>XQST0040</td>
<td>Two or more attributes in an element have the same node name.</td>
<td><code>&lt;elem a=&quot;1&quot; a=&quot;12&quot;/&gt;</code></td>
</tr>
<tr>
<td>XQST0045</td>
<td>A user-defined function uses a reserved namespace.</td>
<td><code>declare function fn:fun() { 1 }; ()</code></td>
</tr>
<tr>
<td>XQST0047</td>
<td>A module was defined more than once.</td>
<td><code>import module ...</code></td>
</tr>
<tr>
<td>XQST0048</td>
<td>A module declaration does not match the namespace of the specified module.</td>
<td><code>import module namespace invalid=&quot;uri&quot;; 1</code></td>
</tr>
<tr>
<td>XQST0049</td>
<td>A global variable was declared more than once.</td>
<td><code>declare variable $a := 1; declare variable $a := 1; $a</code></td>
</tr>
<tr>
<td>XQST0054</td>
<td>A global variable depends on itself. This may be triggered by a circular variable definition.</td>
<td><code>declare variable $a := local:a(); declare function local:a() { $a }; local:a(); local:fun(1,2)</code></td>
</tr>
<tr>
<td>XQST0055</td>
<td>The mode for copying namespaces was declared more than once.</td>
<td><code>declare copy-namespaces ...</code></td>
</tr>
<tr>
<td>XQST0057</td>
<td>The namespace of a schema import may not be empty.</td>
<td><code>import schema &quot;&quot;; ()</code></td>
</tr>
<tr>
<td>XQST0059</td>
<td>The schema or module with the specified namespace cannot be found or processed.</td>
<td><code>import module &quot;unknown&quot;; ()</code></td>
</tr>
<tr>
<td>XQST0060</td>
<td>A user-defined function has no namespace.</td>
<td><code>declare default function namespace &quot;&quot;; declare function x() { 1 }; 1</code></td>
</tr>
<tr>
<td>XQST0065</td>
<td>The ordering mode was declared more than once.</td>
<td><code>declare ordering ...</code></td>
</tr>
<tr>
<td>XQST0065</td>
<td>The default namespace mode for elements or functions was declared more than once.</td>
<td><code>declare default element namespace ...</code></td>
</tr>
<tr>
<td>XQST0067</td>
<td>The construction mode was declared more than once.</td>
<td><code>declare construction ...</code></td>
</tr>
<tr>
<td>XQST0068</td>
<td>The mode for handling boundary spaces was declared more than once.</td>
<td><code>declare boundary-space ...</code></td>
</tr>
<tr>
<td>XQST0069</td>
<td>The default order for empty sequences was declared more than once.</td>
<td><code>declare default order empty ...</code></td>
</tr>
<tr>
<td>XQST0070</td>
<td>A namespace declaration overwrites a reserved namespace.</td>
<td><code>declare namespace xml=&quot;&quot;; ()</code></td>
</tr>
<tr>
<td>XQST0071</td>
<td>A namespace is declared more than once in an element constructor.</td>
<td><code>&lt;a xmlns=&quot;uri1&quot; xmlns=&quot;uri2&quot;/&gt;</code></td>
</tr>
<tr>
<td>XQST0075</td>
<td>The query contains a validate expression (validation is not supported by BaseX).</td>
<td><code>validate strict { () }</code></td>
</tr>
<tr>
<td>XQST0076</td>
<td>A group by or order by clause specifies an unknown collation.</td>
<td><code>for $i in 1 to 10 order by $i</code></td>
</tr>
</tbody>
</table>
XQuery Errors

XQST0079 A pragma was specified without the expression that is (# xml:a #) ()
to be evaluated.

XQST0085 An empty namespace URI was specified. <pref:elem xmlns:pref=""/>

XQST0087 An unknown encoding was specified. Note that the xquery version "1.0" encoding "a b"; ()

XQST0088 An empty module namespace was specified. import module ""; ()

XQST0089 Two variables in a for or let clause have the same name. for $a at $a in 1 return $i

XQST0090 A character reference specifies an invalid character. "&#0;"

XQST0093 A module depends on itself. This may be triggered by a import module ...
circular module definition.

XQST0094 group by references a variable that has not been declared before. for $a in 1 group by $b return $a

XQST0097 A decimal-format property is invalid. declare default decimal-format digit = "xxx"; 1

XQST0098 A single decimal-format character was assigned to multiple properties. declare default decimal-format digit = "%"; 1

XQST0099 The context item was declared more than once. declare context item ...

XQST0105 An annotation has been declared twice in a variable or function declaration. declare %updating %updating function ...

XQST0108 Output declarations may only be specified in the main module. Module: declare output ...

XQST0109 The specified serialization parameter is unknown. declare option output:unknown "..."; 1

XQST0110 A serialization parameter was specified more than once in the output declarations. declare option output:indent "no"; declare option output:indent "no"; 1

XQST0111 A decimal format was declared more than once. declare decimal-format ...

XQST0113 Context item values may only be in the main module. Module: declare context item := 1;

XQST0114 A decimal-format property has been specified more than once. declare decimal-format EN NaN="!" NaN="?"; ()

Type Errors

Error Codes: XPTY, XQTY

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Examples</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 be raised both statically (during query compilation) or dynamically (at runtime).</td>
</tr>
<tr>
<td>XQTY0018</td>
<td>The result of the last step in a path expression contains both nodes and atomic values.</td>
<td>doc('input.xml')/*, 1)</td>
</tr>
<tr>
<td>XQTY0019</td>
<td>The result of a step (other than the last step) in a path expression contains an atomic values.</td>
<td>(1 to 10)/*</td>
</tr>
<tr>
<td>XQTY0024</td>
<td>An attribute node cannot be bound to its parent element.</td>
<td>&lt;elem&gt;text { attribute a as other nodes of a different type were specified before. ( &quot;val&quot; ) } &lt;/elem&gt;</td>
</tr>
</tbody>
</table>
XQuery Errors

XQTY0105 A function item has been specified as content of an element.

<X>{ false#0 } </X>

Dynamic Errors

Error Codes: XPDY, XQDY

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>XPDY0002</td>
<td>No value has been defined for an external variable.</td>
<td>declare variable $x external;</td>
</tr>
<tr>
<td></td>
<td>or no context item has been set before the query was executed.</td>
<td>$x descendant::*</td>
</tr>
<tr>
<td>XPDY0050</td>
<td>The operand type of a treat expression does not match the type of the argument, or the root of the context item must be a document node.</td>
<td>&quot;string&quot; treat as xs:int</td>
</tr>
<tr>
<td>XQDY0025</td>
<td>Two or more attributes in a constructed element have the same node name.</td>
<td>element x { attribute a { &quot;&quot; } attribute a { &quot;&quot; } }</td>
</tr>
<tr>
<td>XQDY0026</td>
<td>The content of a computed processing instruction contains &quot;&gt;&quot;.</td>
<td>processing-instruction pi { &quot;?&gt;&quot; }</td>
</tr>
<tr>
<td>XQDY0041</td>
<td>The name of a processing instruction is invalid.</td>
<td>processing-instruction { &quot;1&quot; } { &quot;&quot; }</td>
</tr>
<tr>
<td>XQDY0044</td>
<td>The node name of an attribute uses reserved prefixes or namespaces.</td>
<td>attribute xmlns { &quot;etc&quot; }</td>
</tr>
<tr>
<td>XQDY0064</td>
<td>The name of a processing instruction equals &quot;XML&quot; (case insensitive).</td>
<td>processing-instruction xml { &quot;etc&quot; }</td>
</tr>
<tr>
<td>XQDY0072</td>
<td>The content of a computed comment contains &quot;--&quot; or ends with &quot;--&quot;.</td>
<td>comment { &quot;one -- two&quot; }</td>
</tr>
<tr>
<td>XQDY0074</td>
<td>The name of a computed attribute or element is invalid, or uses an unbound prefix.</td>
<td>element { &quot;x y&quot; } { &quot;&quot; }</td>
</tr>
<tr>
<td>XQDY0095</td>
<td>A sequence with more than one item was bound to a group by clause.</td>
<td>let $a := (1,2) group by $a return $a</td>
</tr>
<tr>
<td>XQDY0096</td>
<td>The node name of an element uses reserved prefixes or namespaces.</td>
<td>element { QName(&quot;uri&quot;, &quot;xml:n&quot;) } { }</td>
</tr>
<tr>
<td>XQDY0101</td>
<td>Invalid namespace declaration.</td>
<td>namespace xmlns { 'x' }</td>
</tr>
<tr>
<td>XQDY0102</td>
<td>Duplicate namespace declaration.</td>
<td>element x { namespace a { 'b' }, namespace a { 'c' } }</td>
</tr>
</tbody>
</table>

Functions Errors

Error Codes: FOAR, FOCA, FOCCH, 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;P1Y&quot;) * xs:double(&quot;NaN&quot;)</td>
</tr>
<tr>
<td>XQuery Errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>FOCH0001</strong> A codepoint was specified that does not represent a valid XML character.</td>
<td>codepoints-to-string(0)</td>
<td></td>
</tr>
<tr>
<td><strong>FOCH0002</strong> An unsupported collation was specified in a function.</td>
<td>compare('a', 'a', 'unknown')</td>
<td></td>
</tr>
<tr>
<td><strong>FOCH0003</strong> An unsupported normalization form was specified in a function.</td>
<td>normalize-unicode('a', 'unknown')</td>
<td></td>
</tr>
<tr>
<td><strong>FODC0001</strong> The argument specified in <code>fn:id()</code> or <code>fn:idref()</code> must have a document node as root.</td>
<td>id(&quot;id0&quot;, &lt;xml/&gt;)</td>
<td></td>
</tr>
<tr>
<td><strong>FODC0002</strong> The specified document resource cannot be retrieved.</td>
<td>doc(&quot;unknown.xml&quot;)</td>
<td></td>
</tr>
<tr>
<td><strong>FODC0004</strong> The specified collection cannot be retrieved.</td>
<td>collection(&quot;unknown&quot;)</td>
<td></td>
</tr>
<tr>
<td><strong>FODC0005</strong> The specified URI to a document resource is invalid.</td>
<td>doc(&quot;&lt;xml/&gt;&quot;)</td>
<td></td>
</tr>
<tr>
<td><strong>FODC0006</strong> The string passed to <code>fn:parse-xml()</code> is not well-formed.</td>
<td>parse-xml(&quot;&lt;x/&gt;&quot;)</td>
<td></td>
</tr>
<tr>
<td><strong>FODC0007</strong> The base URI passed to <code>fn:parse-xml()</code> is invalid.</td>
<td>parse-xml(&quot;&lt;x/&gt;&quot;, &quot;:&quot;)</td>
<td></td>
</tr>
<tr>
<td><strong>FODF1280</strong> The name of the decimal format passed to <code>fn:format-number()</code> is invalid.</td>
<td>format-number(1, &quot;0&quot;, &quot;invalid&quot;)</td>
<td></td>
</tr>
<tr>
<td><strong>FODF1310</strong> The picture string passed to <code>fn:format-number()</code> is invalid.</td>
<td>format-number(1, &quot;invalid&quot;)</td>
<td></td>
</tr>
<tr>
<td><strong>FODT0001</strong> An arithmetic duration operation causes an underflow.</td>
<td>xs:date('2000-01-01') + xs:duration('P99999Y')</td>
<td></td>
</tr>
<tr>
<td><strong>FODT0002</strong> A duration declaration or operation causes an underflow.</td>
<td>implicit-timezone() div 0</td>
<td></td>
</tr>
<tr>
<td><strong>FODT0003</strong> An invalid timezone was specified.</td>
<td>adjust-time-to-timezone(xs:time(&quot;01:01:01&quot;), xs:dayTimeDuration(&quot;PT20H&quot;))</td>
<td></td>
</tr>
<tr>
<td><strong>FOER0000</strong> Error triggered by the <code>fn:error()</code> function.</td>
<td>error()</td>
<td></td>
</tr>
<tr>
<td><strong>FOFD1340</strong> The picture string passed to <code>fn:format-date()</code>, <code>fn:format-time()</code> or <code>fn:format-datetime()</code> is invalid.</td>
<td>format-date(current-date(), &quot;]&quot;)</td>
<td></td>
</tr>
<tr>
<td><strong>FOFD1350</strong> The picture string passed to <code>fn:format-time()</code> or <code>fn:format-datetime()</code> specifies a non-available component.</td>
<td>format-time(current-time(), &quot;[Y2]&quot;)</td>
<td></td>
</tr>
<tr>
<td><strong>FONS0004</strong> A function has a QName as argument that specifies an unbound prefix.</td>
<td>resolve-QName(&quot;x:e&quot;, &lt;e/&gt;)</td>
<td></td>
</tr>
<tr>
<td><strong>FOR0001</strong> A value cannot be cast to the required target type.</td>
<td>xs:integer(&quot;A&quot;) l + &lt;x&gt;a&lt;/x&gt;</td>
<td></td>
</tr>
<tr>
<td><strong>FOR0002</strong> The URI passed to <code>fn:resolve-uri()</code> is invalid.</td>
<td>resolve-uri(&quot;:&quot;)</td>
<td></td>
</tr>
<tr>
<td><strong>FOR0003</strong> <code>fn:zero-or-one()</code> was called with more than one item.</td>
<td>zero-or-one((1, 2))</td>
<td></td>
</tr>
<tr>
<td><strong>FOR0004</strong> <code>fn:one-or-more()</code> was called with zero items.</td>
<td>one-or-more()</td>
<td></td>
</tr>
<tr>
<td><strong>FOR0005</strong> <code>fn:exactly-one()</code> was called with zero or more items.</td>
<td>exactly-one((1, 2))</td>
<td></td>
</tr>
<tr>
<td><strong>FOR0006</strong> A wrong argument type was specified in a function call.</td>
<td>sum((1, &quot;string&quot;))</td>
<td></td>
</tr>
<tr>
<td><strong>FOR0008</strong> The arguments passed to <code>fn:dateTime()</code> have different timezones.</td>
<td>dateTime(xs:date(&quot;2001-01-01+01:01&quot;), current-time())</td>
<td></td>
</tr>
<tr>
<td><strong>FORX0001</strong> A function specifies an invalid regular expression flag.</td>
<td>matches('input', 'query', 'invalid')</td>
<td></td>
</tr>
</tbody>
</table>
### XQuery Errors

**FORX0002 A function specifies an invalid regular expression.**
```
matches('input', '[')
```

**FORX0003 A regular expression matches an empty string.**
```
tokenize('input', '.?')
```

**FORX0004 The replacement string of a regular expression is invalid.**
```
replace("input", "match", ")"
```

**FOTY0012 An item has no typed value.**
```
count#1
```

**FOTY0013 Functions items cannot be atomized, have no defined equality, and have no string representation.**
```
data(false#0)
```

**FOTY0014 Function items have no string representation.**
```
string(map ({}))
```

**FOTY0015 Function items cannot be compared.**
```
deep-equal(false#0, true#0)
```

**FOUT1170 Function argument cannot be used to retrieve a text resource.**
```
unparsed-text(':')
```

**FOUT1190 Encoding to retrieve a text resource is invalid or not supported.**
```
unparsed-text('file.txt', 'InvalidEncoding')
```

### 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>omit-xml-declaration is set to yes, and standalone has a value other than omit.</td>
<td></td>
</tr>
<tr>
<td>SEPM0010</td>
<td>method is set to xml, undeclare-prefixes is set to yes, and version is set to 1.0.</td>
<td></td>
</tr>
<tr>
<td>SERE0014</td>
<td>method is set to html, and an invalid HTML character is found.</td>
<td></td>
</tr>
<tr>
<td>SERE0015</td>
<td>method is set to html, 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 fn:put() must be a document node or element.</td>
<td>fn:put(text 'file.txt') ( 1 ),</td>
</tr>
<tr>
<td>FOUP0002</td>
<td>The second argument of fn:put() is not a valid URI.</td>
<td>fn:put(&lt;a/&gt; , '//' )</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 &lt;target/&gt; with &lt;new/&gt;</td>
</tr>
<tr>
<td>XUDY0014</td>
<td>The expression updated by the modify clause was not created by the copy 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>
</tbody>
</table>

119
<table>
<thead>
<tr>
<th>Error Code</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
</table>
| XUDY0016   | In a replace expression, a target is replaced more than once.               | let $a := <x>x</x>/node() return (replace node $a with <a/>, replace node $a with <b/>)

| XUDY0017   | In a replace value of expression, a target is replaced more than once.      | let $a := <x/> return (replace value of node $a with 'a', replace value of node $a with 'a')

| XUDY0021   | The resulting update expression contains duplicate attributes.              | copy $c := <x a='a'/> modify insert node attribute a {""} into $c return $c

| 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 ...

| XUST0004   | The query contains a revalidation 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 instruction is expected as target of an insert attribute a { "" }

| XUTY0007   | Only nodes can be deleted.                                                 | delete node "string"

| XUTY0008   | A single element, text, attribute, comment or processing expression is expected as target of a replace > ) with <b/> | replace node document ( <a/ instruction is expected as target of a replace > ) with <b/>
XQuery Errors

XUTY0010 In a replace expression, in which no attributes are targeted, the replacing nodes must not be attributes as well.

XUTY0011 In the replace expression, in which attributes are targeted, the replacing nodes must be attributes as well.

XUTY0012 In a rename expression, the target nodes must be an element, attribute or processing instruction.

XUTY0013 An expression in the copy clause must return a single node.

XUTY0022 An attribute must not be inserted into a document node.

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 string exclude.</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>
Chapter 28. 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**

```xml
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**

```xml
delete node //node
```

The example query deletes all <node> elements in your database. Note that, in contrast to other updating expressions, the delete expression allows multiple nodes as a target.

**replace**

```xml
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.

```xml
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**

```xml
for $n in //node
```
return rename node $n as 'renamedNode'

All node 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
copy $c := doc('example.xml')//node[@id = 1]
modify rename node $c as 'copyOfNode'
return $c
```

The node 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>
  <text>Copy of: Transform expression example</text>
  <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 (ex:1)
        <extrude>1</extrude>,
        <altitudeMode>relativeToGround</altitudeMode>
    ) before $d//*:coordinates
```
Functions

fn:put()

fn:put() is also part of the XQUF and enables the user 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.

Database Functions

Some additional, updating database functions exist in order to perform updates on document and database level.

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:

insert
insert into
insert into last
insert attribute
insert into first
replace value
rename
put
replace
delete
insert before
db:add()
db:store()
db:replace()
db:rename()
db:delete()
db:optimize()
db:flush()
db:drop()
db:create()

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 on bottom 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 compiling and evaluating the query. As a result, 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 Update

```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**

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. The only way to perform any updating queries and return a result at the same time is to use the BaseX-specific `db:output()` function, which caches the results of its arguments at runtime and returns them after all updates have been processed.

Example: Perform update and return success message.

```xquery
db:output("Update successful."), insert node <c/> into doc('factbook')/mondial
```

If you want to modify temporary nodes in main memory without storing them in a database, you can use the `transform` expression.

**Function Declaration**

To use updating expressions within a function, the `%updating` annotation has to be added to the function declaration. A correct declaration of a function that contains updating expressions (or one that calls updating functions) looks like this:

```xquery
declare %updating function { ... }
```

**Effects**

**Original Files**

In BaseX, all updates are performed on database nodes or in main memory. Update operations thus never affect the original input file. You can, however, use the EXPORT command or the `fn:put()` function to create an updated XML file.

Moreover changes in your database are propagated back to the original input file if the database option WRITEBACK is turned on. This option can also be activated on command line via `-u`. Make sure you back up the original data before running your queries.

**Indexes**

By default, index structures are discarded after an update operation, and their maintenance is left to the user. See more in Indexes: Updates.

**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.
Part VI. XQuery Modules
Chapter 29. Admin Module

This XQuery Module contains functions for performing operations that are restricted to users with Admin Permissions. Existing users can be listed, and soon more.

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:users

| Signatures | admin:users() as element(user)* admin:users($db as xs:string) as element(user)* |
| Summary     | Returns an element sequence, containing all registered users along with their access permissions. If a database $db is specified, users registered for a particular database will be returned. The output of this function is similar to the SHOW USERS command. |
| Examples    | • admin:users() returns <user permission="admin">admin</user> if no additional users have been created.  
• admin:users("factbook") returns all users that have been registered for the specified database. |

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 is similar to the SHOW SESSIONS command. |
| 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)* |
| 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. An empty sequence will be returned if no logging data exists for the specified date. |
| Examples    | • admin:logs() may return <file date="2013-01-23" size="834367"/> if a single log file exists.  
• for $i in admin:logs() return admin:logs($i/@date) lists the contents of all log files. |

Changelog

The Module was introduced with Version 7.5.
Chapter 30. 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 item()) 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:

```xml
<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, which can either be specified

- as children of an <archive:options/> element:

```xml
<archive:options>
  <archive:format value="zip"/>
  <archive:algorithm value="deflate"/>
</archive:options>
```

- as map, which contains all key/value pairs:

```xml
map { "format" := "zip", "algorithm" := "deflate" }
```

Currently, the following combinations are supported (all others will be rejected):

- zip: algorithm may be stored or deflate
Archive Module

• gzip: algorithm may be deflate

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.
FORG0006: an argument has a wrong type.

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:

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: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

An example:

<archive:entry size="1840" last-modified="2009-03-20T03:30:32" compressed-size="672">
  doc/index.html
</archive:entry>

Errors
ARCH9999: archive creation failed for some other reason.

Examples
Sums up the file sizes of all entries of a JAR file:

sum(archive:entries(file:read-binary('jar.zip'))/@size)

archive:options

Signatures
archive:options($archive as xs:base64Binary) as element(archive:options)

Summary
Returns the options of the specified $archive in the format specified by archive:create.

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:
archive:extract-text

Signatures
archive:extract-text($archive as xs:base64Binary) as xs:string*
archive:extract-text($archive as xs:base64Binary, $entries as item(*) as xs:string*) as xs:string*

Summary
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 archive:create (attributes will be ignored). The encoding of the input files can be specified via $encoding.

Errors
ARCH0004: the specified encoding is invalid or not supported, or the string conversion failed.
ARCH9999: archive creation failed for some other reason.

Examples
The following expression extracts all .txt files from an archive:

```xml
<archive:options xmlns:archive="http://basex.org/modules/archive">
  <archive:format value="zip"/>
  <archive:algorithm value="deflate"/>
</archive:options>

define archive:extract-text

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
archive:extract-binary($archive as xs:base64Binary) as xs:base64Binary*
archive:extract-binary($archive as xs:base64Binary, $entries as item*) as xs:base64Binary*

Summary
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 archive:create (attributes will be ignored).

Errors
ARCH9999: archive creation failed for some other reason.

