

Comparison of fast, multiplatform XML libraries

1. The eXpat XML Parser (<http://expat.sourceforge.net/>)
2. The XML C Library for Gnome (<http://www.xmlsoft.org/>)
3. Apache Xerces for C++ parser (<http://xml.apache.org/xerces-c/index.html>)
4. Language Technology Group RXP XML parser (<http://www.ltg.ed.ac.uk/software/xml/index.html>)
5. Trolltech QT XML Module (<http://doc.trolltech.com/3.1/xml.html>)
6. IBM alphaWorks XML4C (<http://www.alphaworks.ibm.com/tech/xml4c>)
7. Oracle 9i XML Developers Kit for C/C++/Java (<http://otn.oracle.com/tech/xml/xdk/>)
8. Java XML pack (<http://java.sun.com/xml/downloads/javaxmlpack.html>)

	<i>Expat</i>	<i>libXML</i>	<i>Xerces C</i>	<i>RXP</i>	<i>QT</i>	<i>XML4C</i>	<i>XDK 9i</i>	<i>J2EE</i>
Author	SourceForge	Gnome	Apache	LT Group	Trolltech	IBM	Oracle	Sun
License	MIT	MIT	Apache	GPL	Comm.	Apache	OTN	Apache
Supported OS	Win Linux QNX LynxOS ? BSD Solaris MacOS X HP-UX IRIX AIX OS/2 OpenUnix OpenVMS	Win Linux QNX BSD Solaris MacOS X OpenVMS MVS	Win NT Linux BSD Solaris MacOS X HP-UX IRIX AIX OS/2 UnixWare PTX AS/400 OS/390	Win Linux MacOS X	Win Linux QNX LynxOS BSD Solaris MacOS X HP-UX IRIX AIX OS/2 UnixWare Tru64 XFree86	Win NT Linux BSD Solaris MacOS X HP-UX IRIX AIX OS/2 UnixWare PTX AS/400 OS/390	Win NT Linux Solaris HP-UX	Java
Language	C/C++	C/C++	C++ Java	C/C++	C++	C++ Java	C/C++	Java
Bindings	Python PHP Perl Tcl/Tk Ruby Delphi	Python PHP Perl Tcl/Tk Ruby Delphi Kylux						
Used By	Mozilla	Gnome	Apache			Apache		Apache

	<i>Expat</i>	<i>libXML</i>	<i>Xerces C</i>	<i>RXP</i>	<i>QT</i>	<i>XML4C</i>	<i>XDK 9i</i>	<i>J2EE</i>
Additional Supported Standards		XML Catalog						
Network Clients		FTP, HTTP	FTP, HTTP			FTP, HTTP		FTP, HTTP
SOAP 1.1	EasySOAP ¹⁰						+ ?	Apache SOAP
SOAP 1.2								
SOAP Attachment								Apache SOAP
WSDL	Planned by EasySOAP ¹⁰							
UDDI								+
ebXML								+
XML GUI					+			
XSD -> Class Generator							+	

1. The XSLT C Library for Gnome (<http://xmlsoft.org/XSLT/>) is distributed under MIT license and claims to support all systems, supported by XML C Library for Gnome.
2. Gnome DOM Engine (<http://gdome2.cs.unibo.it/>) is distributed under LGPL license and claims to support all systems, supported by XML C Library for Gnome.
3. XML Security Library (<http://www.aleksey.com/xmlsec/>) is distributed under MIT license and requires OpenSSL 0.9.6c or 0.9.7 (preferred).
4. Sablotron (http://www.gingerall.com/charlie/ga/xml/p_sab.xml) is distributed under GPL and Mozilla Open License 1.1 licenses. Sablotron is written in C++ and depends on expat XML parser.
5. Arabica (<http://www.jezuk.co.uk/cgi-bin/view/arabica>) is distributed under BSD license, written in C++ and claims to support Linux, Windows and different Unixes equipped with GNU compiler 3.
6. Xalan for C++ (<http://xml.apache.org/xalan-c/index.html>) is distributed under Apache license and known to work on Windows, Linux, Solaris, HP-UX and AIX platforms.
7. Pathan (<http://software.decisionsoft.com/pathanIntro.html>) is distributed under BSD style license, written in C++ and known to support Windows, Linux and Solaris platforms.
8. Center Point XML (<http://www.cpointc.com/XML/>) is distributed under Netscape Public License and claims to support many platforms, including Windows, Linux, BSD, Solaris, HP-UX, MacOS X. Written in C++.
9. IBM alphaWorks LotusXSL (<http://www.alphaworks.ibm.com/tech/lotusxsl>) is IBM shipped version of xalan library.
10. EasySOAP++ (<http://easysoap.sourceforge.net/>) is distributed under LGPL license, supports Linux and Windows platforms. And claims to be compatible with different SOAP implementations including, but not limited, Microsoft SOAP toolkit, Microsoft .NET platform, Apache, White Mesa, Phalanx.
11. iconv (<http://www.gnu.org/software/libiconv/>) is part of GNU c library and distributed under