Examples
This example unzips all files of an archive to the current directory:

```xml
let $archive := file:read-binary('archive.zip')
let $entries := archive:entries($archive)
let $contents := archive:extract-binary($archive)
for $entry at $p in $entries
return {
  file:create-dir(replace($entry, "[^/]*",$"")),
  file:write-binary($entry, $contents[$p])
}
```

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
**Archive Module**

<table>
<thead>
<tr>
<th>Examples</th>
<th>This example replaces texts in a Word document:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><img src="https://example.com/code.png" alt="Code" /></td>
</tr>
</tbody>
</table>

**archive:delete**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>archive:delete($archive as xs:base64Binary, $entries as item()* as xs:base64Binary)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Deletes entries from an $archive. The format of $entries is the same as for archive:create.</td>
</tr>
<tr>
<td>Errors</td>
<td>ARCH0005: the entries of the given archive cannot be modified. ARCH9999: archive creation failed for some other reason.</td>
</tr>
</tbody>
</table>

**Examples**

| This example deletes all HTML files in an archive and creates a new file: |
| ![Code](https://example.com/code.png) |

**archive:write**

*Introduced with Version 7.7:*

<table>
<thead>
<tr>
<th>Signatures</th>
<th>archive:write($path as xs:string, $archive as xs:base64Binary) as xs:string* archive:write($path as xs:string, $archive as xs:base64Binary, $entries as item()* as empty-sequence())</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>This convenience function directly writes files of an $archive to the specified directory $path. The entries to be written can be limited via $entries. The format of the argument is the same as for archive:create (attributes will be ignored).</td>
</tr>
<tr>
<td>Errors</td>
<td>FILE0001: a specified path does not exist. ARCH9999: archive creation failed for some other reason.</td>
</tr>
</tbody>
</table>

**Examples**

| This example unzips all files of an archive to the current directory: |
| ![Code](https://example.com/code.png) |

**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 levels.</td>
</tr>
</tbody>
</table>
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.

ARCH0006 The chosen archive format only allows single entries.

ARCH9999 Archive processing failed for some other reason.

**Changelog**

Version 7.7

- Added: `archive:write`

The module was introduced with Version 7.3.
Chapter 31. 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.

Rules

Version 7.7.2: the conversion rules have been updated and aligned with the JSON parser:

The conversion of CSV data is based on the following rules:

1. The resulting document has a <csv/> root node.
2. Rows are represented via <record/> nodes.
3. Fields are either named entry or (if the CSV header is parsed) named by the corresponding column name:
   a. Empty field names are represented by a single underscore (_{_}..._{_}).
   b. Underscore characters are rewritten to two underscores (___).
   c. A character that cannot be represented as NCName character is rewritten to an underscore and its four-digit Unicode.

Functions

csv:parse

Signatures

csv:parse($input as xs:string) as element(csv)
csv:parse($input as xs:string, $options as item()) as element(csv)

Summary

Converts the CSV data specified by $input to XML, and returns the result as element(csv) value. The $options argument can be used to control the way the input is converted. The following options are available:

- separator defines the character which separates columns in a row. By default, this is a comma (,).
- header specifies if the input contains a header row. The default value is false.

Options can either be specified

- as children of an <csv:options/> element; e.g.:

  <csv:options>
    <csv:separator value=';'/>
    ...
  </csv:options>

- or as map, which contains all key/value pairs:
csv:serialize

Signatures

```xml
csv:serialize($input as node(), $options as item()) as xs:string
```

Summary

Serializes the node specified by $input as CSV data, and returns the result as `xs:string`. XML documents can also be serialized as CSV if the Serialization Option "method" is set to "csv". The $options argument can be used to control the way the node is serialized. The following options are available:

- **separator** defines the character which separates columns in a row. By default, this is a comma (,).
- **header** specifies if the input element names are to be interpreted as header names. The default value is false.

Options can either be specified

- as children of an `<csv:options/>` element; e.g.:

  ```xml
  <csv:options>
    <csv:separator value=';'/>
    ...
  </csv:options>
  ```

- or as map, which contains all key/value pairs:

  ```xml
  { 'separator' : ';', ... }
  ```

Errors

BXCS0002: the input cannot be serialized. BXCS0003: the specified separator must be a single character.

Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXCS0001</td>
<td>The input cannot be converted.</td>
</tr>
<tr>
<td>BXCS0002</td>
<td>The node cannot be serialized.</td>
</tr>
<tr>
<td>BXCS0001</td>
<td>The specified separator must be a single character.</td>
</tr>
</tbody>
</table>

Changelog

The module was introduced with Version 7.7.2.
Chapter 32. 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:

```xml
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 a previous 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
Variables and the context item can be declared via $bindings. The specified keys must be QNames or strings, the values can be arbitrary items:

- variables specified as QNames will be directly interpreted as variable name.
- variables specified as xs:string 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.

### 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:

```
let $c := client:connect('localhost', 1984, 'admin', 'admin')
return client:query($c, "$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:

```
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**

<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>
</tbody>
</table>

### Changelog

Version 7.5

- Added: `client:info`

The module was introduced with Version 7.3.
Chapter 33. 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: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.

Examples
- `convert:integer-to-base(-1, 16)` returns the hexadecimal string 'ffffffffffffffff'.
- `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 $str$ contains more than 64 bits of information, the result is truncated arbitrarily.

**Examples**

- `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**

**Signatures**

`convert:integer-to-dateTime($ms as xs:integer) as xs:dateTime`

**Summary**

Converts the specified number of milliseconds since 1 Jan 1970 to an item of type `xs:dateTime`.

**Examples**

- `convert:integer-to-dateTime(0)` returns `1970-01-01T00:00:00Z`.

**convert:dateTime-to-integer**

**Signatures**

`convert:dateTime-to-integer($dateTime as xs:dateTime) as xs:integer`

**Summary**

Converts the specified item of type `xs:dateTime` to the number of milliseconds since 1 Jan 1970.

**Examples**

- `convert:dateTime-to-integer(xs:dateTime('1970-01-01T00:00:00Z'))` returns `0`.

**convert:integer-to-dayTime**

**Signatures**

`convert:integer-to-dayTime($ms as xs:integer) as xs:dayTimeDuration`

**Summary**

Converts the specified number of milliseconds to an item of type `xs:dayTimeDuration`.

**Examples**

- `convert:integer-to-dayTime(1234)` returns `PT1.234S`.

**convert:dayTime-to-integer**

**Signatures**

`convert:dayTime-to-integer($dayTime as xs:dayTimeDuration) as xs:integer`

**Summary**

Converts the specified item of type `xs:dayTimeDuration` to milliseconds represented by an integer.

**Examples**

- `convert:dayTime-to-integer(xs:dayTimeDuration('PT1S'))` returns `1000`.

### 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>
</tbody>
</table>

### Changelog

Version 7.5
Conversion Module

- 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 34. 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://expat.org/ns/crypto namespace, which is statically bound to the crypto prefix. All errors are assigned to the http://expat.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:

```xquery
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**

```
<Signature>
  <SignedInfo>
    <CanonicalizationMethod/>
    <SignatureMethod/>
    <Reference>
      <Transforms/>
      <DigestMethod/>
      <DigestValue/>
  </Reference>
</SignedInfo>
```
Cryptographic Module

- **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` function:

- **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, $ext as item()) as node()`
  - `crypto:generate-signature($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, $Certificate as node()) as node()` crypto:generate-signature($input as node(), $canonicalization as xs:string, $digest as xs:string, $signature as xs:string, $prefix as xs:string, $type as xs:string, $XPath as xs:string, $Certificate as node() as node())

- **Summary**
  - `$canonicalization must either be inclusive-with-comments, inclusive, exclusive-with-comments or exclusive. Default is inclusive-with-comments` 
  - `$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**. $\text{xpath}$ is an arbitrary XPath expression which specifies a subset of the document that is to be signed. $\text{certificate}$ is the digital certificate used to sign the input document. $\text{ext}$ may either be an $\text{xpath}$ expression or a $\text{certificate}$.

### Errors

**CX0001**: the canonicalization algorithm is not supported.<br>
**CX0002**: the digest algorithm is not supported.<br>
**CX0003**: the signature algorithm is not supported.<br>
**CX0004**: the $\text{xpath}$-expression is invalid.<br>
**CX0005**: the root name of $\text{digital-certificate}$ is not 'digital-certificate'.<br>
**CX0007**: the key store is null.<br>
**CX0012**: the key cannot be found in the specified key store.<br>
**CX0023**: the certificate alias is invalid.<br>
**CX0024**: an invalid algorithm is specified.<br>
**CX0025**: an exception occurs while the signing the document.<br>
**CX0026**: an exception occurs during key store initialization.<br>
**CX0027**: an IO exception occurs.<br>
**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>
      </Reference>
    </SignedInfo>
    <SignatureValue>Pl/Jr44WBcdARff2UVYEiwyYW1563XdqnU87nusAIaHgzd+U3SrjVJhPFjL0e0DfJfxVtYzLFAznTYE P3ddeoFmyA==</SignatureValue>
    <KeyInfo>
      <KeyValue>
        <RSAKeyValue>
          <Modulus>rtvpFSbCIE2BJeP1VYLlRIjX10RE5r2+D+JOVKn7AM7V2bcbRDPeqRbjskEz1HWC/N067JdBqH 4/4FPT9bGQ==</Modulus>
          <Exponent>AQB</Exponent>
        </RSAKeyValue>
      </KeyValue>
    </KeyInfo>
  </Signature>
</a>
```

**crypto:validate-signature**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>crypto:validate-signature($input-doc as node()) as xs:boolean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Summary</strong></td>
<td>Checks if the given node contains a <code>Signature</code> element and whether the signature is valid. In this case <code>true</code> is returned. If the signature is invalid the function returns <code>false</code>.</td>
</tr>
</tbody>
</table>
| **Errors** | **CX0015**: the signature element cannot be found.<br>
**CX9994**: an unspecified problem occurs during validation.<br>
**CX9996**: an IO exception occurs during validation. |
| **Example** | **Validates an XML Signature. Query:** |

---

144
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 35. 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

With Version 7.7, the functions no longer accept database nodes as reference. Instead, the name of a database must now be specified.

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**

**Signatures**

```
db:system() as element(system)
```

**Summary**

Returns information on the database system, such as the database path and current database settings. The output is similar to the INFO command.

**Errors**

BXDB0002: The addressed database does not exist or could not be opened.

**db:info**

**Signatures**

```
db:info($db as xs:string) as element(database)
```

**Summary**

Returns meta information on the database $db. The output is similar to the INFO DB command.

**Errors**

BXDB0002: The addressed database does not exist or could not be opened.

**db:list**

**Signatures**

```
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*
```

**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**

**Signatures**

```
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)*
```

```
```
**Summary**
Returns an element sequence with the names of all databases together with their database path, the number of stored resources and the date of modification:

- If a database $db$ is specified, all documents and raw files of the specified database together with their content-type, the modification date and the resource type 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-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.

---

### db:event

**Signatures**

```
db:event($name as xs:string, $query as item()) as empty-sequence()
```

**Summary**
Executes a $query and sends the resulting value to all clients watching the Event with the specified $name. The query may also perform updates; no event will be sent to the client that fired the event.

**Errors**
BXDB0010: the specified event is unknown. SEPM0016: serialization errors occurred while sending the value.

---

### 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**
Returns a sequence with all document nodes contained in the database $db. The document nodes to be returned can be restricted by 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("docs", "one")` returns all documents from the database named docs in the subpath one.

---

### 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 values 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**

\[ \text{db:open-id}($\text{db}$ as xs:string, $\text{id}$ as xs:integer) as node() \]

**Summary**

Opens the database $\text{db}$ and returns the node with the specified $\text{id}$ 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**

\[ \text{db:node-pre}($\text{nodes}$ as node()*) as xs:integer* \]

**Summary**

Returns the **pre** values of the nodes supplied by $\text{nodes}$, 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: $\text{nodes}$ contains a node which is not stored in a database.

**Examples**

- \[ \text{db:node-pre(doc("input"))} \]
  returns 0 if the database input contains a single document.

### db:node-id

**Signatures**

\[ \text{db:node-id}($\text{nodes}$ as node()*) as xs:integer* \]

**Summary**

Returns the **id** values of the nodes supplied by $\text{nodes}$, 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: $\text{nodes}$ contains a node which is not stored in a database.

### db:retrieve

**Signatures**

\[ \text{db:retrieve}($\text{db}$ as xs:string, $\text{path}$ as xs:string) as xs:base64Binary \]

**Summary**

Returns a **binary resource** addressed by the database $\text{db}$ and $\text{path}$ 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**

- \[ \text{declare option output:method 'raw';db:retrieve("DB", "music/01.mp3")} \]
  returns the specified audio file as raw data.
- \[ \text{stream:materialize(db:retrieve("DB", "music/01.mp3")}) \]
  returns a materialized representation of the streamable result.

### db:export

*Added with Version 7.7:*

**Signatures**

\[ \text{db:export}($\text{db}$ as xs:string, $\text{path}$ as xs:string) as empty-sequence() \]
\[ \text{db:export}($\text{db}$ as xs:string, $\text{path}$ as xs:string, $\text{params}$ as item()) as empty-sequence() \]

**Summary**

Exports the specified database $\text{db}$ to the specified file $\text{path}$. Existing files will be overwritten. The $\text{params}$ argument contains serialization parameters (see **Serialization** for more details), which can either be specified.
Database Module

- 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", ... }
```

**Errors**

BXDB0002: The addressed database does not exist or could not be opened.

### Contents

**db:text**

<table>
<thead>
<tr>
<th>Signatures</th>
<th><code>db:text($db as xs:string, $string as item()) as text()*</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns all text nodes of the database $db that have $string as their string value. If available, the value index is used to speed up evaluation.</td>
</tr>
<tr>
<td>Errors</td>
<td>BXDB0002: The addressed database does not exist or could not be opened.</td>
</tr>
<tr>
<td>Examples</td>
<td>• <code>db:text(&quot;DB&quot;, &quot;QUERY&quot;)/..</code> returns the parents of all text nodes of the database DB that match the string QUERY.</td>
</tr>
</tbody>
</table>

**db:text-range**

<table>
<thead>
<tr>
<th>Signatures</th>
<th><code>db:text-range($db as xs:string, $min as xs:string, $max as xs:string) as text()*</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns all text nodes of the database $db that are located in between the $min and $max strings. If available, the value index is used to speed up evaluation.</td>
</tr>
<tr>
<td>Errors</td>
<td>BXDB0002: The addressed database does not exist or could not be opened.</td>
</tr>
<tr>
<td>Examples</td>
<td>• <code>db:text-range(&quot;DB&quot;, &quot;2000&quot;, &quot;2001&quot;)</code> returns all text nodes of the database DB that are found in between 2000 and 2001.</td>
</tr>
</tbody>
</table>

**db:attribute**

<table>
<thead>
<tr>
<th>Signatures</th>
<th><code>db:attribute($db as xs:string, $string as item()) as attribute()*</code></th>
</tr>
</thead>
<tbody>
<tr>
<td><code>db:attribute($db as xs:string, $string as item(), $attname as xs:string) as attribute()*</code></td>
<td></td>
</tr>
<tr>
<td>Summary</td>
<td>Returns all attribute nodes of the database $db that have $string as string value. If available, the value index is used to speed up evaluation. If $attname is specified, the resulting attribute nodes are filtered by their attribute name.</td>
</tr>
<tr>
<td>Errors</td>
<td>BXDB0002: The addressed database does not exist or could not be opened.</td>
</tr>
<tr>
<td>Examples</td>
<td>• <code>db:attribute(&quot;DB&quot;, &quot;QUERY&quot;, &quot;id&quot;)/..</code> returns the parents of all id attribute nodes of the database DB that have QUERY as string value.</td>
</tr>
</tbody>
</table>

**db:attribute-range**

<table>
<thead>
<tr>
<th>Signatures</th>
<th><code>db:attribute-range($db as xs:string, $min as xs:string, $max as xs:string) as attribute()*</code></th>
</tr>
</thead>
<tbody>
<tr>
<td><code>db:attribute-range($db as xs:string) as attribute()*</code></td>
<td></td>
</tr>
<tr>
<td>Summary</td>
<td>Returns all attribute nodes of the database $db that are located in between the $min and $max strings. If available, the value index is used to speed up evaluation.</td>
</tr>
<tr>
<td>Errors</td>
<td>BXDB0002: The addressed database does not exist or could not be opened.</td>
</tr>
<tr>
<td>Examples</td>
<td>• <code>db:attribute-range(&quot;DB&quot;, &quot;2000&quot;, &quot;2001&quot;)</code> returns all attribute nodes of the database DB that are found in between 2000 and 2001.</td>
</tr>
</tbody>
</table>
$\min$ as $\text{xs:string}$, $\max$ as $\text{xs:string}$, $\text{attname}$ as $\text{xs:string}$) as attribute()*

| Summary | Returns all attributes of the database $\text{db}$, the string values of which are larger than or equal to $\min$ and smaller or equal to $\max$. If available, the value index is used to speed up evaluation. |
| Errors | BXDB0002: The addressed database does not exist or could not be opened. |
| Examples | • $\text{db:attribute-range("DB", "id456", "id473", 'id')}$ returns all $\text{id}$ attributes of the database $\text{DB}$ that have a string value in between $\text{id456}$ and $\text{id473}$. |

**db:fulltext**

| Signatures | $\text{db:fulltext($db$ as $\text{xs:string}$, $\text{terms}$ as $\text{xs:string}$) as text()*} |
| Summary | Returns all text nodes from the full-text index of the database $\text{db}$ that contain the text specified as $\text{terms}$. The options used for 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. |
| Errors | BXDB0002: The addressed database does not exist or could not be opened. BXDB0004: the full-text index is not available. |
| Examples | • $\text{db:fulltext("DB", "QUERY")}$ returns all text nodes of the database $\text{DB}$ that contain the string $\text{QUERY}$. |

**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**

*Updated with Version 7.7: $\text{options}$ argument added*

| Signatures | $\text{db:create($db$ as $\text{xs:string}$) as empty-sequence()}$ $\text{db:create($db$ as $\text{xs:string}$, $\text{inputs}$ as $\text{item()}*$) as empty-sequence()}$ $\text{db:create($db$ as $\text{xs:string}$, $\text{inputs}$ as $\text{item()}*$, $\text{paths}$ as $\text{xs:string}*$) as empty-sequence()}$ $\text{db:create($db$ as $\text{xs:string}$, $\text{inputs}$ as $\text{item()}*$, $\text{paths}$ as $\text{xs:string}*$, $\text{options}$ as $\text{item()}$) as empty-sequence()}$ |
| Summary | Creates a new database with name $\text{db}$ and adds initial documents specified via $\text{inputs}$ to the specified $\text{paths}$. $\text{inputs}$ may be strings or nodes different than attributes. If the $\text{input source}$ is not a file or a folder, the $\text{paths}$ argument is mandatory. Please note that $\text{db:create}$ will be placed last on the Pending Update List. As a consequence, a newly created database cannot be addressed in the same query. The $\text{options}$ argument can be used to change the indexing behavior. Allowed options are all **Indexing Options** and **Full-Text Options** in lower case. Options can be specified either...

* as children of an $\text{<options/>}$ element, e.g.

```
<options>
  <textindex value='true'/>
  <maxcats value='128'/>
</options>
```

* or as map, which contains all key/value pairs:

```
map { "textindex" := true(), "maxcats" = 128 }
```
**Errors**  
FODC0002: $inputs points to an unknown resource. FOUP0001: $inputs is neither string nor a document node. BXDB0007: $db is opened by another process. BXDB0011: $db 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 and creates 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: $db 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()`

**Summary**

Adds documents specified by $input to the database $db and the specified $path. $input may be a string or a node different than attribute. If the $input source is not a file or a folder, $path must be specified.

**Errors**

BXDB0002: The addressed database does not exist or could not be opened. FODC0002: $input points to an unknown resource. FOUP0001: $input is neither string nor a document node.

**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.
- `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**

`db:delete($db as xs:string, $path as xs:string) as empty-sequence()`

**Summary**

Deletes document(s), specified by $path, from the database $db.

**Errors**

BXDB0002: The addressed database does not exist or could not be opened.

**Examples**

- `db:delete("DB", "docs/dir/doc.xml")` deletes the document docs/dir/doc.xml in the database DB.
- `db:delete("DB", "docs/dir")` deletes all documents with paths beginning with docs/dir in the database DB.
### db:optimize

*Updated with Version 7.7: $options argument added*

**Signatures**

<table>
<thead>
<tr>
<th>Signature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>db:optimize($db as xs:string) as empty-sequence()</td>
<td></td>
</tr>
<tr>
<td>db:optimize($db as xs:string, $all as xs:boolean) as empty-sequence()</td>
<td></td>
</tr>
<tr>
<td>db:optimize($db as xs:string, $all as xs:boolean, $options as item()) as empty-sequence()</td>
<td></td>
</tr>
</tbody>
</table>

**Summary**

Optimizes the meta data and indexes of the database $db. If $all is set to true(), the complete database will be rebuilt. The usage of the $options argument is identical to the db:create function, except that the UPDINDEX option is not supported.

**Errors**

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXDB0002</td>
<td>The addressed database does not exist or could not be opened.</td>
</tr>
<tr>
<td>BXDB0008</td>
<td>New document names would be empty.</td>
</tr>
<tr>
<td>BXDB0014</td>
<td>$path points to a directory.</td>
</tr>
<tr>
<td>FODC0002</td>
<td>$input is a string representing a path, which cannot be read.</td>
</tr>
<tr>
<td>FOUP0001</td>
<td>$input is neither a string nor a document node.</td>
</tr>
</tbody>
</table>

**Examples**

- db:optimize("DB") optimizes the database structures of the database DB.
- db:optimize("DB", true(), map { 'ftindex' := true() }) optimizes all database structures of the database DB and creates a full-text index.

### db:rename

**Signatures**

<table>
<thead>
<tr>
<th>Signature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>db:rename($db as xs:string, $path as xs:string, $newpath as xs:string) as empty-sequence()</td>
<td></td>
</tr>
</tbody>
</table>

**Summary**

Renames document(s), specified by $path to $newpath in the database $db.

**Errors**

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXDB0002</td>
<td>The addressed database does not exist or could not be opened.</td>
</tr>
<tr>
<td>BXDB0008</td>
<td>New document names would be empty.</td>
</tr>
<tr>
<td>BXDB0014</td>
<td>$path points to a directory.</td>
</tr>
<tr>
<td>FODC0002</td>
<td>$input is a string representing a path, which cannot be read.</td>
</tr>
</tbody>
</table>

**Examples**

- db:rename("DB", "docs/dir/doc.xml", "docs/dir/newdoc.xml") renames the document docs/dir/doc.xml to docs/dir/newdoc.xml in the database DB.
- db:rename("DB", "docs/dir", "docs/newdir") renames all documents with paths beginning with docs/dir to paths beginning with docs/newdir in the database DB.

### db:replace

**Signatures**

<table>
<thead>
<tr>
<th>Signature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>db:replace($db as xs:string, $path as xs:string, $input as item()) as empty-sequence()</td>
<td></td>
</tr>
</tbody>
</table>

**Summary**

Replaces a document, specified by $path, in the database $db with the content of $input, or adds it as a new document.

**Errors**

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXDB0002</td>
<td>The addressed database does not exist or could not be opened.</td>
</tr>
<tr>
<td>BXDB0014</td>
<td>$path points to a directory.</td>
</tr>
<tr>
<td>FODC0002</td>
<td>$input is a string representing a path, which cannot be read.</td>
</tr>
<tr>
<td>FOUP0001</td>
<td>$input is neither a string nor a document node.</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.

### db:store

**Signatures**

<table>
<thead>
<tr>
<th>Signature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>db:store($db as xs:string, $path as xs:string, $input as item()) as empty-sequence()</td>
<td></td>
</tr>
</tbody>
</table>
### 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.

## db:output

### Signatures
`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.

## db:flush

### Signatures
`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

*Added with Version 7.7:*

### 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

*Added with Version 7.7:*

### 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($db as xs:string) as xs:boolean`
- `db:exists($db as xs:string, $path as xs:string) as xs:boolean`

### Summary
Checks if the database $db or the resource specified by $path exists. false is returned if a database directory has been addressed.

### Examples
- `db:exists("DB")` returns true if the database DB exists.
Database Module

• 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

**Summary**
Checks if the specified resource in the database $db and the path $path exists, and if it is a binary resource.

**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

**Errors**

BXDB0002: The addressed database does not exist or could not be opened.

**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**

BXDB0002: The addressed database does not exist or could not be opened. FODC0002: the addressed resource is not found or cannot be retrieved.

**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.

**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>Database paths cannot be renamed to empty strings.</td>
</tr>
<tr>
<td>BXDB0009</td>
<td>The addressed database id or pre value is out of range.</td>
</tr>
<tr>
<td>BXDB0010</td>
<td>The specified event is unknown.</td>
</tr>
</tbody>
</table>
The name of the specified database is invalid.

A database can only be created once.

The number of specified inputs and paths differs.

Path points to a directory.

**Changelog**

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 36. 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.

The module has initially been inspired by Zorba’s Fetch Module.

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:$text

**Signatures**

fetch:text($uri as xs:string) as xs:string  
fetch:text($uri as xs:string, $encoding as xs:string) as xs:string

**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:binary

**Signatures**

fetch:binary($uri as xs:string) as xs:base64Binary

**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: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 file extension will be guessed based on the file extension.
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>

Examples


Changelog

The module was introduced with Version 7.6.
Chapter 37. File Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions and variables related to file system operations, such as listing, reading, or writing files.

This module is based on the EXPath File Module. With Version 7.7, all returned strings that refer to existing directories will be suffixed with a directory separator.

Conventions

All functions in this module are assigned to the http://expath.org/ns/file namespace, which is statically bound to the file prefix. All errors are assigned to the http://expath.org/ns/error namespace, which is statically bound to the experr prefix.

Read Operations

file:list

Signatures

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*

Summary

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 (,).

Errors

FILE0003: the specified path does not point to a directory.
FILE9999: the operation fails for some other reason.

file:read-binary

Signatures

file:read-binary($path as xs:string) as xs:base64Binary

Summary

Reads the binary content of the file specified by $path and returns it as streamable xs:base64Binary.

Errors

FILE0001: the specified file does not exist.
FILE0004: the specified path is a directory.
FILE9999: the operation fails for some other reason.

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

FILE0001: the specified file does not exist.
FILE0004: the specified path is a directory.
FILE0005: the specified encoding is not supported, or unknown.
FILE9999: 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
FILE0001: the specified file does not exist.FILE0004: the specified path is a directory.FILE0005: the specified encoding is not supported, or unknown.FILE9999: 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
FILE0002: a file with the same path already exists.FILE9999: the operation fails for some other reason.

file:create-temp-dir

Introduced with Version 7.7:

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
FILE0003: the specified directory points to a file.FILE9999: the directory could not be created.

file:create-temp-file

Introduced with Version 7.7:

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.

Errors
FILE0003: the specified directory points to a file.FILE9999: the directory could not be created.

file:delete

Signatures
file:delete($path as xs:string) as empty-sequence()
file:delete($path as xs:string, $recursive as xs:boolean) as empty-sequence()

Summary
Recursively deletes a file or directory specified by $path. The optional parameter $recursive specifies whether sub-directories will be deleted, too.
## Errors

<table>
<thead>
<tr>
<th>FILE0001</th>
<th>the specified path does not exist.</th>
</tr>
</thead>
<tbody>
<tr>
<td>FILE9999</td>
<td>the operation fails for some other reason.</td>
</tr>
</tbody>
</table>

## file:write

### Signatures

```
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()
```

### 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.:

  ```xml
  <output:serialization-parameters>
    <output:method value='xml'/>
    <output:cdata-section-elements value="div"/>
    ...
  </output:serialization-parameters>
  ```

- as map, which contains all key/value pairs:

  ```xml
  map { "method" := "xml", "cdata-section-elements" := "div", ... }
  ```

### Errors

<table>
<thead>
<tr>
<th>FILE0003</th>
<th>the parent of specified path is no directory.</th>
</tr>
</thead>
<tbody>
<tr>
<td>FILE0004</td>
<td>the specified path is a directory.</td>
</tr>
<tr>
<td>FILE9999</td>
<td>the operation fails for some other reason.</td>
</tr>
</tbody>
</table>

## file:write-binary

### Signatures

```
file:write-binary($path as xs:string, $value as xs:anyAtomicType) as empty-sequence()
```

### Summary

Writes a binary item (xs:base64Binary, xs:hexBinary) to the specified file. If the file already exists, it will be overwritten.

### Errors

<table>
<thead>
<tr>
<th>FILE0003</th>
<th>the parent of specified path is no directory.</th>
</tr>
</thead>
<tbody>
<tr>
<td>FILE0004</td>
<td>the specified path is a directory.</td>
</tr>
</tbody>
</table>

## file:write-text

### Signatures

```
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()
```

### 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

<table>
<thead>
<tr>
<th>FILE0003</th>
<th>the parent of specified path is no directory.</th>
</tr>
</thead>
<tbody>
<tr>
<td>FILE0004</td>
<td>the specified path is a directory.</td>
</tr>
<tr>
<td>FILE0005</td>
<td>the specified encoding is not supported, or unknown.</td>
</tr>
<tr>
<td>FILE9999</td>
<td>the operation fails for some other reason.</td>
</tr>
</tbody>
</table>

## file:write-text-lines

### Signatures

```
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

- `FILE0003`: the parent of specified path is no directory.
- `FILE0004`: the specified path is a directory.
- `FILE0005`: the specified encoding is not supported, or unknown.
- `FILE9999`: the operation fails for some other reason.

### file:append

**Signatures**

```
file:append($path as xs:string, $items as item()* as empty-sequence())
```

**Summary**

Appends a serialized sequence of items to the specified file. If the file does not exists, a new file is created.

**Errors**

- `FILE0003`: the parent of specified path is no directory.
- `FILE0004`: the specified path is a directory.
- `FILE9999`: the operation fails for some other reason.

### file:append-binary

**Signatures**

```
file:append-binary($path as xs:string, $value as xs:anyAtomicType)
```

**Summary**

Appends a binary item (xs:base64Binary, xs:hexBinary) to the specified file. If the file does not exists, a new one is created.

**Errors**

- `FILE0003`: the parent of specified path is no directory.
- `FILE0004`: the specified path is a directory.

### file:append-text

**Signatures**

```
file:append-text($path as xs:string, $value as xs:string) as empty-sequence()
```

**Summary**

Appends a string to a file specified by `$path`. If the specified file does not exists, a new file is created. The optional parameter `$encoding` defines the output encoding (default: UTF-8).

**Errors**

- `FILE0003`: the parent of specified path is no directory.
- `FILE0004`: the specified path is a directory.
- `FILE0005`: the specified encoding is not supported, or unknown.
- `FILE9999`: the operation fails for some other reason.

### file:append-text-lines

**Signatures**

```
file:append-text-lines($path as xs:string, $values 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**

- `FILE0003`: the parent of specified path is no directory.
- `FILE0004`: the specified path is a directory.
- `FILE0005`: the specified encoding is not supported, or unknown.
- `FILE9999`: the operation fails for some other reason.

### file:copy

**Signatures**

```
file:copy($source as xs:string, $target as xs:string) as empty-sequence()
```

**Summary**

Copies a file 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.
### File Module

#### Errors
- FILE0001: the specified source does not exist.
- FILE0002: the specified source is a directory and the target is a file.
- FILE0003: the parent of the specified target is no directory.
- FILE9999: the operation fails for some other reason.

#### file: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**
- FILE0001: the specified source does not exist.
- FILE0002: the specified source is a directory and the target is a file.
- FILE0003: the parent of the specified target is no directory.
- FILE9999: the operation fails for some other reason.

#### File Properties

#### file: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.

#### file: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.

#### file: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**
- FILE0001: 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 $path.

**Errors**
- FILE0001: the specified file does not exist.
- FILE0004: the specified file points to a directory.

#### Path Functions

#### file:base-name

**Signatures**
```
file:base-name($path as xs:string) as xs:string
file:base-name($path as xs:string, $suffix as xs:string) as xs:string
```

---

162
### Summary

Returns the base-name of the path specified by $path, which is the component after the last directory separator. If $suffix is specified, it will be trimmed from the end of the result.

**file:dir-name**

<table>
<thead>
<tr>
<th>Signatures</th>
<th><code>file:dir-name($path as xs:string) as xs:string</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the parent directory of the path specified by $path, which is the component before the last directory separator.</td>
</tr>
<tr>
<td>Examples</td>
<td>• <code>file:dir-name(static-base-uri())</code> returns the directory of the current XQuery module.</td>
</tr>
</tbody>
</table>

**file:path-to-native**

<table>
<thead>
<tr>
<th>Signatures</th>
<th><code>file:path-to-native($path as xs:string) as xs:string</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Transforms the $path argument to its native representation on the operating system.</td>
</tr>
<tr>
<td>Errors</td>
<td>FILE9999: the specified path cannot be transformed to its native representation.</td>
</tr>
</tbody>
</table>

**file:resolve-path**

<table>
<thead>
<tr>
<th>Signatures</th>
<th><code>file:resolve-path($path as xs:string) as xs:string</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Transforms the $path argument to an absolute operating system path.</td>
</tr>
</tbody>
</table>

**file:path-to-uri**

<table>
<thead>
<tr>
<th>Signatures</th>
<th><code>file:path-to-uri($path as xs:string) as xs:string</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Transforms the path specified by $path into a URI with the file:// scheme.</td>
</tr>
</tbody>
</table>

### System Properties

**file:dir-separator**

<table>
<thead>
<tr>
<th>Signatures</th>
<th><code>file:dir-separator() as xs:string</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the directory separator used by the operating system, such as / or .</td>
</tr>
</tbody>
</table>

**file:path-separator**

<table>
<thead>
<tr>
<th>Signatures</th>
<th><code>file:path-separator() as xs:string</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the path separator used by the operating system, such as ; or :.</td>
</tr>
</tbody>
</table>

**file:line-separator**

<table>
<thead>
<tr>
<th>Signatures</th>
<th><code>file:line-separator() as xs:string</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the line separator used by the operating system, such as \n, \r, \n, or \t.</td>
</tr>
</tbody>
</table>

**file:temp-dir**

*Introduced with Version 7.7:*

<table>
<thead>
<tr>
<th>Signatures</th>
<th><code>file:temp-dir() as xs:string</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the system’s default temporary-file directory.</td>
</tr>
</tbody>
</table>
Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FILE0001</td>
<td>A specified path does not exist.</td>
</tr>
<tr>
<td>FILE0002</td>
<td>A file with the same path already exists.</td>
</tr>
<tr>
<td>FILE0003</td>
<td>The specified path does not point to a directory.</td>
</tr>
<tr>
<td>FILE0004</td>
<td>The specified path is a directory.</td>
</tr>
<tr>
<td>FILE0005</td>
<td>The specified encoding is not supported, or unknown.</td>
</tr>
<tr>
<td>FILE9999</td>
<td>The operation fails for some other reason specific to the operating system.</td>
</tr>
</tbody>
</table>

Changelog

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 38. Full-Text Module

This XQuery Module extends the W3C Full Text Recommendation with some useful functions: The index can be directly accessed, full-text 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

Signatures

| ft:search($db as xs:string, $terms as item()*) as text()* |
| ft:search($db as xs:string, $terms as item()*, $options as item()) as text()* |

Summary

Returns all text nodes from the full-text index of the database $db that contain the specified $terms. The options used for 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 overwrite the default full-text options, which can be either specified

- as children of an <options/> element, e.g.:

  `<options>
  <key1 value='value1'/>
  ...
  </options>`

- as map, which contains all key/value pairs:

  `map { "key1" := "value1", ... }`

The following keys are supported:

- mode: determines the search mode (also called AnyAllOption). 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 an empty string or true, or false. By default, fuzzy querying is turned off.

- wildcards: turns wildcard querying on or off. Allowed values are an empty string or true, or false. By default, wildcard querying is turned off.

Errors

- BXDB0002: The addressed database does not exist or could not be opened.
- BXFT0001: the full-text index is not available.
- BXFT0003: both fuzzy and wildcard querying was selected.

Examples

- `ft:search("DB", "QUERY")` returns all text nodes of the database DB that contain the term QUERY.

- `ft:search("DB", ("2010","2011"), map { 'mode':'all' })` returns all text nodes of the database DB that contain the numbers 2010 and 2011.
• The last example iterates over five databases and returns all elements containing terms similar to Hello World in the text nodes:

```xml
let $terms := "Hello Worlds"
let $fuzzy := true()
let $options :=
  <options>
    <fuzzy>{ $fuzzy }</fuzzy>
  </options>
for $db in 1 to 3
let $dbname := 'DB' || $db
return ft:search($dbname, $terms, $options)/..
```

**ft:mark**

**Signatures**

```xml
ft:mark($nodes as node()) as node()*
```

**Summary**

Puts a marker element around the resulting $nodes of a full-text index request. The default tag name of the marker element is mark. An alternative tag name can be chosen via the optional $tag argument. Please note that:

- 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 2.
- the full-text expression, which computes the token positions, must be specified within 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 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":

```xml
ft:mark(db:open('DB')//*[text() contains text 'hello'])
```

**Example 2**: The following expression returns `<p><b>word</b></p>:

```xml
copy $p := <p>word</p>
modify ()
return ft:mark($p[text() contains text 'word'], 'b')
```

**Example 3**: The following expression loops through the first ten full-text results and marks the results in a second expression:

```xml
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 }])
}
```

**ft:extract**

**Signatures**

```xml
ft:extract($nodes as node()) as node()*
```

```xml
ft:extract($nodes as node()*, $tag as x:s:string) as node()*
```

```xml
ft:extract($nodes as node()*, $tag as x:s:string, $length as x:s:integer) as node()*
```
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 tag name of the marker element is mark. An alternative tag name can be chosen via the optional $tag 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":

  ft:extract(db:open('DB')//*[text() contains text 'hello'], 'b', 1)

ft:count

Signatures

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

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

ft:tokens($db as xs:string) as element(value)* ft:tokens($db as xs:string, $prefix 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.

ft:tokenize

Signatures

ft:tokenize($input as xs:string) as xs:string*

Summary

Tokenizes the given $input string, using the current default full-text options.

Examples

• ft:tokenize("No Doubt") returns the two strings no and doubt.
  • declare ft-option using stemming; ft:tokenize("GIFTS") returns a single string gift.

Errors

Code   Description

BXFT0001 Both wildcards and fuzzy search have been specified as search options.

Changelog

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 39. 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

**Signatures**

```
geo:dimension($geometry as element(*)) as xs:integer
```

**Summary**

Returns the dimension 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';
let $point := <gml:Point><gml:coordinates>1,2</gml:coordinates></gml:Point>
return geo:dimension($point)
```

geo:geometry-type

**Signatures**

```
geo:geometry-type($geometry as element(*)) as xs:QName
```

**Summary**

Returns the name of the geometry type of given geometry $geometry, if the geometry is not recognized with an error massage.

**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
Query:

import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
```
let $point := <gml:Point><gml:coordinates>1,2</gml:coordinates></gml:Point>
return geo:geometry-type($point)

Result:
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.
GEO0005: the output object cannot be written as an element by writer for some reason.

Example  Query:

import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';

let $line := <gml:LinearRing><gml:coordinates>1,1 20,1 20,20 1,20 1,1</gml:coordinates></gml:LinearRing>
return geo:envelope($line)

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  Query:

import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';

let $point := <gml:Point><gml:coordinates>1,2</gml:coordinates></gml:Point>
### geo:as-text

**Signatures**
`geo:as-text($point)`

**Summary**
Returns the GeoJSON text representation of the given point `$point`.

**Example**
```xml
let $point := <gml:Point><gml:coordinates>1,2</gml:coordinates></gml:Point>
return geo:as-text($point)
```

**Result:**
`POINT (1 2)`

### 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';
let $point := <gml:Point><gml:coordinates>1,2</gml:coordinates></gml:Point>
return geo:as-text($point)
```

**Result:**
`AAAAAAE/8AAAAAAAEEEEEEEEEEA`

### geo:is-simple

**Signatures**
`geo:is-simple($geometry as element(*)) as xs:boolean`

**Summary**
Returns whether the given geometry is simple `$geometry` and does not have any 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)).

**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';
let $line := <gml:MultiLineString>
  <gml:LineString><gml:coordinates>1,1 0,0 2,1</gml:coordinates></gml:LineString>
  <gml:LineString><gml:coordinates>2,1 3,3 4,4</gml:coordinates></gml:LineString>
</gml:MultiLineString>
return geo:is-simple($line)
```

**Result:**
`true`

### geo:boundary

**Signatures**
`geo:boundary($geometry as element(*)) as element(*)?`

**Summary**
Returns the boundary of the given geometry `$geometry`. If `$geometry` is a simple line or point, the result is an empty element. Otherwise, the result is an element containing the boundary.

**Example**
```xml
import module namespace geo = "http://expath.org/ns/geo";
let $line := <gml:MultiLineString>
  <gml:LineString><gml:coordinates>1,1 0,0 2,1</gml:coordinates></gml:LineString>
  <gml:LineString><gml:coordinates>2,1 3,3 4,4</gml:coordinates></gml:LineString>
</gml:MultiLineString>
return geo:boundary($line)
```
Summary
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.

Errors
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.

Example
Query:

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
let $Line := <gml:LineString><gml:coordinates>1,1 0,0 2,1</gml:coordinates></gml:LineString>
return geo:boundary($Line)
```

Result:

```
<gml:MultiPoint>
  <gml:pointMember>
    <gml:Point>
      <gml:coordinates>1.0,1.0</gml:coordinates>
    </gml:Point>
  </gml:pointMember>
  <gml:pointMember>
    <gml:Point>
      <gml:coordinates>2.0,1.0</gml:coordinates>
    </gml:Point>
  </gml:pointMember>
</gml:MultiPoint>
```

geo:num-geometries

Signatures
geo:num-geometries($geometry as element(*)) as xs:integer

Summary
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, ...).

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
Query:

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
let $Line := <gml:MultiLineString>
  <gml:LineString><gml:coordinates>1,1 0,0 2,1</gml:coordinates></gml:LineString>
  <gml:LineString><gml:coordinates>2,1 3,3 4,4</gml:coordinates></gml:LineString>
</gml:MultiLineString>
return geo:num-geometries($Line)
```

Result:
Query:

import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';

let $Line := <gml:LineString><gml:coordinates>1,1 0,0 2,1</gml:coordinates></gml:LineString>

return geo:num-geometries($Line)

Result:

1

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 $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 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

Query:

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)

Result:

<gml:LineString>
    <gml:coordinates>1.0,1.0 0.0,0.0 2.0,1.0</gml:coordinates>
</gml:LineString>

Query:

import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';

let $Line := <gml:LineString><gml:coordinates>1,1 0,0 2,1</gml:coordinates></gml:LineString>

return geo:geometry-n($Line, 1)
Result:

```xml
<gml:LineString>
  <gml:coordinates>1.0,1.0 0.0,0.0 2.0,1.0</gml:coordinates>
</gml:LineString>
```

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**

Query:

```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)
```

**Result:**

4

Query:

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
return geo:length($Line)
```

**Result:**

3.6502815398728847

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 return 1. This is an implementation different from the EXPath geo specification, as it limits the input geometry type only to lines.

**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**

Query:
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';

return geo:num-points($Line)

Result:
3

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 | Query:
---|---
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)

Result:
1

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 | Query:
---|---
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>
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 the surface. Otherwise, the point may lie on the boundary of the geometry.

**Errors**
GEO00001: the given element(s) is not recognized as a valid geometry (QName).
GEO00002: the given element cannot be read by reader for some reason.
GEO00005: the output object cannot be written as an element by writer for some reason.

**Example**

**Query:**

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
let $Line ::= <gml:LineString><gml:coordinates>1,1 55,99 2,1</gml:coordinates></gml:LineString>
return geo:point-on-surface($Line)
```

**Result:**

```xml
<gml:Point>
  <gml:coordinates>55.0,99.0</gml:coordinates>
</gml:Point>
```
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:point-on-surface($Polygon)

Result:

<gml:Point>
    <gml:coordinates>1.5,1.5</gml:coordinates>
</gml:Point>

Spatial Predicate Functions

geo:equals

Signatures
geo:equals($geometry1 as element(*), $geometry2 as element(*)) as xs:boolean

Summary
Returns whether geometry1 $geometry1 is spatially equal to $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
Query:

import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';

let $Line1:= <gml:LineString><gml:coordinates>1,1 55,99 2,1</gml:coordinates></gml:LineString>

let $Line2:= <gml:LineString><gml:coordinates>1,1 1,1 55,99 2,1</gml:coordinates></gml:LineString>

return geo:equals($Line1, $Line2)

Result:

true

geo:disjoint

Signatures
geo:disjoint($geometry1 as element(*), $geometry2 as element(*)) as xs:boolean

Summary
Returns whether geometry1 $geometry1 is spatially disjoint from $geometry2 $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
Query:

import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';

let $Line1:= := <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>

let $Line2:= <gml:LineString><gml:coordinates>0,0 2,1 3,3</gml:coordinates></gml:LineString>
return geo:disjoint($Line1, $Line2)

Result:
false

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 Query:
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';

let $Line1:= := <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 $Line2:= <gml:LineString><gml:coordinates>0,0 2,1 3,3</gml:coordinates></gml:LineString>
return geo:intersects($Line1, $Line2)

Result:
true

geo:touches

Signatures geo:touches($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 Query:
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:touches($Line, $Polygon)

Result:

true

geo:crosses

Signatures geo: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.

Example Query:

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)

Result:

false

geo:within

Signatures geo:within($geometry1 as element(*), $geometry2 as element(*)) as xs:boolean

Summary Returns whether geometry1 $geometry1 is spatially within $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 Query:
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:within($Line, $Polygon)

Result:
false

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 Query:

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)

Result:
false

geo:overlaps

Signatures geo:overlaps($geometry1 as element(*), $geometry2 as element(*)) as xs:boolean

Summary Returns whether $geometry1 is spatially 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 Query:
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: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)

Result:
false

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
$geometry1 and $geometry2 $geometry2 match the pattern specified in $intersectionMatrix
$geometry2, 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.

Example Query:

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>
        <gml:LinearRing><gml:coordinates>10,10 20,10 20,40
        10,40 10,10</gml:coordinates></gml:LinearRing>
    </gml:outerBoundaryIs>
</gml:Polygon>
Geo Module

```
return geo:relate($Point, $Polygon)

Result:
true
```

Analysis Functions

**geo:distance**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>geo:distance($geometry1 as element(<em>), $geometry2 as element(</em>)) as xs:double</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the shortest distance, in the units of the spatial reference system of $geometry1, between the geometries, where that distance is the distance between a point on each of the geometries.</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.</td>
</tr>
<tr>
<td>Example</td>
<td>Query:</td>
</tr>
</tbody>
</table>

```
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)

Result:
60
```

**geo:buffer**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>geo:buffer($geometry as element(<em>), $distance as xs:double) as element(</em>)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>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.</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. GEO0005: the output object cannot be written as an element by writer for some reason.</td>
</tr>
<tr>
<td>Example</td>
<td>Query:</td>
</tr>
</tbody>
</table>

```
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:buffer($Line, $Polygon)

Result:
```
Geo Module

---

import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';

let $Polygon:= <gml:Polygon>
    <gml:outerBoundaryIs>
        <gml:LinearRing><gml:coordinates>10,10 20,10 30,40 20,40 10,10</gml:coordinates></gml:LinearRing>
    </gml:outerBoundaryIs>
</gml:Polygon>
return geo:buffer($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**

**Query:**

import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';

let $Line:= <gml:LinearRing>
    <gml:coordinates>10,400 20,200 30,100 20,100 10,400</gml:coordinates>
</gml:LinearRing>
return geo:convex-hull($Line)

**Result:**

<gml:Polygon>
    <gml:outerBoundaryIs>
        <gml:LinearRing>
            <gml:coordinates>20.0,100.0 10.0,400.0 30.0,100.0 20.0,100.0</gml:coordinates>
        </gml:LinearRing>
    </gml:outerBoundaryIs>
</gml:Polygon>

---

**geo:intersection**

**Signatures**

geo:intersection($geometry1 as element(*), $geometry2 as element(*)) as element(*)?

**Summary**

Returns the intersection geometry of geometry1 $geometry1 with geometry2 $geometry2, in GML or empty sequence if there is no intersection of these 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. GEO0005: the output object cannot be written as an element by writer for some reason.

**Example**

**Query:**

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 $Point := <gml:Point><gml:coordinates>1.00,1.00</gml:coordinates></gml:Point>

return geo:intersection($Line, $Point)

Result:

<gml:Point>
  <gml:coordinates>1.0,1.0</gml:coordinates>
</gml:Point>

---

**geo:union**

**Signatures**
geo:union($geometry1 as element(*), $geometry2 as element(*)) as element(*)

**Summary**
Returns the union geometry of geometry1 $geometry1 with geometry2 $geometry2, in GML.

**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**

Query:

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 $Point := <gml:Point><gml:coordinates>1.00,1.00</gml:coordinates></gml:Point>

return geo:union($Line, $Point)

Result:

<gml:LineString>
  <gml:coordinates>1.0,1.0 55.0,99.0 2.0,1.0</gml:coordinates>
</gml:LineString>

---

**geo:difference**

**Signatures**
geo:difference($geometry1 as element(*), $geometry2 as element(*)) as element(*)?

**Summary**
Returns the difference geometry of geometry1 $geometry1 with geometry2 $geometry2, in GML, or empty sequence if the difference is empty, as a set of point in geometry1 $geometry1 and not included in 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.
GEO0005: the output object cannot be written as an element by writer for some reason.

**Example**

Query:
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>

return geo:difference($Point, $Line)

Result:

<gml:Point>
  <gml:coordinates>1.0,1.0</gml:coordinates>
</gml:Point>

geo:sym-difference

Signatures
geo:sym-difference($geometry1 as element(*), $geometry2 as element(*)) as element(*)?

Summary
Returns the symmetric difference geometry of geometry1 $geometry1 with geometry2 $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

Query:

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>

return geo:sym-difference($Point, $Line)

Result:

<gml:MultiGeometry>
  <gml:geometryMember>
    <gml:Point>
      <gml:coordinates>1.0,1.0</gml:coordinates>
    </gml:Point>
  </gml:geometryMember>
  <gml:geometryMember>
    <gml:LineString>
      <gml:coordinates>2.0,1.0 3.0,3.0 4.0,4.0</gml:coordinates>
    </gml:LineString>
  </gml:geometryMember>
</gml:MultiGeometry>

Functions Specific to Geometry Type

geo:x

Signatures
geo:x($point as element(*)) as xs:double
### Geo Module

**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**

**Query:**

```
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';

let $Point := <gml:Point><gml:coordinates>1.00,1.00</gml:coordinates></gml:Point>
return geo:x($Point)
```

**Result:**

1

---

#### geo:y

**Signatures**
geo:y($point as element(*)) as xs:double?

**Summary**
Returns the y 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**

**Query:**

```
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';

let $Point := <gml:Point><gml:coordinates>1.00,2.00</gml:coordinates></gml:Point>
return geo:y($Point)
```

**Result:**

2

---

#### 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**

**Query:**

```
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';

let $Point := <gml:Point><gml:coordinates>1.00,1.00,3.00</gml:coordinates></gml:Point>
return geo:z($Point)
```
**Geo Module**

---

### Result:

| 3 |

---

**geo:start-point**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>geo:start-point($line as element(<em>)) as element(</em>)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the starting point of the given line $line. $line has to be a single line, LineString or LinearRing.</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 line. Other geometries are not accepted.</td>
</tr>
</tbody>
</table>

**Example**

#### Query:

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
return geo:start-point($Line)
```

**Result:**

```xml
<gml:Point>
  <gml:coordinates>2.0,1.0</gml:coordinates>
</gml:Point>
```

---

**geo:end-point**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>geo:end-point($line as element(<em>)) as element(</em>)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the ending point of the given line $line. $line has to be a single line, LineString or LinearRing.</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 line. Other geometries are not accepted.</td>
</tr>
</tbody>
</table>

**Example**

#### Query:

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
return geo:end-point($Line)
```

**Result:**

```xml
<gml:Point>
  <gml:coordinates>4.0,4.0</gml:coordinates>
</gml:Point>
```

---

**geo:is-closed**

| Signatures | geo:is-closed($line as element(*)) as xs:boolean |

---

187
### geo:is-closed

**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**

**Query:**

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';

return geo:is-closed($Line)
```

**Result:**
```
false
```

### 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**

**Query:**

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';

return geo:is-ring($Line)
```

**Result:**
```
false
```

### 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**

**Query:**

```xml
import module namespace geo = "http://expath.org/ns/geo";
```
| declare namespace gml='http://www.opengis.net/gml'; |
| return geo:point-n($Line,1) |
| **Result:** |
| <gml:Point> |
|  | <gml:coordinates>2.0,1.0</gml:coordinates> |
|  | </gml:Point> |

### geo:exterior-ring

**Signatures**

geo:exterior-ring($polygon as element(*)) as element(*)

**Summary**

Returns the outer ring of the given polygon $geometry, as a gml:LineString.

**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 polygon. Other geometries are not accepted.GEO0005: the output object cannot be written as an element by writer for some reason.

**Example**

**Query:**

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
let $Polygon:= <gml:Polygon>
  <gml:outerBoundaryIs>
    <gml:LinearRing><gml:coordinates>10,10 20,10 30,40 20,40 10,10</gml:coordinates></gml:LinearRing>
  </gml:outerBoundaryIs>
</gml:Polygon>
return geo:exterior-ring($Polygon)
```

**Result:**

```xml
<gml:LineString>
  <gml:coordinates>10.0,10.0 20.0,10.0 30.0,40.0 20.0,40.0 10.0,10.0</gml:coordinates>
</gml:LineString>
```

### geo:num-interior-ring

**Signatures**

geo:num-interior-ring($polygon as element(*)) as xs:integer

**Summary**

Returns the number of interior rings in the given polygon $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.GEO0003: the given element has to be a polygon. Other geometries are not accepted.

**Example**

**Query:**

```xml
import module namespace geo = "http://expath.org/ns/geo";
declare namespace gml='http://www.opengis.net/gml';
let $Polygon:= <gml:Polygon>
  
```

189
<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>

return geo:num-interior-ring($Polygon)

Result:

2

geo:interior-ring-n

Signatures
geo:interior-ring-n($polygon as element(*)) as element(*)

Summary
Returns the outer ring of the given polygon $geometry, as a gml:LineString.

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 polygon. Other geometries are not accepted.
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

Query:

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 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>

return geo:interior-ring-n($Polygon, 1)

Result:

<gml:LineString>
  <gml:coordinates>2.0,2.0 3.0,2.0 3.0,3.0 2.0,3.0 2.0,2.0</gml:coordinates>
</gml:LineString>

Errors

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
</table>

190
Geo Module

GEO0001  Unrecognized Geo type.
GEO0002  The input GML node cannot be read by GMLreader.
GEO0003  Input geometry is not an appropriate geometry for this function.
GEO0004  The input index is out of range.
GEO0005  The result geometry can not be written by GMLwriter.

Changelog

The module was introduced with Version 7.6.
Chapter 40. 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 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, which can be specified...

- as children of an <html:options/> element; e.g.:

```xml
<html:options>
  <html:key1 value='value1'/>
  ...
</html:options>
```

- as map, which contains all key/value pairs:

```xml
map { "key1" := "value1", ... }
```

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"/>
    ...
</html>

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 41. 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()+

**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.

**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.

**Notes**

The attribute auth-method of $request is not considered in our implementation because we are handling only basic authentication.

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: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:

```
<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
    <meta http-equiv="content-type" content="text/html; charset=ISO-8859-1"/>
    <title>Google</title>
    <script>window.google={kEI:"rZB-...
    </script>
  </head>
  <body>
    ...
  </body>
</html>
```

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). The result can then be further processed:

Query:

```
let $binary := http:send-request{
  <http:request method='get'
    override-media-type='application/octet-stream'
    href='http://www.google.com'/>
}[2]
return try {
  html:parse($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="Age" value="9260"/>
  <http:header name="Date" value="Mon, 14 Mar 2011 19:17:10 GMT"/>
  <http:header name="Content-Type" value="image/svg+xml"/>
  <http:header name="X-Cache-Lookup" value="MISS from knsq22.knams.wikimedia.org:80"/>
  <http:header name="Server" value="Sun-Java-System-Web-Server/7.0"/>
  <http:header name="X-Cache" value="MISS from knsq22.knams.wikimedia.org"/>
  <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">
        <![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: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 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 42. 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 or xs:base64Binary.

Errors

FORG0006: the specified value is neither a string nor a binary item.

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 or xs:base64Binary.

Errors

FORG0006: the specified value is neither a string nor a binary item.

Examples

• xs:hexBinary(hash:sha1("BaseX")) returns 3AD5958F0F27D5AFFDCA2957560F121D0597A4ED.

• hash:sha1(xs:base64Binary("")) returns 2jmj715rSw0yVb/v1WAYkY/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 or xs:base64Binary.

Errors

FORG0006: the specified value is neither a string nor a binary item.

Examples

• xs:hexBinary(hash:sha256("BaseX")) returns 15D570763DEB75D728BB69643392873B835CCC94A2F1E881909DA47662821A3.

• hash:sha256(xs:base64Binary("")) returns 47DEQpj8HBSa+/TImW+5JCeuQeRkmSNMpJWZG3hSuFU=.

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 or xs:base64Binary. The following three algorithms are supported: MD5, SHA-1, and SHA-256.

Errors
HASH0001: the specified hashing algorithm is unknown. FORG0006: the specified value is neither a string nor a binary item.

Examples
• `xs:hexBinary(hash:md5("", "MD5"))` returns D41D8CD98F00B204E9800998ECF8427E.
• `hash:md5("", "")` 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 43. 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 $sort-by := function($f, $seq) {
    for $x in $seq
        order by $f($x)
        return $x
}
let $sort := $sort-by(hof:id#1, ?),
    $reverse-sort := $sort-by(function($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())
```

**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, $seq2) {
    sum(map-pairs($f, $seq1, $seq2))
}
let $sum-all := $zip-sum(function($a, $b) { $a + $b }, ?, ?),
```


Higher-Order Functions Module

$sum-left := \text{\$zip-sum(hof:const\#2, ?, ?)}$
return (  
$\text{\$sum-all((1, 1, 1, 1, 1), 1 \ to \ 5),}$  
$\text{\$sum-left((1, 1, 1, 1, 1), 1 \ to \ 5))}$
)

• Another use-case: When inserting a key into a map, $f$ decides how to combine the new value with a possibly existing old one. $\text{\hof:const}$ here means ignoring the old value, so that's normal insertion.

let $\text{\$insert-with := function($f, $map, $k, $v)}$
{  
let $\text{\$old := $map($k),}$  
$\text{\$new := if($\text{\$old}) then $f($v, $\text{\$old}) else $v}$  
return $\text{\map:new(((\text{\$map, map( $k := $new }))}$
}
let $\text{\$map := map( 'foo' := 1 )}$
let $\text{\$add := $\text{\$insert-with(function($a, $b) {$a + $b}), ?, ?, ?),}$  
$\text{\$insert := $\text{\$insert-with(hof:const\#2, ?, ?, ?))}$
return (  
$\text{\$add($map, 'foo', 2)('foo'),}$  
$\text{\$insert($map, 'foo', 42)('foo')}$
)
returns 3 42

\text{\hof:fold-left1}

Updated with Version 7.7: the arguments have been swapped, similar to the standard Higher-Order Functions.

Signatures \text{\hof:fold-left1($seq \ as \ item()\+, \ $f \ as \ function(item()\*, \ item()) \ as \ item()\*\+) as \ item()\*\)}

Summary Works the same as \text{\fn:fold-left}, but doesn’t need a seed, because the sequence must be non-empty.

Examples • \text{\hof:fold-left1(1 \ to \ 10, \ function($a, $b) \{ \$a + \$b \}) \ returns \ 55.}$

• \text{\hof:fold-left1((), \ function($a, $b) \{ \$a + \$b \}) \ throws \ XPTY0004, because $seq \ has \ to \ be \ non-empty.}$

\text{\hof:until}

Signatures \text{\hof:until($pred \ as \ function(item()\*) \ as \ xs:boolean, \ $f \ as \ function(item()\*) \ as \ item()\*, \ $start \ as \ item()\*) \ as \ item()\*\)}

Summary Applies the function $f$ to the initial value $start$ until the predicate $pred$ applied to the result returns true().

Examples • \text{\hof:until(function($x) \{ \$x \ ge \ 1000 \}, \ function($y) \{ \ 2 * \$y \}, 1) \ returns \ 1024.}$

• Calculating the square-root of a number by iteratively improving an initial guess:

let $\text{\$sqrt := function($x \ as \ xs:double) \ as \ xs:double (}$
\text{\hof:until{}}
function($\text{\$res) \{ abs($\text{\$res * $\text{\$res - $x) < 0.00001 \),}}$
function($\text{\$guess) \{ ($\text{\$guess + $x \ div \$\text{\$guess) \ div \ 2 \),}$
\text{\$x}$
}\})
return $\text{\$sqrt(25)}$
returns 5.000000000053722.
**hof:top-k-by**

**Signatures**

```
$hof:top-k-by($seq as item()*, $sort-key as function(item()) as item(), $k as xs:integer) as item()*
```

**Summary**

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:

```
(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**

```
$hof:top-k-with($seq as item()*, $lt as function(item(), item()) as xs:boolean, $k as xs:integer) as item()*
```

**Summary**

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 7.2

- Added: `hof:top-k-by, hof:top-k-with`
- Removed: `hof:iterate`

Version 7.0

- module added
Chapter 44. 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

Signature

index:facets($db as xs:string) as xs:string
index:facets($db as xs:string, $type as xs:string) as xs:string

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

Signature

index:texts($db as xs:string) as element(value)*
index:texts($db as xs:string, $prefix as xs:string) as element(value)*
index:texts($db as xs:string, $start as xs:string, $ascending as xs:boolean) as element(value)*

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 text index is not available.

index:attributes

Signature

index:attributes($db as xs:string) as element(value)*
index:attributes($db as xs:string, $prefix as xs:string) as element(value)*
index:attributes($db as xs:string, $start as xs:string, $ascending as xs:boolean) as element(value)*

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 attribute index is not available.
**index:element-names**

<table>
<thead>
<tr>
<th>Signatures</th>
<th><code>index:element-names($db as xs:string) as element(value)*</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns all element names stored in the Name Index of the database $db, along with their number of occurrences.</td>
</tr>
<tr>
<td>Errors</td>
<td>BXDB0002: The addressed database does not exist or could not be opened.</td>
</tr>
</tbody>
</table>

**index:attribute-names**

<table>
<thead>
<tr>
<th>Signatures</th>
<th><code>index:attribute-names($db as xs:string) as element(value)*</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns all attribute names stored in the Name Index of the database $db, along with their number of occurrences.</td>
</tr>
<tr>
<td>Errors</td>
<td>BXDB0002: The addressed database does not exist or could not be opened.</td>
</tr>
</tbody>
</table>

**Changelog**

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 45. 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

<table>
<thead>
<tr>
<th>Signatures</th>
<th>inspect:functions() as function(<em>)</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns function items for all user-defined functions that are known in the current query context.</td>
</tr>
<tr>
<td>Examples</td>
<td>The following query returns 42:</td>
</tr>
<tr>
<td></td>
<td>$f = declare %private function local:solution() { 42 }; for $f in inspect:functions() return $f()</td>
</tr>
</tbody>
</table>

Documentation

inspect:function

<table>
<thead>
<tr>
<th>Signatures</th>
<th>inspect:function($function as function(*)) as element(function)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>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.</td>
</tr>
<tr>
<td>Examples</td>
<td>The query inspect:function(count#1) yields:</td>
</tr>
<tr>
<td></td>
<td>&lt;function name=&quot;count&quot; uri=&quot;http://www.w3.org/2005/xpath-functions&quot;&gt;</td>
</tr>
<tr>
<td></td>
<td>&lt;argument type=&quot;item()&quot; occurrence=&quot;*&quot;/&gt;</td>
</tr>
<tr>
<td></td>
<td>&lt;return type=&quot;xs:integer&quot;/&gt;</td>
</tr>
<tr>
<td></td>
<td>&lt;/function&gt;</td>
</tr>
<tr>
<td></td>
<td>The function...</td>
</tr>
<tr>
<td></td>
<td>(~</td>
</tr>
<tr>
<td></td>
<td>: This function simply returns the specified integer.</td>
</tr>
<tr>
<td></td>
<td>: @param $number number to return</td>
</tr>
<tr>
<td></td>
<td>: @return specified number</td>
</tr>
<tr>
<td></td>
<td>)</td>
</tr>
<tr>
<td></td>
<td>declare %private function local:same($number as xs:integer) as</td>
</tr>
<tr>
<td></td>
<td>xs:integer {</td>
</tr>
<tr>
<td></td>
<td>$number</td>
</tr>
<tr>
<td></td>
<td>}</td>
</tr>
<tr>
<td></td>
<td>...is represented by inspect:function(local:same#1) as...</td>
</tr>
<tr>
<td></td>
<td>&lt;function name=&quot;local:same&quot; uri=&quot;http://www.w3.org/2005/xquery-local-functions&quot;&gt;</td>
</tr>
</tbody>
</table>
**inspect:context**

**Signatures**  
inspect:context() as element(context)

**Summary**  
Generates an element that describes all variables and functions in the current query context.

**Examples**  
Evaluate all user-defined functions with zero arguments in the query context:

```xml
inspect:context()/function ! function-lookup(QName(@uri, @name), 0) ! . ()
```

Return the names of all private functions in the current context:

```xml
for $f in inspect:context()/function
where $f/annotation/@name = 'private'
return $f/@name/string()
```

**inspect:module**

**Signatures**  
inspect:module($uri as xs:string) as element(module)

**Summary**  
Retrieves the string from the specified $input, parses it as XQuery module, and generates an element that describes its structure.

**Errors**  
FODC0002: the addressed resource cannot be retrieved.

**Examples**  
An example is shown below.

**inspect:xqdoc**

**Signatures**  
inspect:xqdoc($uri as xs:string) as element(xqdoc:xqdoc)

**Summary**  
Retrieves the string from the specified $input, parses it as XQuery module, and generates an xqDoc element. xqDoc provides a simple vendor neutral solution for generating a 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):

- 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:
Inspection Module

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';

declare variable $samples:test-string as xs:string := 'this is a string';

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:

```
<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:

```
<xqdoc:xqdoc xmlns:xqdoc="http://www.xqdoc.org/1.0">
  <xqdoc:control>
    <xqdoc:date>2013-06-01T16:59:33.654+02:00</xqdoc:date>
    <xqdoc:version>1.1</xqdoc:version>
  </xqdoc:control>
  <xqdoc:module type="library">
    <xqdoc:uri>http://basex.org/modules/samples</xqdoc:uri>
    <xqdoc:name>sample.xqm</xqdoc:name>
    <xqdoc:comment>
      <xqdoc:description>This module provides some sample functions to demonstrate the features of the Inspection Module.</xqdoc:description>
      <xqdoc:author>BaseX Team</xqdoc:author>
      <xqdoc:see>http://docs.basex.org/wiki/XQDoc_Module</xqdoc:see>
      <xqdoc:version>1.0</xqdoc:version>
    </xqdoc:comment>
  </xqdoc:module>
  <xqdoc:namespaces>
    <xqdoc:namespace prefix="samples" uri="http://basex.org/modules/samples"/>
  </xqdoc:namespaces>
</xqdoc:xqdoc>
```
Changelog

This module was introduced with Version 7.7.
Chapter 46. 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 two sets of functions in this module:

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.

Formats

BaseX

The standard BaseX JSON format facilitates a lossless conversion from JSON to XML and back. The transformation is based on the following rules:

1. The resulting document has a <json/> root node.
2. Names (keys) of objects are represented as elements:
   a. Empty names are represented by a single underscore (_<.../></_>).
   b. Underscore characters are rewritten to two underscores (___).
   c. A character that cannot be represented as NCName character is rewritten to an underscore and its four-digit Unicode.
3. As the members of arrays have no names, <value/> is used as element name.
4. JSON values are represented as text nodes.
5. The types of values are represented in attributes:
   a. The value types number, boolean, null, object and array are represented by a type attribute.
   b. The string type is omitted, as it is treated as default type.
   c. 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. The attributes are named by their type in the plural (numbers, booleans, nulls, objects and arrays), and the attribute value contains all names with that type, separated by whitespaces.

JsonML

The JsonML format can be used to transform 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.

Functions

json:parse

Version 7.7.2: $options argument added
### json:parse

**Signatures**

`json:parse($input as xs:string) as element(json)`

`json:parse($input as xs:string, $options as item()) as element(json)`

**Summary**

Converts the JSON document specified by `$input` to XML, and returns the result as `element(json)` instance. The converted XML document is both well readable and lossless, i.e., the converted document can be serialized back to the original JSON representation. The `$options` argument can be used to control the way the input is converted. The following options are available (they bear resemblance to the ones specified in XSLT 3.0):

- **spec** determines the used JSON specification. Allowed values are RFC4627, ECMA-262, and liberal; the default is (RFC4627).
- **unescape** determines whether escape sequences (marked by a backslash) in the input are expanded. The default is true.
- **format** determines the conversion format. Allowed values are `json`, `jsonml`, and `map`; the default is `json`.

Options can either be specified

- as children of an `<json:options/>` element, e.g.:

  ```xml
  <json:options>
     <json:format value='jsonml'/>
     ...
  </json:options>
  ```

- or as map, which contains all key/value pairs:

  ```xml
  { 'format' : 'jsonml', ... }
  ```

**Errors**

- BXJS0001: the specified input cannot be parsed as JSON document.
- BXJS0003: a specified option cannot be parsed.

### json:serialize

**Version 7.7.2**: `$options` argument added

**Signatures**

`json:serialize($input as node()) as xs:string`  
`json:serialize($input as node(), $options as item()) as xs:string`

**Summary**

Serializes the node specified by `$input` as JSON, and returns the result as `xs:string` instance. The serialized node must conform to the syntax specified by the `json:parse()` function. XML documents can also be serialized as JSON if the Serialization Option method is set to `json`. The `$options` argument can be used to control the way the input is serialized. The following options are available (they bear resemblance to the ones specified in XSLT 3.0):

- **spec** determines the used JSON specification. Allowed values are RFC4627, ECMA-262, and liberal; the default is (RFC4627).
- **unescape** determines whether escape sequences (marked by a backslash) in the input are expanded. The default is true.
- **format** determines the conversion format. Allowed values are `json` and `jsonml`; the default is `json`.

Options can either be specified

- as children of an `<json:options/>` element, e.g.:
**JSON Module**

```
<json:options>
  <json:format value='jsonml'/>
  ...
</json:options>
```

• or as map, which contains all key/value pairs:

```
{ 'format' : 'jsonml', ... }
```

**Errors**

BXJS0002: the specified node cannot be serialized as JSON document.

**json:serialize-ml**

*Version 7.7.2: deprecated; please use json:serialize instead.*

**Signatures**

`json:serialize-ml($input as node()) as xs:string`

**Summary**

Serializes the node specified by `$input` and returns the result as `xs:string` instance. XML documents can also be output in the JsonML format by setting the **Serialization Option** "method" to "jsonml".

**Errors**

BXJS0002: the specified value cannot be serialized.

**json:parse-ml**

*Version 7.7.2: deprecated; please use json:parse instead.*

**Signatures**

`json:parse-ml($input as xs:string) as element()`

**Summary**

Converts the JsonML document specified by `$input` to XML, and returns the result as `element()` instance. The JSON input must conform to the JsonML specification to be successfully converted.

**Errors**

BXJS0001: the specified input cannot be parsed as JsonML instance.

**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, document { $json }, $name)
```

**Example 2: Converts a simple JSON string to XML and back**

**Query:**

```
json:parse('{}')
```

**Result:**

```
<json objects="json"/>
```

**Query:**
Example 3: Converts a JSON string with simple objects and arrays

Query:

```json
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": "2009-02-02T11:10:27Z",
    "generator": "http://www.flickr.com/
}')
```

Result:

```html
<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>2009-02-02T11:10:27Z</modified>
<generator>http://www.flickr.com/</generator>
```

Example 4: Converts a JSON string with different data types

Query:

```json
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
        }
    ]
}')
```

Result:

```html
<first__name>John</first__name>
```
<last__name>Smith</last__name>
<age>25</age>
<address>
  <street>21 2nd Street</street>
  <city>New York</city>
  <code>10021</code>
</address>
<phone>
  <value>
    <type>home</type>
    <number>212 555-1234</number>
  </value>
  <value>
    <type>mobile</type>
    <number type="number">1327724623</number>
  </value>
</phone>
</json>

JsonML Format

Example 1: Converts all XML documents in a database to JsonML and writes them to disk

Query:

```
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:

```
json:serialize(<xml/>, { 'format': 'jsonml' })
```

Result:

`["xml"]`

Example 3: Converts an XML document with elements and text

Query:

```
json:serialize(doc('flickr.xml'), { 'format': 'jsonml' })
```

flickr.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>2009-02-02T11:10:27Z</modified>
</flickr>
```

Result:

`["flickr",`
Example 4: Converts a document with nested elements and attributes

Query:

```javascript
json:serialize(doc('input.xml'), { 'format': 'jsonml' })
```

input.xml:

```xml
<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>
<tr>
<td>BXJS0003</td>
<td>A specified option cannot be parsed.</td>
</tr>
</tbody>
</table>

Changelog

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 47. Map Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions for manipulating maps. The following documentation is derived from an XQuery 3.0 Functions and Operators working draft proposal written by Michael H. Kay, and is not part of the official recommendation yet.

With Version 7.7.1, the map syntax has changed (the old syntax is still supported): the map keyword is omitted, and a colon is used as key/value delimiter (e.g.: { 'key' : 'value' }). The new syntax will most probably be standardized with XQuery 3.1.

Introduction

A map is an additional kind of item. It comprises a collation and a set of entries. Each entry comprises a key which is an arbitrary atomic value, and an arbitrary sequence called the associated value. Within a map, no two entries have the same key, when compared using the eq operator under the map's collation. It is not necessary that all the keys should be mutually comparable (for example, they can include a mixture of integers and strings). Key values will never be of type xs:untypedAtomic, and they will never be the xs:float or xs:double value NaN.

The function call map:get($map, $key) can be used to retrieve the value associated with a given key.

A map can also be viewed as a function from keys to associated values. To achieve this, a map is also a function item. The function corresponding to the map has the signature function($key as xs:anyAtomicType) as item() *. Calling the function has the same effect as calling the get function: the expression $map($key) returns the same result as map:get($map, $key). 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:

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

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

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.

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.

Maps may be compared using the fn:deep-equal function. The semantics for this function are extended so that when two items are compared, at any level of recursion, the items compare equal if they are both maps, if both use the same collation, if both contain the same set of keys (compared using the eq operator), without regard to ordering, and if for each key that is present in both maps, the associated values are deep-equal. When comparing maps, the maps' collation is used rather than the collation supplied as an argument to the fn:deep-equal function.

There is no operation to atomize a map or convert it to a string. The following XQuery snippet shows how the contents of a map can be serialized:

```
let $map := map { 1:='a', 2:='b' }
```
return string-join(
    for $m in map:keys($map)
    return concat($m, ':=', $map($m)), ', '
)

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"
};
```

**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**

**map:collation**

**Signatures**

map:collation($map as map(*)) as xs:string

**Summary**

Returns the collation URI of the map supplied as $map.

**map:contains**

**Signatures**

map:contains($map as map(*), $key as item()) as xs:boolean

**Summary**

Returns true if the map supplied as $map contains an entry with a key equal to the supplied value of $key; otherwise it returns false. The equality comparison uses the map’s collation; no error occurs 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. If the supplied key is the xs:float or xs:double value NaN, the function returns false.

**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 item(), $value as item()* as map(*)

**Summary**

Creates a new map containing a single entry. The collation of the new map is the default collation from the static context. The key of the entry in the new map is $key, and its associated value is $value. If the supplied key is xs:untypedAtomic, it is converted to xs:string. If the supplied key is the xs:float or xs:double value NaN, the supplied $map is returned unchanged. The function map:entry is intended primarily for use in conjunction with the function map:new. For example, a map containing seven entries may be constructed like this:
Unlike the map{ ... } expression, this technique can be used to construct a map with a variable number of entries, for example:

map:new(for $b in //book return map:entry($b/isbn, $b))

Examples

* map:entry("M", "Monday") creates a map with the values { "M":="Monday" }.

map:get

Signatures

map:get($map as map(*), $key as item()) 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 $map that has a key equal to the supplied value of $key. If there is such an entry, it returns the associated value; otherwise it returns an empty sequence. The equality comparison uses the map's collation; no error occurs 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. If the supplied key is the xs:float or xs:double value NaN, the function returns an empty sequence. 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 $map is a map, the expression $map($K) is equivalent to get($map, $K). Similarly, the expression get (get (get ($map, 'employee'), 'name'), 'first') can be written as $map('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($map 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 $map argument and returns the keys that are present in the map as a sequence of atomic values, in implementation-dependent order.

Examples

* map:keys(map{ 1:="yes", 2:="no" }) returns some permutation of (1, 2) (the result is in implementation-dependent order).

map:new

Signatures

map:new() as map(*) map:new($maps as map(*)) as map(*)
map:new($maps as map(*), $coll as xs:string) as map(*)

Summary

Constructs and returns a new map. The zero-argument form of the function returns an empty map whose collation is the default collation in the static context. It is equivalent to calling the one-argument form of the function with an empty sequence as the value of the first argument. The one-
argument form of the function returns a map that is formed by combining the contents of the maps supplied in the $maps argument. It is equivalent to calling the two-argument form of the function with the default collation from the static context as the second argument. The two-argument form of the function returns a map that is formed by combining the contents of the maps supplied in the $maps argument. The collation of the new map is the value of the $coll argument. The supplied 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 with $coll as the collation.
2. The associated value for each such key is taken from the last map in the input sequence $maps that contains an entry with this key. If this map contains more than one entry with this key (which can happen if its collation is different from that of the new map) then it is implementation-dependent which of them is selected.

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:new()` creates an empty map.
- `map:new(())` creates an empty map.
- `map:new(map:entry(0, "no"), map:entry(1, "yes"))` creates a map with the values `{ 0:="no", 1:="yes" }.

### map:remove

**Signatures**

map:remove($map as map(*), $key as item()) as map(*)

**Summary**

Constructs a new map by removing an entry from an existing map. The collation of the new map is the same as the collation of the map supplied as $map. The entries in the new map correspond to the entries of $map, 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:remove($week, 4)` creates a map with the values `{ 0:="Sonntag", 1:="Montag", 2:="Dienstag", 3:="Mittwoch", 5:="Freitag", 6:="Samstag" }.

### map:size

**Signatures**

map:size($map as map(*)) as xs:integer

**Summary**

Returns a the number of entries in the supplied map. The function takes any map as its $map argument and returns the number of entries that are present in the map.

**Examples**

- `map:size(map:new())` returns 0.
- `map:size(map{ "true":1, "false":0 })` returns 2.
Changelog

Version 7.7.1

• Updated: alternative map syntax with without map keyword and : as key/value delimiter (e.g.: { 'key' : 'value' })
Chapter 48. Math Module

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

math:pi

Signatures | math:pi() as xs:double
--- | ---
Summary | Returns the xs:double value of the mathematical constant π whose lexical representation is 3.141592653589793.
Examples | • 2*math:pi() returns 6.283185307179586e0.
 | • 60 * (math:pi() div 180) converts an angle of 60 degrees to radians.

math:sqrt

Signatures | math:sqrt($arg as xs:double?) as xs:double?
--- | ---
Summary | 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

Signatures | math:sin($arg as xs:double?) as xs:double?
--- | ---
Summary | 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

Signatures | math:cos($arg as xs:double?) as xs:double?
--- | ---
Summary | 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.

math:tan

Signatures | math:tan($arg as xs:double?) as xs:double?
--- | ---
Summary | Returns the tangent of $arg, expressed 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.
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:log(100) returns 2.
Additional Functions

math:e

Signatures  | math:e() as xs:double
Summary  | Returns the xs:double value of the mathematical constant e whose lexical representation is 2.718281828459045.
Examples  | • 5*math:e() returns 13.591409142295225.

math:sinh

Signatures  | math:sinh($arg as xs:double?) as xs:double?
Summary  | 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.
Examples  | • math:sinh(0) returns 0.

math:cosh

Signatures  | math:cosh($arg as xs:double?) as xs:double?
Summary  | 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.
Examples  | • math:cosh(0) returns 1.

math:tanh

Signatures  | math:tanh($arg as xs:double?) as xs:double?
Summary  | 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.
Examples  | • math:tanh(100) returns 1.

math:crc32

Signatures  | math:crc32($str as xs:string) as xs:hexBinary
Summary  | Calculates the CRC32 check sum of the given string $str.
Examples  | • math:crc32("") returns '00000000'.
| • math:crc32("BaseX") returns '4C06FC7F'.

Changelog

Version 7.5
• Moved: math:random and math:uuid have been move to Random Module.

Version 7.3
• Added: math:crc32 and math:uuid have been adopted from the obsolete Utility Module.
Chapter 49. 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>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>

Changelog

Introduced with Version 7.3. Functions have been adopted from the obsolete Utility Module.
Chapter 50. Process Module

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

Signatures

<table>
<thead>
<tr>
<th>Signature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>proc:system($cmd as xs:string) as xs:string</td>
<td>Executes the specified command in a separate process and returns the result as string. Additional command arguments may be specified via $args. The result can be explicitly converted to a specified $encoding. If no encoding is specified, the system’s default encoding is used.</td>
</tr>
<tr>
<td>proc:system($cmd as xs:string, $args as xs:string*) as xs:string</td>
<td></td>
</tr>
<tr>
<td>proc:system($cmd as xs:string, $args as xs:string*, $encoding as xs:string) as xs:string</td>
<td></td>
</tr>
</tbody>
</table>

Summary

Executes the specified command in a separate process and returns the result as string. Additional command arguments 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.

Errors

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.

Examples

- proc:system('date') returns the current date on a Linux system.
- The following example returns "Command not found", if the command "xyz" cannot be located or executed:
  
  ```xml
  try {
    proc:system('xyz')
  } catch bxerr:BXPR0002 {
    'Command not found.'
  }
  ```

proc:execute

Signatures

<table>
<thead>
<tr>
<th>Signature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>proc:execute($cmd as xs:string) as element(result)</td>
<td>Executes the specified command in a separate process and returns the result as element. Additional command arguments may be specified via $args. The result can be explicitly converted to a specified $encoding. If no encoding is specified, the system’s default encoding is used. A result has the following structure:</td>
</tr>
<tr>
<td>proc:execute($cmd as xs:string, $args as xs:string*) as element(result)</td>
<td></td>
</tr>
<tr>
<td>proc:execute($cmd as xs:string, $args as xs:string*, $encoding as xs:string) as element(result)</td>
<td></td>
</tr>
</tbody>
</table>

Summary

Executes the specified command in a separate process and returns the result as element. Additional command arguments 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:

```xml
<result>
  <output>...result...</output>
  <error>0</error>
</result>
```

Errors

BXPR9999: the specified encoding does not exist or is not supported.

Examples

- proc:execute('dir', '\') returns the root directory on a Windows system.
* proc:execute('ls', ('-l', '-a')) executes the famous ls -la command on Unix systems.

**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**

The module was introduced with Version 7.3.
Chapter 51. Profiling Module

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:mb("1 to 100000") may output 0 Bytes.  
• prof:mb("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

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>•</td>
<td>prof:human(16384) returns 16K.</td>
</tr>
</tbody>
</table>

### prof:dump

<table>
<thead>
<tr>
<th>Signatures</th>
<th>prof:dump($expr as item()) as empty-sequence() prof:dump($expr as item(), $label as xs:string) as empty-sequence()</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>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.</td>
</tr>
<tr>
<td>Properties</td>
<td>In contrast to fn:trace(), the consumed expression will not be passed on.</td>
</tr>
</tbody>
</table>

### prof:current-ms

<table>
<thead>
<tr>
<th>Signatures</th>
<th>prof:current-ms() as xs:integer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>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.</td>
</tr>
<tr>
<td>Properties</td>
<td>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.</td>
</tr>
</tbody>
</table>

### prof:current-ns

<table>
<thead>
<tr>
<th>Signatures</th>
<th>prof:current-ns() as xs:integer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns the current value of the most precise available system timer in nanoseconds.</td>
</tr>
<tr>
<td>Properties</td>
<td>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.</td>
</tr>
</tbody>
</table>

### prof:void

*Introduced with Version 7.7:*

<table>
<thead>
<tr>
<th>Signatures</th>
<th>prof:void($value as item()*) as empty-sequence()</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>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.</td>
</tr>
<tr>
<td>Properties</td>
<td>The function is non-deterministic: evaluation order will be preserved by the compiler.</td>
</tr>
<tr>
<td>Examples</td>
<td>• prof:void(fetch:binary('<a href="http://my.rest.service">http://my.rest.service</a>')) performs an HTTP request and ignores the result.</td>
</tr>
</tbody>
</table>

### Changelog

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 52. 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

**Signatures**  
`rest:base-uri()` as `xs:anyURI`

**Summary**  
This function returns the implementation defined base URI of the resource function.

rest:uri

**Signatures**  
`rest:uri()` as `xs:anyURI`

**Summary**  
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.

rest:wadl

**Signatures**  
`rest:wadl()` as `element(wadl:application)`

**Summary**  
This (unofficial) function returns a WADL description of all available REST services.

Changelog

This module was introduced with Version 7.7.
Chapter 53. Random Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains non-deterministic functions for returning random values.

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**

Signatures: `random:double() as xs:double`

Summary: Returns a double value between 0.0 (inclusive) and 1.0 (exclusive).

**random:integer**

Signatures: `random:integer() as xs:integer`  
`random:integer($max as xs:integer) as xs:integer`

Summary: Returns an integer value, either in the whole integer range or between 0 (inclusive) and the given maximum (exclusive).

**random:seeded-double**

Signatures: `random:seeded-double($seed as xs:integer, $num as xs:integer) as xs:double*`

Summary: 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.

**random:seeded-integer**

Signatures: `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*`

Summary: 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.

**random:gaussian**

Signatures: `random:gaussian($num as xs:integer) as xs:double*`

Summary: 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.

**random:uuid**

Signatures: `random:uuid() as xs:string`
Random Module

<table>
<thead>
<tr>
<th>Summary</th>
<th>Creates a random universally unique identifier (UUID), represented as 128-bit value.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Examples</td>
<td>• random:uuid() eq random:uuid() will (most probably) return the boolean value false.</td>
</tr>
</tbody>
</table>

**Changelog**

The module was introduced with Version 7.5. It includes some functionality which was previously located in the Math Module.
Chapter 54. Repository Module

Read this entry online in the BaseX Wiki.

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 invalid 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 version is not supported.
BXRE0007: 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

Code    Description

BXRE0001 The addressed package does not exist.
BXRE0002 A package uses an invalid namespace URI.
BXRE0003 The package to be installed requires a package which is not installed yet.
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.
BXRE0007 The package cannot be deleted.
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 55. 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 upcoming EXQuery Request Module.

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:

```xquery
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):

<table>
<thead>
<tr>
<th>scheme</th>
<th>hostname</th>
<th>port</th>
<th>path</th>
<th>query</th>
</tr>
</thead>
<tbody>
<tr>
<td>foo://example.com:8042/over/there?name=ferret</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

General Functions

request:method

**Signatures**

`request:method() as xs:string`

**Summary**

Returns the Method of the HTTP request.

request:attribute

**Introduced with Version 7.7:**

**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`
### Summary

<table>
<thead>
<tr>
<th>Request</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>request:hostname</code></td>
<td>Returns the Hostname component of the URI of an HTTP request.</td>
</tr>
<tr>
<td><code>request:port</code></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><code>request:path</code></td>
<td>Returns the Path component of the URI of an HTTP request.</td>
</tr>
<tr>
<td><code>request:query</code></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><code>request:uri</code></td>
<td>Returns the complete URI of an HTTP request as it has been specified by the client.</td>
</tr>
</tbody>
</table>

### Example

For the example given in the introduction, this function would return:

- **`request:hostname`**: `example.com`
- **`request:port`**: `8042`
- **`request:path`**: `/over/there`
- **`request:query`**: `name=ferret`
- **`request:uri`**: `foo://example.com:8042/over/there?name=ferret`

### Connection Functions

<table>
<thead>
<tr>
<th>Request</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>request:address</code></td>
<td>Returns the IP address of the server.</td>
</tr>
<tr>
<td><code>request:remote-hostname</code></td>
<td>Returns the fully qualified hostname of the client that sent the request.</td>
</tr>
<tr>
<td><code>request:remote-address</code></td>
<td>Returns the IP address of the server.</td>
</tr>
</tbody>
</table>
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 parameters available from the HTTP request. If RESTXQ is used, this function may help to find query parameters that have not been bound by %restxq:query-param annotations.

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 parameter in an HTTP request. If the parameter does not exist, an empty sequence or the optionally specified default value is returned instead.

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 may help to find headers that have not been bound by %restxq:header-param annotations.

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 may help to find cookies that have not been bound by %restxq:cookie-param annotations.
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 7.7

- Added: request:attribute

This module was introduced with Version 7.5.
Chapter 56. SQL Module

Read this entry online in the BaseX Wiki.

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.

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($url as xs:string) as xs:integer
sql:connect($url as xs:string, $user as xs:string, $password as xs:string) as xs:integer
sql:connect($url as xs:string, $user as xs:string, $password as xs:string, $options as item()) 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, which can either be specified

• as children of an <sql:options/> element, e.g.:

  <sql:options>
    <sql:autocommit value='true'/>
    ...
  </sql:options>

• as map, which contains all key/value pairs:

  map { "autocommit" := "true", ... }

Errors

BXSQ0001: an SQL exception occurs, e.g. missing JDBC driver or not existing relation.

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()*
### SQL Module

<table>
<thead>
<tr>
<th>Function</th>
<th>Signature</th>
<th>Summary</th>
<th>Errors</th>
</tr>
</thead>
</table>
| `sql:execute-prepared` | `sql:execute-prepared($id as xs:integer, $params as element(sql:parameters)) as element()*` | This function executes a prepared statement. The parameter $id specifies a prepared statement handle. The optional parameter $params is an element `<sql:parameters/>` representing the parameters for a prepared statement along with their types and values. The following schema shall be used:  
```xml
<element sql:parameters {
  <element sql:parameter {
    attribute type { "int"|"string"|"boolean"|"date"|"double"|"float"|"short"|"time"|"timestamp" },
    attribute null { "true"|"false" }?,
    text ?
  }
}
```
**Errors**  
BXSQ0001: an SQL exception occurs, e.g. not existing relation is retrieved. BXSQ0002: a wrong prepared statement handle is passed. BXSQ0003: the number of `<sql:parameter/>` elements in `<sql:parameters/>` differs from the number of placeholders in the prepared statement. BXSQ0004: the type of a parameter for a prepared statement is not specified. BXSQ0005: an attribute different from `type` and `null` is set for a `<sql:parameter/>` element. BXSQ0006: a parameter is from type date, time or timestamp and its value is in an invalid format. |
| `sql:prepare`     | `sql:prepare($connection as xs:integer, $statement as xs:string) as xs:integer` | This function prepares a statement and returns a handle to it. The parameter $connection indicates the connection handle to be used. The parameter $statement is a string representing an SQL statement with one or more '?' placeholders. If the value of a field has to be set to NULL, then the attribute `null` of the element `<sql:parameter/>` has to be true. |
| `sql:commit`      | `sql:commit($connection as xs:integer) as empty-sequence()` | This function commits the changes made to a relational database. $connection specifies the connection handle. |
| `sql:rollback`    | `sql:rollback($connection as xs:integer) as empty-sequence()` | This function rolls back the changes made to a relational database. $connection specifies the connection handle. |
| `sql:close`       | `sql:close($connection as xs:integer) as empty-sequence()` | |

---

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  
BXSQ0001: an SQL exception occurs, e.g. not existing relation is retrieved. BXSQ0002: a wrong connection handle is passed.
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:

```sql
let $conn := sql:connect("jdbc:postgresql://localhost:5432/coffeehouse")
return sql:execute($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>
<sql:row xmlns:sql="http://basex.org/modules/sql">
  <sql:column name="cof_name">French_Roast_Decaf</sql:column>
  <sql:column name="sup_id">49</sql:column>
  <sql:column name="price">7.5</sql:column>
  <sql:column name="sales">10</sql:column>
  <sql:column name="total">14</sql:column>
</sql:row>
<sql:row xmlns:sql="http://basex.org/modules/sql">
  <sql:column name="cof_name">Colombian_Decaf</sql:column>
  <sql:column name="sup_id">101</sql:column>
  <sql:column name="price">8.75</sql:column>
  <sql:column name="sales">6</sql:column>
  <sql:column name="total">12</sql:column>
  <sql:column name="date">2010-10-10 13:56:11.0</sql:column>
</sql:row>
```

Prepared Statements

A prepared select statement can be executed in the following way:

```sql
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 <code>&lt;sql:parameter/&gt;</code> elements in <code>&lt;sql:parameters/&gt;</code> differs from the 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 <code>type</code> and <code>null</code> is set for a <code>&lt;sql:parameter/&gt;</code> 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 57. 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:

```
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:

```
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.
# Session Module

## 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**

```xquery
session:get($key as xs:string) as xs:string?
session:get($key as xs:string, $default as xs:string) as xs:string
```

**Summary**

Returns the value of a variable bound to the current session. If the variable does not exist, 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**

```xquery
session:set($key as xs:string, $value as xs:string) as empty-sequence()
```

**Summary**

Assigns a value to a session variable.

**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**

```xquery
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**

```xquery
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>
A function was called outside the servlet context.

**Changelog**

This module was introduced with Version 7.5.
Chapter 58. 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:

  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

Signatures  
sessions:ids() as xs:string

Summary  
Returns the IDs of all registered sessions.

sessions:created

Signatures  
sessions:created($id as xs:string) as xs:dateTime

Summary  
Returns the creation time of the session specified by $id.

sessions:accessed

Signatures  
sessions:accessed($id as xs:string) as xs:dateTime

Summary  
Returns the last access time of the session specified by $id.

sessions:names

Signatures  
sessions:names($id as xs:string) as xs:string*

Summary  
Returns the names of all variables bound to the session specified by $id.

sessions:get

Signatures  
sessions:get($id as xs:string, $key as xs:string) as xs:string?
sessions:get($id as xs:string, $key as xs:string, $default as xs:string) as xs:string
### Sessions Module

| **Summary** | Returns the value of a variable bound to the session specified by $id$. If the variable does not exist, an empty sequence or the optionally specified default value is returned. |
| **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 xs:string) as empty-sequence() |
| **Summary** | Assigns a value to a variable bound 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

<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>
<tr>
<td>BXSE0003</td>
<td>A function was called outside the servlet context.</td>
</tr>
<tr>
<td>BXSE0004</td>
<td>The specified session was not found.</td>
</tr>
</tbody>
</table>

#### Changelog

This module was introduced with Version 7.5.
Chapter 59. Streaming Module

Read this entry online in the BaseX Wiki.

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:

```
```

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**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>stream:materialize($item as item()) as item()</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Returns a materialized instance of the specified $item:</td>
</tr>
<tr>
<td></td>
<td>• if an item is streamable, its value will be retrieved, and a new item containing the value will be returned.</td>
</tr>
</tbody>
</table>
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).

<table>
<thead>
<tr>
<th>Example</th>
<th>In the following example, a file will be deleted before its content is returned. To avoid a &quot;file not found&quot; error, the content will first be materialized:</th>
</tr>
</thead>
</table>
|         | let $file := 'data.txt'
|         | let $data := stream:materialize(file:read-text($file))
|         | return (file:delete($file), $data) |

**stream:is-streamable**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>stream:is-streamable($item as item()) as item()</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Checks whether the specified $item is streamable.</td>
</tr>
</tbody>
</table>

**Changelog**

This module was introduced with Version 7.7.
Chapter 60. Unit Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains annotations and functions for performing Unit tests with XQuery.

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, automatized way for testing software. The XUnit frameworks (such as SUnit or JUnit) allow testing of atomic unit of a program, such as single functions and algorithms.

This module borrows heavily from the existing frameworks: it introduces various new 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.

Please note that this module is still in beta stage, and its functionality is still subject to change. Your feedback is welcome.

Conventions

Both 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", <ERROR>)

Summary

With this annotation, a function can be marked as unit test. It will be evaluated whenever a test report is created for the module in which this function is located. If an optional error code is specified and if the function expression does not raise that error, the test will fail.

%unit:before

Syntax

%unit:before

Summary

A function decorated with this annotation will be evaluated before each unit test.

%unit:after

Syntax

%unit:after

Summary

A function decorated with this annotation will be evaluated after each unit test.

%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()*, $message as xs:string) 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. If the optional error $message can be specified as second argument.

**Errors**  
UNIT0001: the assertion failed, or an error was raised.

#### unit:fail

**Signatures**  
unit:fail($message as xs:string) as empty-sequence()

**Summary**  
Raises a unit error with the specified message.

**Errors**  
UNIT0001: default error raised by this function.

#### unit:test

**Signatures**  
unit:test() as element(testsuite)*  
unit:test($functions as function(*)) as element(testsuite)*

**Summary**  
Runs all functions, or the specified $functions, in the current query context that have unit annotations. A test report is generated and returned, which resembles the format returned by other xUnit testing frameworks, such as the Maven Surefire Plugin.

**Errors**  
UNIT0002: a test function must have no arguments.UNIT0003: a test function must not be updating.UNIT0004: an annotation was declared twice.UNIT0005: an annotation has invalid arguments.

**Examples**  
unit:test(inspect:context()) runs all unit tests found in the query context.

#### unit:test-uris

**Signatures**  
unit:test-uris($uris as xs:string*) as element(testsuites)

**Summary**  
Runs all functions in the specified modules that have unit annotations. A test report is generated and returned, which resembles the format returned by other xUnit testing frameworks, such as the Maven Surefire Plugin.

**Errors**  
UNIT0002: a test function must have no arguments.UNIT0003: a test function must not be updating.UNIT0004: an annotation was declared twice.UNIT0005: an annotation has invalid arguments.
Example

The following XQuery main module creates a test report. It contains all available unit annotations:

Query:

```xquery
(:~ Initializing function, which is called once before all tests. :) declare %unit:before-module function local:before-all-tests() { } ; (:~ Initializing function, which is called once after all tests. :) declare %unit:after-module function local:after-all-tests() { } ; (:~ Initializing function, which is called before each test. :) declare %unit:before function local:before() { } ; (:~ Initializing function, which is called after each test. :) declare %unit:after function local:after() { } ; (:~ Function demonstrating a successful test. :) declare %unit:test function local:success-function() { unit:assert(1 + 2 = 3) } ; (:~ Function demonstrating a failure. :) declare %unit:test function local:failure-function() { unit:assert(4 + 5 = 6) } ; (:~ Function demonstrating an expected error. :) declare %unit:test("expected", "FORG0001") function local:expected-success() { } ; (:~ Function demonstrating an expected error. :) declare %unit:test("expected", "FORG0001") function local:expected-error() { 1 + <a/> } ; (:~ Function demonstrating an error. :)```
declare
%unit:test
function local:error-function() {
  1 + a
};

(:~ Skipping a test. :) declare
declare %unit:test %unit:ignore("Skipped!")
function local:skipped-function() {
  ()
};

(: run all tests :) unit:test()

Result:
<testsuite name="/path/to/tests.xq" time="PT0S" tests="6" failures="2" errors="1"
skipped="1">
  <testcase name="success-function" time="PT0S"/>
  <testcase name="failure-function" time="PT0S">
    <failure message="Assertion failed." type="UNIT0001"/>
  </testcase>
  <testcase name="expected-success" time="PT0S">
    <failure message="Error expected." type="FORG0001"/>
  </testcase>
  <testcase name="expected-error" time="PT0S"/>
  <testcase name="error-function" time="PT0S">
    <error message="Invalid xs:double cast: ." type="FORG0001"/>
  </testcase>
  <testcase name="skipped-function" time="PT0S">
    <skipped message="Skipped!"/>
  </testcase>
</testsuite>

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>
<tr>
<td>UNIT0003</td>
<td>A test function must not be updating.</td>
</tr>
<tr>
<td>UNIT0004</td>
<td>An annotation was declared twice.</td>
</tr>
<tr>
<td>UNIT0005</td>
<td>An annotation has invalid arguments.</td>
</tr>
</tbody>
</table>

Changelog

This module was introduced with Version 7.7.
Chapter 61. Validation Module

Read this entry online in the BaseX Wiki.

This XQuery Module contains functions to perform validations against XML Schema and Document Type Declarations. By default, this module uses Java’s standard validators. As an alternative, Saxon XSLT Processor is used if (saxon9he.jar, saxon9pe.jar or saxon9ee.jar) is added to the classpath.

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.

Functions

validate:xsd

Signatures

validate:xsd($input as item()) as empty-sequence()
validate:xsd($input as item(), $schema as item()) as empty-sequence()

Summary

Validates the document specified by $input. Both $input and $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.

$schema can be used to specify the schema for validation. If no schema is given, $input is required to contain an xsi:(noNamespace)schemaLocation attribute as defined in W3C XML Schema.

Errors

BXVA0001: the validation fails. BXVA0002: the validation process cannot be started.

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:

  let $doc := <simple:root xmlns:simple='http://basex.org/simple'/>
  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 document specified by $input and returns warning, errors and fatal errors in a string sequence. $input and $schema can be specified as:

- xs:string, containing the path to the resource,
- xs:string, containing the resource in its string representation, or
**Validation Module**

- node(), containing the resource itself.

$schema can be used to specify the schema for validation. If no schema is given, $input is required to contain an xsi:(noNamespace)schemaLocation attribute as defined in W3C XML Schema.

**Errors**

- BXVA0002: the validation process cannot be started.

**validate:dtd**

**Signatures**

- validate:dtd($input as item()) as empty-sequence()
- validate:dtd($input as item(), $dtd as xs:string) as empty-sequence()

**Summary**

Validates the document specified by $input. $input 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.

$schema can be used to specify the DTD for validation. If no DTD is given, $input is required to contain a DTD doctype declaration.

**Errors**

- BXVA0001: the validation fails.
- BXVA0002: the validation process cannot be started.

**Examples**

- validate:xsd('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 $dtd := '<!ELEMENT root (#PCDATA)'>
    return validate:dtd($doc, $dtd)
} catch BXVA0001 {
    'DTD Validation failed.'
}
```

**validate:dtd-info**

**Signatures**

- validate:dtd-info($input as item()) as xs:string*
- validate:dtd-info($input as item(), $dtd as xs:string) as xs:string*

**Summary**

Validates the document specified by $input and returns warning, errors and fatal errors in a string sequence. $input 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.

$schema can be used to specify the DTD for validation. If no DTD is given, $input is required to contain a DTD doctype declaration.

**Errors**

- BXVA0002: the validation process cannot be started.

**Errors**

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
</table>

254
BXVA0001 The document cannot be validated against the specified DTD or XML Schema.

BXVA0002 The validation cannot be started.

**Changelog**

Version 7.6

- **Added**: validate:xsd-info, validate:dtd-info

The module was introduced with Version 7.3.
Chapter 62. 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**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>xquery:eval($query as xs:string) as item()* xquery:eval($query as xs:string, $bindings as map(<em>)) as item()</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>Evaluates $query as XQuery expression at runtime and returns the resulting items. Variables and context items can be declared via $bindings. The specified keys must be QNames or strings, the values can be arbitrary item sequences:</td>
</tr>
<tr>
<td></td>
<td>• variables specified as QNames will be directly interpreted as variable name.</td>
</tr>
<tr>
<td></td>
<td>• variables specified as xs:string may be prefixed with a dollar sign. Namespace can be specified using the Clark Notation.</td>
</tr>
<tr>
<td></td>
<td>• If the specified string is empty, the value will be bound to the context item.</td>
</tr>
<tr>
<td>Errors</td>
<td>BXXQ0001: the query contains updating expressions.</td>
</tr>
<tr>
<td>Examples</td>
<td>• xquery:eval(&quot;1+3&quot;) returns 4.</td>
</tr>
<tr>
<td></td>
<td>• You can bind the context and e.g. operate on a certain database only: xquery:eval(&quot;//country&quot;, map{ '' := db:open('factbook') })</td>
</tr>
<tr>
<td></td>
<td>• The following expressions use strings as keys. All of them return 'XML': xquery:eval(&quot;.&quot;#, map{ '' := 'XML' }), xquery:eval(&quot;$xml&quot;, map{ 'xml' := 'XML' }), xquery:eval(&quot;$xml&quot;, map{ '$xml' := 'XML' }), xquery:eval(&quot;declare namespace pref='URI'; $pref:xml&quot;, map{ '{URI}xml' := 'XML' }),</td>
</tr>
<tr>
<td></td>
<td>• The following expressions use QNames as keys. All of them return 'XML': declare namespace pref = 'URI'; xquery:eval(&quot;$xml&quot;, map{ xs:QName('xml') := 'XML' }), xquery:eval(&quot;declare namespace pref='URI'; $pref:xml&quot;, map{ xs:QName('pref:xml') := 'XML' }),</td>
</tr>
</tbody>
</table>

**xquery:invoke**

<table>
<thead>
<tr>
<th>Signatures</th>
<th>xquery:invoke($uri as xs:string) as item()* xquery:invoke($uri as xs:string, $bindings as map(<em>)) as item()</em></th>
</tr>
</thead>
</table>
**Summary**
Opens $uri as file, evaluates it as XQuery expression at runtime, and returns the resulting items. The semantics of the $bindings parameter is the same as for xquery:eval.

**Errors**
BXXQ0001: the query contains updating expressions.

### 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**

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXXQ0001</td>
<td>A dynamically evaluated query must not contain any updating expressions.</td>
</tr>
</tbody>
</table>

**Changelog**

This module was introduced with Version 7.3. Functions have been adopted from the obsolete Utility Module.
Chapter 63. XSLT Module

Read this entry online in the BaseX Wiki.

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:

```
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 item()) 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, which can either be specified as

- children of an <xslt:parameters/> element; e.g.:
XSLT Module

• or as map, which contains all key/value pairs:

map { "key1" := "value1", ... }

Note that only strings are supported when using Saxon (XSLT 2.0).

Error

BXSL0001: an error occurred during the transformation process.

XSLT Transform-text

Signatures

xslt:transform-text($input as item(), $stylesheet as item()) as xs:string
xslt:transform-text($input as item(), $stylesheet as item(), $params as item()) as xs:string

Summary

Transforms the document specified by $input, using the XSLT template specified by $stylesheet, and returns the result as string. The parameters are the same as described 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:

xslt:transform-text(<dummy/>, 'basic.xslt')

Result:

123

Example 2: XSLT transformation of an input document

Query:

(: 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>
</books>
XSLT Module

Example 3: Assigning a variable to an XSLT stylesheet

Query:

```xml
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

```xml
<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 64. ZIP Module

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. It may soon be replaced by the Archive Module.

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

Signatures

| zip:binary-entry($uri as xs:string, $path as xs:string) as xs:base64Binary |

Summary

Extracts the binary file at $path within the ZIP file located at $uri and returns it as an xs:base64Binary item.

Errors

ZIP0001: the specified path does not exist.ZIP0003: the operation fails for some other reason.

zip:text-entry

Signatures

| zip:text-entry($uri as xs:string, $path as xs:string) as xs:string |
| zip:text-entry($uri as xs:string, $path as xs:string, $encoding as xs:string) as xs:string |

Summary

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.

Errors

ZIP0001: the specified path does not exist.ZIP0003: the operation fails for some other reason.

zip:xml-entry

Signatures

| zip:xml-entry($uri as xs:string, $path as xs:string) as document-node() |

Summary

Extracts the XML file at $path within the ZIP file located at $uri and returns it as a document node.

Errors

FODC0006: the addressed file is not well-formed.ZIP0001: the specified path does not exist.ZIP0003: the operation fails for some other reason.

zip:html-entry

Signatures

| zip:html-entry($uri as xs:string, $path as xs:string) as document-node() |

Summary

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.

Errors

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.
ZIP Module

zip:entries

Signatures  
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:

```
```

zip:zip-file

Signatures  
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  
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:

```
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 65. 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

JavaDoc

The project’s JavaDoc can be explored online.

HTTP Services

- **RESTXQ**: Write web services with XQuery
- **REST**: Access and update databases via HTTP requests
- **WebDAV**: Access databases from your filesystem

APIs

- **Clients**: Communicate with BaseX using C#, PHP, Python, Perl, C, ...
- **Java Examples**: Code examples for developing with BaseX
- **XQJ API**, implemented by Charles Foster
- **XQuery for Scala API**, based on XQJ and written by Dino Fancellu
Chapter 66. 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 being developed with the Eclipse environment. Other IDEs are used as well in our community, but are not supported by our team.

- The EGit plugin can be used to check out the latest sources from our repository within Eclipse.

- The m2eclipse plugin is required to work with packages other than the main project; it adds Maven support to Eclipse.

- Additional coding guidelines are defined via Checkstyle and can be integrated with the eclipse-es plugin.

- Other Eclipse plugins we frequently use are FindBugs to analyze Java byte code, and Core Tools to find unreferenced members.

Check Out

To get some help on how to check out BaseX from the GitHub Repository, please have a look at our Git Tutorial.

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

With the Maven plugin from Eclipse, it may require 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 67. 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. If forking sounds unfamiliar to you, we suggest to check out the git introduction below.

Using Git & Eclipse

1. (Optional) Head over to https://github.com/BaseXdb and create an account
2. Fork BaseX, so you have a version on your own
3. Make yourself familiar with git (see the end of this page)
4. Open Eclipse
5. Install egit (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.
• 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 -1 $PWD/*

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
[master 0fde1fb] Added TODO in section "USING GIT"
1 files changed, 4 insertions(+), 0 deletions(-)
```
### 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:

```
$ git push
```

Check out the links at the end of the page for more git options.
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 68. 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` repository to start the command-line mode,

• type in `mvn jetty:run` in the `basex-api` repository 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 69. 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

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

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/

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/

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/

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

• Ubuntu: http://launchpad.net/ubuntu/+source/basex

• Arch Linux: http://aur.archlinux.org/packages.php?ID=38645
Chapter 70. 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
- **Indonesian**: Andria Arisal
- **Italian**: Massimo Franceschet
- **Japanese**: Toshio HIRAI and Kazuo KASHIMA
- **Mongolian**: Tuguldur Jamiyansharav
- **Romanian**: Adrian Berila
- **Russian**: Oleksandr Shpak and Max Shamaev

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 Maven, 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

**Updating BaseX.jar**

You may 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 int 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 may also change the language in the .basetx configuration file, which is placed in your home directory. In order to see where the all text keys are used within BaseX, you may temporarily set the LANGKEY option to true.
Chapter 71. 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 distributed XML resources. 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/. By default, the HTTP server is started with admin credentials (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

<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

<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:database>

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.

Operations

GET and POST requests support the following operations:

- **Query**: Evaluates an XPath/XQuery expression. If a database or database path is specified in the URL, it is used as initial query context.
- **Run**: Runs a query file located on the server. The query directory is defined by the WEBPATH option.
- **Command**: Executes a database command.
- **Get**: This is the default operation for the GET operation (it is not possible to use this operation in a POST request). It returns a list of all databases, the resources of a database or the addressed resource.
Parameters

Additionally, 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.
- **Wrap**: The wrap parameter encloses all query results with XML elements, using the http://basex.org/rest namespace.

While Options can be specified for all operations, the remaining parameters will only make sense for **Query** and **Run**.

Request Methods

GET Requests

If the GET method is used, all query parameters are directly specified within the URL.

Examples

- The first example lists all resources found in the tmp path of the factbook database:
  ```
  http://localhost:8984/rest/factbook/tmp
  ```
- The first example prints the city names from the factbook database and encloses all results with a `<rest:result/>` element:
  ```
  ```
- In the next request, US-ASCII is chosen as output encoding, and the query eval.xq is evaluated:
  ```
  ```
- The next URL turns on XML wrapping and lists all database users that are known to BaseX:
  ```
  http://localhost:8984/rest?command=show+users
  ```
- The last example includes an to disallow XQuery 3.0 expressions:
  ```
  http://localhost:8984/rest?query=12345&xquery3=false
  ```

POST Requests

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:
  ```xml
  <query xmlns="http://basex.org/rest">
    <text><![CDATA[ (//city/name)[position() <= 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:
  ```xml
  <query xmlns="http://basex.org/rest">
    <text></text>
  </query>
  ```
<rest:query xmlns:rest="http://basex.org/rest">  
  <rest:text>for $i in .//text() return string-length($i)</rest:text>  
</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 Requests

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 : If application/octet-stream is chosen as content-type, the input data is added as raw.

• 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.
  - application/jsonml+json: Stores JSONML input 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.

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 Requests

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 Requests

All query parameters that have not been processed before will be treated as variable assignments:

#### Examples

- The following request binds a single variable to the query to be processed:
  ```
  http://localhost:8984/rest?query=$text&$text=Hello+World
  ```

- 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 Requests

If **query** or **run** is used as operation, external variables can be specified via the `<variable/>` element:

```
<query xmlns="http://base64.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>
```
Content Type

As the content type of a REST response cannot be dynamically determined in all cases, it can be manually adjusted by the user. The final content type of a REST response is chosen in several steps:

1. By default, the content type of a response depends on the chosen operation:
   - Query/Run → application/xml
   - Command → text/plain
   - Get → application/xml, or content type of the addressed resource

2. The default content type is overwritten if a serialization method is specified, either as query parameters or within the XQuery expression. The following method/content-type mappings are available:
   - xml → application/xml
   - xhtml → text/html
   - html → text/html
   - text → text/plain
   - raw → application/octet-stream
   - json or jsonml → application/json

3. The content type is overwritten in any case if a specific media-type is chosen, again as query parameter or within the query.

The following three example requests will all return `<a/>` as result and use application/xml as content-type:

`http://localhost:8984/rest?query=%3Ca/%3E`  
`http://localhost:8984/rest?query=%3Ca/%3E&method=xml`  
`http://localhost:8984/rest?query=%3Ca/%3E&mediatype=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();
}

Content-Types

The content-type of the input can easily be included, just add the following property to the connection (in this example we explicitly store the input file as raw):

// 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 7.2

- Removed: direct evaluation of addresses resources with `application/xquery` as content type

Version 7.1.1

- Added: `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 72. REST: POST Schema

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>
```
Chapter 73. 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 a new 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.

Note that some extensions presented in this documentation are BaseX-specific; they may be integrated in future versions of the RESTXQ draft.

With Version 7.7, the following features have been added:

• RESTXQ functions may also be specified in main modules (extension: .xq)
• multipart types are now supported, including multipart/form-data
• a new %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 RESTXQ namespace prefix has been changed to rest
• by default, RESTXQ is now available on top level via http://localhost:8984/

Usage


All RESTXQ annotations are assigned to the http://exquery.org/ns/restxq namespace, which is statically bound to the rest prefix (before Version 7.7: restxq). 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 library modules (detected by the extension .xqm) and functions with RESTXQ annotations.

Since Version 7.7, also main modules (detected by .xq) will be parsed; the main expression of the main module will never be evaluated. All modules will be cached and parsed again when their timestamp changes.

A simple RESTXQ module is shown below. It is part of a clean installation and available at http://localhost:8984/.

(module namespace page = 'http://basex.org/examples/web-page';

declare %rest:path("hello/{$world}")
%rest:GET
%rest:header-param("User-Agent", "{$agent}")
function page:hello($world as xs:string, $agent as xs:string*) {

<title> Hello {$world}! </title>
<info>You requested this page with {$agent}. </info>

</response>

);
If the URI http://localhost:8984/hello/World is accessed, the result will be similar to:

```xml
<response>
  <title>Hello World!</title>
  <time>The current time is: 18:42:02.202+02:00</time>
</response>
```

The RESTXQ module contains yet another function:

```xml
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/)... 

`curl -i -X POST --data "content='CONTENT'" http://localhost:8984/form`

...you will receive something similar to the following result:

```
HTTP/1.1 200 OK
Content-Type: application/xml; charset=UTF-8
Content-Length: 107
Server: Jetty(8.1.11.v20130520)

<response type="form">
  <message>(no message)</message>
  <user-agent>cURL/7.31.0</user-agent>
</response>
```

**Requests**

This section shows how annotations are used to handle and process HTTP requests.

**Constraints**

Constraints restrict the HTTP requests that a resource function may process.

**Paths**

A resource function must have a single Path Annotation with a single string as argument. The function will be called if a URL matches the path segments and templates of the argument. Path templates contain variables in curly brackets, and map the corresponding segments of the request path to the arguments of the resource function.

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 $variable will be cast to an xs:integer before being bound:
declare %rest:path("/a/path/{$with}/some/{$variable}")
  function page:test($with, $variable as xs:integer) { ... }

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:

  %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:

  %rest:produces("application/atom+xml")

By default, both mime types are */*. 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

The HTTP method annotations are equivalent to all HTTP request methods except for 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:

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:

declare %rest:PUT("{$body}") %rest:path("/put")
  function page:put($body) { "Request body: " || $body };

If a content-type is specified in the request, the content is converted to the following XQuery type:

<table>
<thead>
<tr>
<th>Content-Type</th>
<th>XQuery type</th>
</tr>
</thead>
<tbody>
<tr>
<td>application/json, +json</td>
<td>document-node() (conversion is described in the JSON Module)</td>
</tr>
<tr>
<td>text/html</td>
<td>document-node() (conversion is described in the HTML Module)</td>
</tr>
<tr>
<td>text/comma-separated-values</td>
<td>document-node()</td>
</tr>
<tr>
<td>text/xml, application/xml</td>
<td>document-node()</td>
</tr>
<tr>
<td>text/*</td>
<td>xs:string</td>
</tr>
<tr>
<td>others</td>
<td>xs:base64Binary</td>
</tr>
<tr>
<td>multipart/*</td>
<td>sequence (see next paragraph)</td>
</tr>
</tbody>
</table>

Multipart Types

*Introduced with Version 7.7:

Some first support for multipart content-types has been added. The 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:
Please note that support for multipart types is still experimental, and it may change in a future version BaseX. Your feedback is welcome.

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):

HTML Form Fields

Form parameters are specified the same way as query parameters. Their values are extracted from GET or POST requests.

File Uploads

Introduced with Version 7.7:

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):

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:
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", "{$user-agent}")
%rest:header-param("Referer", "{$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")

Responses
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 Responses
Custom responses can be built from within XQuery by returning a 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:

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 Web Applications, precisely 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

<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**

```xml
<rest:redirect>{ $location }</rest:redirect>
```

...is basically an abbreviation for...

```xml
<http:response status="302" message="Temporary Redirect">
  <http:header name="location" value="{ $location }"/>
</http:response>
```

The client decides whether to follow this redirection. Browsers usually will, tools like `curl` won’t unless `-L` is specified.

### Output

Similar to the REST interface, result serialization can be modified via XQuery 3.0 serialization parameters; in RESTXQ, serialization parameters may be specified in the query prolog, via annotations, or within REST response element. Global parameters are overwritten by more local parameters.

#### 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:

```xml
declare option output:media-type 'text/plain';
```

```xml
declare %rest:path("version1") function page:version1() {
  'Keep it simple, stupid'
};
```

#### Annotations

The serialization can also be parameterized via annotations:

```xml
declare %output:media-type("text/plain") %rest:path("version2") function page:version2() {
  'Still somewhat simple.'
};
```

#### Response Element

The following example demonstrates how serialization parameters can be dynamically set within a query:

```xml
declare %rest:path("version3") function page:version3() {
  <rest:response>
    <output:serialization-parameters>
```
The content type can also be overwritten by specifying an output method. The following method mappings are available:

- xml → application/xml
- xhtml → text/html
- html → text/html
- text → text/plain
- raw → application/octet-stream
- json or jsonml → application/json

By default, application/xml is returned as content type. In the following example, XHTML headers will be generated, and text/html will be set as content type:

```restxq
declare %rest:path("")
%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:start() {
    <html xmlns="http://www.w3.org/1999/xhtml">
        <body>done</body>
    </html>
}
```

### Error Handling

*Introduced with Version 7.7:*

#### XQuery Errors

XQuery runtime errors can be processed via error annotations. A single argument must be supplied, which represents the QName of the error to be caught. A wildcard * may be specified to catch all possible errors. A function can only have a single error annotation:

```restxq
declare %rest:error("*") function page:error() {
    "An error occurred while processing your RESTXQ code!"
}
```

The XQuery try/catch construct assigns error information to a number of pre-defined variables (code, description, value, module, line-number, column-number, additional). These variables 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 the following example shows:

```restxq
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) {
        'User "' || $user || '" is unknown'
    };

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):

<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 additional
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). With Version 7.7, a RESTXQ Module was added that 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
RESTXQ has been proposed by Adam Retter. More information on all specifics can be found in the following
two documents:

• RESTful XQuery, Standardised XQuery 3.0 Annotations for REST . Paper, XMLPrague, 2012
• RESTXQ . Slides, MarkLogic User Group London, 2012
• RESTXQ Specification , Unofficial Draft
• Web Application, RESTXQ Development . Web Application Development with RESTXQ Slides from
XMLPrague 2013

Changelog
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 74. 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 WebDAV service is activated and is accessible at http://localhost:8984/webdav/. Generally it can be accessed using either with a http://<httphost>:<httpport>/webdav/ or webdav://<httphost>:<httpport>/webdav/ URL, depending on the used 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.

Locking

BaseX WebDAV implementation supports WebDAV locking since Version 7.7. 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. Implementations should not rely on the existence of the database, since it may not be accessible in the future.

WebDAV Clients

Please check out the following tutorials to get WebDAV running on different operating systems and with oXygen:

- Windows 7
- 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 75. 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 76. 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:
WebDAV: KDE

![WebDAV Interface]

A screenshot of a WebDAV interface showing directories and files.
Chapter 77. 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:
Chapter 78. WebDAV: Windows 7

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."

• Click "Next", select "Choose a custom network location" and click "Next" again.

• Enter the URL address of the BaseX WebDAV Server (e.g. http://localhost:8984/webdav) and click "Next".
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 79. WebDAV: Windows XP

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":

• Confirm the upcoming introductory dialog:

• Select "Choose another network location" in the next dialog:
WebDAV: Windows XP

• 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:
### WebDAV: Windows XP

![WebDAV window](Image)

<table>
<thead>
<tr>
<th>Name</th>
<th>Internet Address</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>factbook</td>
<td><a href="http://localhost:8994/webdav/factbook">http://localhost:8994/webdav/factbook</a></td>
<td>Web Folder</td>
</tr>
</tbody>
</table>

1 objects
Part IX. Client APIs
Chapter 80. Clients

Read this entry online in the BaseX Wiki.

This page is part of the Developer Section. It describes how to use BaseX from other programming languages.

With the following light-weight bindings in different programming languages, you will be able to connect to a running BaseX server instance, execute database commands, perform queries or listen to events. Most clients offer the following two classes:

- **Standard Mode**: connecting to a server, sending commands
- **Query Mode**: defining queries, binding variables, iterative evaluation

Please have a look at our Server Protocol for more information on the clients and the underlying protocol. Bindings for other languages are easy to write; your contributions are welcome.

Currently, we offer bindings for the following programming languages:

Object oriented

- **C#**, **VB**, **Scala**, **Java**
  - **Scala**: contributed by Manuel Bernhardt
  - **ActionScript**: contributed by Manfred Knobloch

Scripting

- **Perl**
  - **PHP** (example) updated by James Ball
  - **Python 3.x, 2.7.3**: contributed by Hiroaki Itoh
  - **Python < 2.7**: improved by Arjen van Elteren
  - **Rebol**: contributed by Sabu Francis
  - **Ruby**

Functional

- **Haskell**: contributed by Leo Wörteler
  - **Lisp**: contributed by Andy Chambers

Others

- **node.js**: contributed by Andy Bunce
  - **node.js**: contributed by Hans Hübner (deviating from client API)
  - **Qt**: contributed by Hendrik Strobelt
  - **C**

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.
• EventExample demonstrates how to watch and unwatch Events.
Chapter 81. 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.
- **ServerEventsGUI.java** is a little GUI example for demonstrating database events.
- **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.
- `EventExample.java` demonstrates how to trigger and receive database events.
- `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 82. PHP Example

Read this entry online in the BaseX Wiki.

This page is referenced from the Clients page. It demonstrates how database commands and XQuery can be executed on a database server with a PHP client. The results of the query are stored into a DOM document and can be processed in several ways.

Requirements

- BaseXClient: BaseX PHP Client
- DOMExample: Example used in this tutorial
- any PHP Server, such as XAMPP

Setting up

- Install and start XAMPP, or choose a PHP Server of your own
- Copy BaseXClient.php and DOMExample.php to the XAMPP folder or upload it to your webserver

Usage

1. Start a Database Server instance on your local or a remote machine. Make sure the host and port settings in DOMExample.php are correct.
2. Call DOMExample.php in a web browser of your choice.
3. Look at the DOM document on the PHP documentation for further information on the DOM document functions.
4. Open DOMExample.php in an editor and edit it for your own needs.
Chapter 83. 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
 */

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();
    }
}
```

} catch (Exception $e) {
    // print exception
    print $e->getMessage();
}

// close session
session_close();

} catch (Exception $e) {
    // print exception
    print $e->getMessage();
}

?>

Changelog

Version 7.2

- Added: context() function
Chapter 84. 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.

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 in)`

• Add a document to the current database from an input stream: `void add(String path, InputStream in)`

• Replace a document with the specified input stream: `void replace(String path, InputStream in)`

• Store raw data at the specified path: `void store(String path, InputStream in)`

• Watch the specified event: `void watch(String name, Event notifier)`

• Unwatch the specified event: `void unwatch(String name)`

• Return process information: `String info()`
• Close the session: `void close()`

**Query**

• Create query instance with session and query:
  `Query(Session s, 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()`

**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.

**General Syntax**

• \x: single byte.

• {...}: utf8 strings or raw data, suffixed with a \0 byte. To avoid confusion with the suffix byte, all \0 and \xFF bytes that occur in raw data will be prefixed with \xFF.

**Authentication (via cram-md5)**

1. Client connects to server socket
2. Server sends timestamp (string representation of the current time in milliseconds):
   `{timestamp}`
3. Client sends username and hashed password/timestamp:
   `{username} {md5(md5(password) + timestamp)}`
4. Server replies with \0 (success) or \1 (error)

**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><code>{command}</code></td>
<td><code>{result} {info} \0</code></td>
<td>Executes a database command.</td>
</tr>
<tr>
<td>QUERY</td>
<td>\0 <code>{query}</code></td>
<td><code>{id} \0</code></td>
<td>Creates a new query instance and returns its id.</td>
</tr>
</tbody>
</table>
### Server Protocol

<table>
<thead>
<tr>
<th>Command</th>
<th>Request Example</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CREATE</td>
<td>\8 {name} (\text{input}) (\text{info}) \0</td>
<td>Creates a new database with the specified input (may be empty).</td>
</tr>
<tr>
<td>ADD</td>
<td>\9 {name} (\text{path}) (\text{info}) \0 (\text{input})</td>
<td>Adds a new resource to the opened database.</td>
</tr>
<tr>
<td>WATCH</td>
<td>\10 {name} (\text{info}) \0</td>
<td>Registers the client for the specified event.</td>
</tr>
<tr>
<td>UNWATCH</td>
<td>\11 {name} (\text{info}) \0</td>
<td>Unregisters the client.</td>
</tr>
<tr>
<td>REPLACE</td>
<td>\12 {path} (\text{input}) (\text{info}) \0</td>
<td>Replaces a resource with the specified input.</td>
</tr>
<tr>
<td>STORE</td>
<td>\13 {path} (\text{input}) (\text{info}) \0</td>
<td>Stores a binary resource in the opened database.</td>
</tr>
</tbody>
</table>

**# error**

\{ partial result \} \{error\} \1

Error feedback.

---

### 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>\2 (\text{id}) \0 \0</td>
<td>Closes and unregisters the query with the specified id.</td>
<td></td>
</tr>
<tr>
<td>BIND</td>
<td>\3 (\text{id}) (\text{name}) \0 \0 (\text{value}) (\text{type})</td>
<td>Binds a value to a variable. An empty string can be specified as data type.</td>
<td></td>
</tr>
<tr>
<td>RESULTS</td>
<td>\4 (\text{id}) \x (\text{item}) (\text{item}) \0</td>
<td>Returns the single items as strings, prefixed by a single byte (\x) that represents the Type ID. This command is called by the <code>more()</code> function of a client implementation.</td>
<td></td>
</tr>
<tr>
<td>EXECUTE</td>
<td>\5 (\text{id}) (\text{result}) \0</td>
<td>Executes the query and returns all results as a single string.</td>
<td></td>
</tr>
<tr>
<td>INFO</td>
<td>\6 (\text{id}) (\text{result}) \0</td>
<td>Returns a string with query compilation and profiling info.</td>
<td></td>
</tr>
<tr>
<td>OPTIONS</td>
<td>\7 (\text{id}) (\text{result}) \0</td>
<td>Returns a string with all query serialization parameters.</td>
<td></td>
</tr>
<tr>
<td>CONTEXT</td>
<td>\14 (\text{id}) (\text{value}) \0 \0 (\text{type})</td>
<td>Binds a value to the context item. An empty string can be specified as data type.</td>
<td></td>
</tr>
<tr>
<td>UPDATING</td>
<td>\30 (\text{id}) (\text{result}) \0</td>
<td>Returns <code>true</code> if the query may perform updates; <code>false</code> otherwise.</td>
<td></td>
</tr>
<tr>
<td>FULL</td>
<td>\31 (\text{id}) XDM (\text{item}) ... XDM (\text{item}) \0</td>
<td>Returns all resulting items as strings, prefixed by the XDM Meta Data. This command is e.g. used by the XQJ API.</td>
<td></td>
</tr>
</tbody>
</table>
As can be seen in the table, all results end with a single \0 byte, which indicates that the process was successful. If an error occurs, an additional byte \1 is sent, which is then followed by the error message string.

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.

- **Client** connects to the database server socket
- **Server** sends timestamp "1369578179679": $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** additionally sends hashed password/timestamp combination: md5(md5("topsecret") + "1369578179679") = "66442c0e3b5af8b9324f7e31b7f5cca8": $36 36 34 ... 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** returns the first result ("1", type xs:integer): $52 31 00
- **Server** sends a single "$0" byte instead of a new result, which indicates that no more results can be expected: $00
- **Server** sends the error code "$1" 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

Existing Clients

- Java client
- C# client
- Python client
- Perl client
- more client implementations are listed on the Clients page.

Changelog

Version 7.2
• Added: Query Commands CONTEXT, UPDATING and FULL

• Added: Client function `context(String value, String type)`
Chapter 85. Server Protocol: Types

Read this entry online in the BaseX Wiki.

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, though: 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:NMSTRING</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>
</tbody>
</table>
### Server Protocol: Types

<table>
<thead>
<tr>
<th>Number</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<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 86. 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 2005-11, 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\n";
} catch (Exception $e) {
    // print exception
    print $e->getMessage();
}
?>
```
Part X. Advanced User's Guide
Chapter 87. 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
- **Indexes**: Available index structures and their utilization
- **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
- **Events**: Description of the event feature
- **Execution Plan**: Analyzing query evaluation
Chapter 88. 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 89. 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:

```xml
<!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
<?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 90. 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.
- **.basexperm** contains user name, passwords, and permissions (see last paragraph).
- **.basexevents** contains all existing events (see Events).
- **.basexhistory** contains commands that have been typed in most recently.
- **.basexhome** can be created by a user to mark a folder as home directory (Introduced with Version 7.7)

Note that, depending on your OS and configuration, files and folders with a ‘.’ prefix may be hidden.

## 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.

## 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 datda.

The database path can be changed as follows:

- GUI: Choose Options → Preferences and choose a new database path.
- General: edit the DBPATH in the .basex configuration file

Note: Existing databases will not be automatically moved to the new destination.

## User and Log Files

Global users, along with their passwords and permissions, are stored in the .basexperm file in the home directory. Local users and permissions are stored inside the database files. Log files are stored in text format in a sub-directory .logs of the database folder (see Logging for more information).
Changelog

Version 7.7

Updated: the .baseXhome file marks a folder as home directory
Chapter 91. Events

Read this entry online in the BaseX Wiki.

This article is part of the Advanced User's Guide. It presents how to trigger database events and notify listening clients.

Introduction

The events feature enables users with admin permissions to create events, which can then be watched by other clients. All clients that have registered for an event will be notified if an event is triggered by another client.

Managing Events

CREATE EVENT [name] Creates an event [name].

DROP EVENT [name] Drops the event with the specified [name].

SHOW EVENTS Shows all events.

Watching/Unwatching Events

The events can currently be watched by the Java and C# clients. See the following Java code example:

Watch events:

```java
// name of the event
String event = "call";
// create new client
BaseXClient client = new BaseXClient("localhost", 1984, "admin", "admin");
// register for an event
client.watch(event, new EventNotifier() {
    @Override
    public void notify(final String value) {
        System.out.println("Received data: " + value);
    }
});
```

Unwatch events:

```java
// unregister from an event
client.unwatch(event);
```

For a complete and self-contained example in Java, you may have a look [1].

Firing Events

Events are triggered via the XQuery function `db:event()`:

```
db:event($name as xs:string, $query as item())
```

Executes a $query and sends the resulting value to all clients watching the event with the specified $name. No event will be sent to the client that fired the event.

Example Scenarios

Basic

1. Client1 creates an event with the name "EVENT"
2. **Client2** and **Client3** call the watch method for event "EVENT"

3. **Client1** executes XQuery
   
   ```xml
   db:event("EVENT", "1 to 2")
   ```

4. **Client2** and **Client3** will receive the result 1 2

5. **Client2** executes XQuery
   
   ```xml
   db:event("EVENT", "2 to 3")
   ```

6. **Client3** will receive the result 2 3

**Included in Update Expression**

1. **Client1** creates an event with the name "DELETED"

2. **Client2** and **Client3** call the watch method for event "DELETED"

3. **Client1** executes XQuery
   
   ```xml
   let $deleted := //nodes return ( delete node $deleted, db:event("DELETED", $deleted) )
   ```

4. **Client2** and **Client3** will receive the deleted nodes.

**Included in Update Expression with Payload**

1. **Client1** creates an event with the name "DELETED"

2. **Client2** and **Client3** call the watch method for event "DELETED"

3. **Client1** executes XQuery
   
   ```xml
   let $deleted := //nodes return ( delete node $deleted, db:event("DELETED",
   <message>  <payload>{count($deleted)} items have been deleted.</payload>
   <items>{$deleted}</items> </message> )
   ```

4. **Client2** and **Client3** will receive the message with the payload and the deleted nodes.
Chapter 92. Execution Plan

Read this entry online in the BaseX Wiki.

This article is part of the Advanced User's Guide. For each execution of a query, BaseX creates an execution plan. This execution plan shows you each step of the query, so that you can evaluate your query and analyse if it accesses any indexes or not. You can activate the execution plan by activating the XMLPLAN or DOTPLAN options.

Examples

Execution plan for original and optimized query execution

Query: for $item in doc('xmark')/descendant::item where $item/payment = 'Creditcard' return $item

Optimized query: doc('xmark')/descendant::item[payment = 'Creditcard']

Execution plan:

Replacing XQuery with equivalent XPath expressions

Execution plan for query execution with full-text index access and without

Query: //MEDIUM/LAN[text() contains text "dt"]

Execution plan:
Query Plan 2
Chapter 93. Indexes

Read this entry online in the BaseX Wiki.

This article is part of the Advanced User's Guide and introduces the available index structures, which are utilized by the query optimizer to rewrite expressions and speed up query evaluation.

Nearly all examples in this article are based on the factbook.xml document. To see how a query is rewritten, please turn on the Info View in the GUI or use the -V flag on command line.

Structural Indexes

Structural indexes will always be present 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 statistics are discarded after database updates and can be recreated with the OPTIMIZE command.

The name index is e.g. applied to pre-evaluate location steps that will never yield results:

```xquery
/ non-existing-name
```

The contents of the name indexes can be directly accessed via 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 the same statistical information as the name index. The statistics are discarded after database updates and can be recreated with the OPTIMIZE command.

The path index is applied to rewrite descendant steps to multiple child steps. Child steps can be evaluated faster, as less nodes have to be accessed:

```xquery
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:

```xquery
/: will be rewritten and pre-evaluated by the path index :)
count( doc('factbook')//country )
```

The contents of the path index can be directly accessed via the XQuery function `index:facets()`.

Resource Index

The resource index contains references to the pre values of all XML document nodes. It speeds up the access to specific documents in a database, and it will be automatically updated when updates are performed.

The following query will be sped up by the resource index:
Indexes

Value Indexes

Value indexes can be optionally created and dropped by the user. The text and attribute index will be created by default.

Text Index

Exact Queries

This index speeds up string-based equality tests on text nodes. The UPDINDEX option can be activated to keep this index up-to-date.

The following queries will all be rewritten for index access:

```xml
(: 1st example :)  
//*[text() = 'Germany'],  
(: 2nd example :)  
doc('factbook.xml')//name[. = 'Germany'],  
(: 3rd st example :)  
for $c in db:open('factbook')//country  
where $c//@city/name = 'Hanoi'  
return $c//@name
```

Text nodes can be directly accessed from the index via the XQuery function `db:text()`. The contents of the index can be accessed via `index:texts()`.

Range Queries

The text index also supports range queries based on string comparisons:

```xml
(: 1st example :)  
db:open('Library')//Medium[Year >= '2005' and Year <= '2007'],  
(: 2nd example :)  
let $min := '2011-04-16T00:00:00'  
let $max := '2011-04-19T23:59:59'  
return db:open('news')//entry[date-time > $min and date-time < $max]
```

Text nodes can be directly accessed 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. The UPDINDEX option can be activated to keep this index up-to-date.

The following queries will all be rewritten for index access:

```xml
(: 1st example :)  
//country[@car_code = 'J'],  
(: 2nd example :)  
//province[@* = 'Hokkaido']//name,  
(: 3rd example :)  
//sea[@depth > '2100' and @depth < '4000']
```

Text nodes can be directly accessed from the index via the XQuery functions `db:attribute()` and `db:attribute-range()`. The contents of the index can be accessed via `index:attributes()`.
Indexes

Full-Text Index

The Full-Text index speeds up queries using the contains text expression. Internally, two index structures are provided: the default index sorts all keys alphabetically by their character length. It is particularly fast if fuzzy searches are performed. The second index is a compressed trie structure, which needs slightly more memory, but is specialized on wildcard searches. Both index structures will be merged in a future version of BaseX.

The following queries are examples for expressions that will be optimized for index access (provided that the relevant index exists in a particular database):

If the full-text index exists, the following queries will all be rewritten for index access:

1. `//country/name[text() contains text 'and']`
2. `//religions[. contains text { 'Catholic', 'Roman' } using case insensitive distance at most 2 words]`

Text nodes can be directly accessed from the index via the XQuery function db:fulltext(). The Full-Text Module contains additional functions for retrieving index data.

Index Construction

If main memory runs out while creating a value index, the currently generated index structures will be partially written to disk and eventually merged. If the used memory heuristics fails for some reason (i.e., because multiple index operations run at the same time), fixed index split sizes may be chosen via the INDEXSPLITSIZE and FTINDEXSPLITSIZE options.

If DEBUG is set to true, and if a new database is created from command-line, the number of index operations will be output to standard output; this might help you to choose proper split size. The following example shows how the output can look like for a document with 111 MB and 128 MB of available main memory:

```
> basex -d -c"set ftindex; create db 111mb 111mb.xml"
Creating Database...
.... 8132.44 ms (17824 KB)
Indexing Text...
.. 979920 operations, 2913.78 ms (44 MB)
Indexing Attribute Values...
.. 381870 operations, 630.61 ms (21257 KB)
Indexing Full-Text...
..|| 3 splits, 12089347 operations, 16420.47 ms (36 MB)
```

The info string 3 splits indicates that three partial full-text index structures were written to disk, and the string 12089347 operations tells that the index construction consisted of appr. 12 mio index operations. If we set FTINDEXSPLITSIZE to the fixed value 4000000 (12 mio divided by three), or a smaller value, we should be able to build the index and circumvent the memory heuristics.

Updates

By default, index structures are discarded after an update operation, and their maintenance is left to the user.

After the execution of update operations, the OPTIMIZE command or the db:optimize function can be called to rebuild the index structures. This way, multiple update operations will be performed faster, as the database meta data is only updated and regenerated once in the updating process.

As an alternative, incremental indexing can be turned on for the text and attribute value index: the UPDINDEX option must be activated before creating a new database. Note that, even with this option, the update of the path and fulltext index will have to be manually triggered.
Changelog

Version 7.2.1

• Added: string-based range queries
Chapter 94. 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>Operation</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></td>
<td></td>
<td></td>
<td></td>
<td>Query executed in 2.38 ms.</td>
</tr>
<tr>
<td></td>
<td>127.0.0.1:4722</td>
<td>jack</td>
<td>OK</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></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>Operation</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>Operation</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></td>
</tr>
</tbody>
</table>

An admin user has accessed the factbook database via REST.
Chapter 95. 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"
```

<table>
<thead>
<tr>
<th>PRE</th>
<th>DIS</th>
<th>SIZ</th>
<th>ATS</th>
<th>ID</th>
<th>NS</th>
<th>KIND</th>
<th>CONTENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>DOC</td>
<td>db.xml</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>ELEM</td>
<td>xml</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>TEXT</td>
<td>HiThere</td>
</tr>
</tbody>
</table>

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"
```

<table>
<thead>
<tr>
<th>PRE</th>
<th>DIS</th>
<th>SIZ</th>
<th>ATS</th>
<th>ID</th>
<th>NS</th>
<th>KIND</th>
<th>CONTENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>DOC</td>
<td>db.xml</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>ELEM</td>
<td>b</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>ELEM</td>
<td>xml</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>TEXT</td>
<td>HiThere</td>
</tr>
</tbody>
</table>

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.basex 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.basex) 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.basex, 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:

<table>
<thead>
<tr>
<th>fpre's</th>
<th>addr's</th>
</tr>
</thead>
<tbody>
<tr>
<td>0, 13, 257</td>
<td>0, 8192, 4096</td>
</tr>
</tbody>
</table>

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 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 96. 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><em>Num</em></td>
<td>Compressed integer (1-5 bytes), specified in Num.java</td>
<td>15 → 0F; 511 → 41 FF</td>
</tr>
<tr>
<td><em>Token</em></td>
<td>Length (Num) and bytes of UTF8 byte representation</td>
<td>Hello → 05 48 65 6c 6c 6f</td>
</tr>
<tr>
<td><em>Double</em></td>
<td>Number, stored as token</td>
<td>123 → 03 31 32 33</td>
</tr>
<tr>
<td><em>Boolean</em></td>
<td>Boolean (1 byte, 00 or 01)</td>
<td>true → 01</td>
</tr>
<tr>
<td><em>Nums, Tokens, Doubles</em></td>
<td>Arrays of values, introduced with the number of entries</td>
<td>1,2 → 02 01 31 01 32</td>
</tr>
<tr>
<td><em>TokenSet</em></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) 2. 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 → Tag Index • ATTS → Attribute Name Index • PATH → Path Index • NS → Namespaces • DOCS → Document Index 2. Empty key as finalizer</td>
<td>DiskData()</td>
</tr>
<tr>
<td>2 a) Name IndexTag/attribute names</td>
<td>1. Token set, storing all names (TokenSet) 2. 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>
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)

PathSummary() PathNode()

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)

Namespaces() NSNode()

2 d) Document Index

Array of integers, representing the distances between all document pre values (Nums)

DocIndex()

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; it uses the same layout.

Full-Text Fuzzy Index: ftxx, ftxy, ftxz

...may soon be reimplemented.
Chapter 97. 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 are some more informations about the transaction management.

Transaction

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.

Note: 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 the shutdown time. So we advise to use the BACKUP command to backup your database regularly. If the worst case occurs, you can try the INSPECT command to check if your database has obvious inconsistencies, and RESTORE to restore a previous version of the database.

Update Transactions

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 locking on database level. Writing transactions do not necessarily block all other transactions any more. The number of parallel transactions can be limited by setting the PARALLEL option.

Transaction Monitor

The transaction monitor ensures that just one writing transaction or an arbitrary amount of reading transactions per database are active at the same time.

Deadlocks are prevented by using preclaiming two phase locking. Execution is starvation-free as lock acquisition is queued per database. Due to the specifics of XQuery Update, all updates are written at the end of the query. Locking is strict with the exception that databases for which BaseX recognizes it will not write to are downgraded to read locks.

Locks are not synchronized between multiple BaseX instances. We generally recommend working with the client/server architecture if concurrent write operations are to be performed.

External Side Effects

Access to external resources (files on hard disk, HTTP requests, ...) is not controlled by BaseX' transaction monitor 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`. It accepts arbitrary strings, for locking. The option may be set multiple times to acquire multiple locks, or use a colon as separator between them. These locks are in another namespace than the database names, locking "factbook" using the options does not lock the database named factbook.

These options will put read locks on `foo`, `bar` and `batz` and a write lock on `quix`. 
declare option query:read-lock "foo,bar";
declare option query:read-lock "batz";
declare option query:write-lock "quix";

Limitations

Commands

Database locking works with all commands unless no glob syntax is used, such as in the following command call:

- DROP DB new*: drop all databases starting with "new"

XQuery

As XQuery is a very powerful language, deciding which databases will be accessed by a query is non-trivial. Optimization is work in progress. The current identification of which databases to lock is limited to queries that access the currently opened database, XQuery functions that explicitly specify a database, and expressions that address no database at all.

Some examples on database-locking enabled queries, all of these can be executed in parallel:

- //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"

Some examples on queries that are not supported by database-locking yet:

- let $db := 'factbook' return doc($db), will read-lock: referencing database names isn’t supported yet
- for $db in ('factbook') return doc($db), will read-lock globally
- doc(doc('test')/reference/text()), will read-lock globally
- let $db := 'test' return insert nodes <test/> into doc($db), will write-lock globally

A list of all locked databases is output if QUERYINFO is set to true. If you think that too much is locked, please give us a note on our mailing list with some example code.

GUI

Database locking is currently disabled if the BaseX GUI is used.

How to disable

In order to disable traditional process locking, the option GLOBALLOCK can be set to false. This can e.g. be done by editing your .basex file (see Options for more details). To enable it again, set it to true.

File-System Locks

Update Operations

During the term of 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 may not be deleted. In this
case, the database cannot be opened anymore using the default commands, 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 caused by write operations running in 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, it will be rejected with the message "Database ... is opened by another process." if no exclusive lock can be acquired.

As the standalone versions of BaseX (command-line, GUI) cannot be synchronized with other BaseX instances, we generally recommend working with the client/server architecture if concurrent write operations are to be performed.

Changelog

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 98. 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 specific database command.

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. Next, local permissions exist, which can be assigned to single databases. Local permission have a higher priority and override global permissions.

All global permissions are stored in the file .basexperm, and local permissions are encoded in the database metadata (inf.base).

Permissions hierarchy User names must follow the valid names constraints.

Commands

Admin permissions are needed to execute all of the following commands:

Creating user 'test' (password will be entered on command line):

> CREATE USER test

Change user 'test' password (password will be entered on command line):

> ALTER USER test

As global permissions, you can set 'none', 'read', 'write', 'create' and 'admin':

Grant all permissions to user 'test':

> GRANT admin TO test

Valid local permissions are 'none', 'read' and 'write':

Granting write permission on database 'factbook' to user 'test':

> GRANT write ON factbook 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) the 'factbook' database. If no local permissions are set, the global rights are inherited.

Showing global permissions:
> SHOW USERS

Showing local permissions on database 'factbook':

> SHOW USERS ON factbook

Dropping of user 'test':

> DROP USER test
Part XI. Use Cases
Chapter 99. 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>Limits</td>
<td>512 GiB</td>
<td>536'870'912</td>
<td>2'147'483'648</td>
<td>32'768'128</td>
<td>32'768'128</td>
<td>16'384'256</td>
<td>no limit</td>
<td>6</td>
<td>128</td>
</tr>
<tr>
<td>RuWikiHist</td>
<td>421 GiB</td>
<td>416 GiB</td>
<td>324'848'508</td>
<td>21</td>
<td>6</td>
<td>2</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZhWikiHist</td>
<td>26 GiB</td>
<td>120 GiB</td>
<td>179'199'662</td>
<td>21</td>
<td>6</td>
<td>2</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EnWiktont</td>
<td>22 GiB</td>
<td>75 GiB</td>
<td>134'380'393</td>
<td>21</td>
<td>6</td>
<td>2</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>XMark</td>
<td>55 GiB</td>
<td>64 GiB</td>
<td>1'615'071'348</td>
<td>74</td>
<td>9</td>
<td>0</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EnWikiMedia</td>
<td>64 GiB</td>
<td>52 GiB</td>
<td>401'456'348</td>
<td>21</td>
<td>6</td>
<td>2</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MedLine</td>
<td>38 GiB</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>
<td></td>
</tr>
<tr>
<td>iProClass</td>
<td>36 GiB</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>
<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>1'104'623'376</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>25 GiB</td>
<td>198'546'747</td>
<td>24</td>
<td>21</td>
<td>2</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>XMark</td>
<td>22 GiB</td>
<td>26 GiB</td>
<td>645'997'962</td>
<td>74</td>
<td>9</td>
<td>0</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>InterPro</td>
<td>14 GiB</td>
<td>19 GiB</td>
<td>860'304'235</td>
<td>7</td>
<td>15</td>
<td>0</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genome1</td>
<td>13 GiB</td>
<td>13 GiB</td>
<td>432'628'1032</td>
<td>26</td>
<td>101</td>
<td>2</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NewYorkTime</td>
<td>18 GiB</td>
<td>13 GiB</td>
<td>280'407'005</td>
<td>41</td>
<td>33</td>
<td>0</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TrEMBL</td>
<td>11 GiB</td>
<td>14 GiB</td>
<td>589'650'538</td>
<td>47</td>
<td>30</td>
<td>2</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>XMark</td>
<td>11 GiB</td>
<td>13 GiB</td>
<td>323'083'402</td>
<td>74</td>
<td>9</td>
<td>0</td>
<td>13</td>
<td></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>61</td>
<td>283</td>
<td>1</td>
<td>93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDMX</td>
<td>6356 MiB</td>
<td>8028 MiB</td>
<td>395'871'872</td>
<td>22</td>
<td>6</td>
<td>3</td>
<td>7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

354
<table>
<thead>
<tr>
<th>Dataset</th>
<th>Size</th>
<th>Unique Size</th>
<th>Count</th>
<th>Size Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>OpenStreet</td>
<td>5171 MiB</td>
<td>6'910'669</td>
<td>19</td>
<td>6</td>
</tr>
<tr>
<td>SwissProt</td>
<td>5422 MiB</td>
<td>241'274'408</td>
<td>70</td>
<td>3</td>
</tr>
<tr>
<td>EURLex</td>
<td>5532 MiB</td>
<td>167'328'032</td>
<td>186</td>
<td>1</td>
</tr>
<tr>
<td>Wikicorpus</td>
<td>4432 MiB</td>
<td>157'948'561</td>
<td>1'257</td>
<td>2</td>
</tr>
<tr>
<td>EnWikiRDF</td>
<td>3537 MiB</td>
<td>98'433'1941</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>CoPhIR</td>
<td>918 MiB</td>
<td>46'401'941</td>
<td>23</td>
<td>3</td>
</tr>
<tr>
<td>MeSH</td>
<td>2410 MiB</td>
<td>101'638'857</td>
<td>42</td>
<td>42</td>
</tr>
<tr>
<td>FreeDB</td>
<td>2462 MiB</td>
<td>102'901'512</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>XMark</td>
<td>1303 MiB</td>
<td>32'298'982</td>
<td>74</td>
<td>9</td>
</tr>
<tr>
<td>DeepFS</td>
<td>850 MiB</td>
<td>44'821'506</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>LibraryUK</td>
<td>918 MiB</td>
<td>46'401'941</td>
<td>23</td>
<td>3</td>
</tr>
<tr>
<td>Twitter</td>
<td>767 MiB</td>
<td>15'309'015</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Organizations</td>
<td>724 MiB</td>
<td>33'112'392</td>
<td>38</td>
<td>9</td>
</tr>
<tr>
<td>DBLP</td>
<td>944 MiB</td>
<td>36'878'181</td>
<td>4</td>
<td>35</td>
</tr>
<tr>
<td>Feeds</td>
<td>604 MiB</td>
<td>5'933'713</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>MedLineS</td>
<td>407 MiB</td>
<td>14'512'851</td>
<td>1</td>
<td>111</td>
</tr>
<tr>
<td>AirBase</td>
<td>273 MiB</td>
<td>14'512'851</td>
<td>111</td>
<td>5</td>
</tr>
<tr>
<td>MedLineD</td>
<td>195 MiB</td>
<td>10'401'847</td>
<td>66</td>
<td>8</td>
</tr>
<tr>
<td>ZDNET</td>
<td>133 MiB</td>
<td>3'060'186</td>
<td>40</td>
<td>90</td>
</tr>
<tr>
<td>JMNEdict</td>
<td>171 MiB</td>
<td>8'592'666</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>XMark</td>
<td>130 MiB</td>
<td>3'221'926</td>
<td>74</td>
<td>9</td>
</tr>
<tr>
<td>Freshmeat</td>
<td>86 MiB</td>
<td>3'832'028</td>
<td>58</td>
<td>1</td>
</tr>
<tr>
<td>DeepFS</td>
<td>93 MiB</td>
<td>4'842'638</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Treebank</td>
<td>92 MiB</td>
<td>3'829'513</td>
<td>250</td>
<td>1</td>
</tr>
<tr>
<td>DBLP2</td>
<td>102 MiB</td>
<td>4'044'649</td>
<td>35</td>
<td>6</td>
</tr>
<tr>
<td>DDI</td>
<td>39 MiB</td>
<td>2'070'157</td>
<td>104</td>
<td>16</td>
</tr>
<tr>
<td>Alfred</td>
<td>86 MiB</td>
<td>3'784'285</td>
<td>60</td>
<td>0</td>
</tr>
<tr>
<td>University</td>
<td>66 MiB</td>
<td>3'468'606</td>
<td>28</td>
<td>4</td>
</tr>
<tr>
<td>MediaUK</td>
<td>45 MiB</td>
<td>1'619'443</td>
<td>21</td>
<td>3</td>
</tr>
<tr>
<td>HClIBIB2</td>
<td>33 MiB</td>
<td>617'023</td>
<td>39</td>
<td>1</td>
</tr>
<tr>
<td>Nasa</td>
<td>25 MiB</td>
<td>845'805</td>
<td>61</td>
<td>8</td>
</tr>
<tr>
<td>MovieDB</td>
<td>19 MiB</td>
<td>868'980</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>KanjiDic2</td>
<td>18 MiB</td>
<td>917'833</td>
<td>27</td>
<td>10</td>
</tr>
<tr>
<td>XMark</td>
<td>13 MiB</td>
<td>324'274</td>
<td>74</td>
<td>9</td>
</tr>
<tr>
<td>Shakespeare</td>
<td>9854 KiB</td>
<td>327'170</td>
<td>59</td>
<td>0</td>
</tr>
<tr>
<td>TreeOfLife</td>
<td>7106 KiB</td>
<td>363'560</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Thesaurus</td>
<td>4088 KiB</td>
<td>201'798</td>
<td>33</td>
<td>9</td>
</tr>
<tr>
<td>MusicXML</td>
<td>2942 KiB</td>
<td>171'400</td>
<td>179</td>
<td>56</td>
</tr>
<tr>
<td>BibDBPub</td>
<td>2359 KiB</td>
<td>80'178</td>
<td>54</td>
<td>1</td>
</tr>
<tr>
<td>Factbook</td>
<td>1560 KiB</td>
<td>77'315</td>
<td>23</td>
<td>32</td>
</tr>
<tr>
<td>XMark</td>
<td>1334 KiB</td>
<td>33'056</td>
<td>74</td>
<td>9</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>EURLEllex</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/iproclass/iproclass.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://dump.wiki.openstreetmap.org/osmwiki-latest-files.tar.gz">http://dump.wiki.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.drze.de/BELIT/thesaurus">http://www.drze.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/">http://www.xml-benchmark.org/</a> generated 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 100. Twitter

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>
</tbody>
</table>
### Twitter Statistics

<table>
<thead>
<tr>
<th>Date</th>
<th>Total Tweets</th>
<th>Average tweets per hour</th>
<th>Average tweets per minute</th>
<th>Average tweets per second</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sun, 6-May-2012</td>
<td>35,982,976</td>
<td>1,499,290</td>
<td>24,988</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,
    "retweet_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": "CODEEDE",
        "name": "BaseX",
        "profile_background_color": "CODEEDE",
        "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_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/>
    </entities>
    <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/>
    <retweet_count>0</retweet_count>
    <created_at>Fri May 04 13:17:16 +0000 2012</created_at>
    <coordinates/>
</json>
```
BaseX Performance

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>