GPL license.

12. IBM International Components for Unicode (<http://oss.software.ibm.com/icu/>) is written in pure C unicode 3.2 library, provided under X Open Source license. ICU claims to support Linux, Windows and variety of Unix platforms.

13. Apache XML Security for C++ (<http://xml.apache.org/security/index.html>) is distributed under Apache license and known to work on Windows, Linux, Solaris, HP-UX and AIX platforms.

	<i>libxslt</i>	<i>GDome</i>	<i>xmlsec</i>	<i>sablotron</i>	<i>cslxml</i>	<i>arabica</i>	<i>xalan</i>	<i>XML Security</i>	<i>pathan</i>
Author	Gnome	Gnome	Alexey Sanin	Ginger Alliance	Connective Software Engineering	Jez UK	Apache	Apache	Decision Soft
License	MIT	LGPL	MIT	GPL Mozilla	Netscape	BSD	Apache	Apache	BSD
Requirements	LibXML2	LibXML2	LibXML2 OpenSSL	Expat	Expat	Expat LibXML Xerces	Xerces	Xerces Xalan	Xerces
Language	C/C++	C/C++	C/C++	C++	C++	C++	C++ Java	C++ Java	C++ Java
Bindings	Perl Ruby Python Tcl/Tk Pascal	Perl Ruby Caml		Perl Ruby Python Php Tcl/Tk Pascal Ada Erlang					Perl
Used By	Gnome	Gnome	Gnome	mozilla			Apache	Apache	
Library	Shared	Shared	Shared	Shared	Static	Static	Shared	Shared	shared
Library Size	255 KB	225 KB	317 KB	805 KB	962 KB	76 KB	3.3 MB	345 KB	991 KB
Thread Safe	Yes	?	?	Partly ¹	Yes	?	Yes	Yes	Yes
Latest Version	1.1.2	0.8.1	1.2.4	1.0.1	2.1.7	Jan04	1.7	1.0.0	1.2r2 2.0alpha
Release Date	24.12.03	10.05.03	27.01.04	25.11.03	15.05.03	03.01.04	18.01.04	10.08.03	04.09.03

1. Expect DOM Modification part

Benchmark Setup

1. Hardware

1. Intel Pentium IV 2200 (22x100) MHZ (Nothwood)
2. Intel D850MV motherboard (i850E + ICH2)

3. 4 x Samsung MR16R 256 MB PC800 Rambus DRAM
4. IBM GXP120 60GB hard drive

2. Software

1. DarkLin III (Mandrake 9.2 Based) Single User Mode
2. Kernel 2.4.22
3. GNU C Library 2.3.2
4. GNU C Compiler 3.3.1
5. Intel C++ Compiler 8.0
6. J2SE 1.4.2 + JAXP 1.2

3. Libraries

1. Expat 1.95.6
2. Sablotron 1.0.1
3. Arabica Jan04
4. CenterPoint XML 2.1.7
5. Gnome XML Library 2.6.5
6. Gnome XML DOM Library 0.8.1
7. Gnome XSLT Library 1.1.2
8. XML Security Library 1.2.4
9. Xerces for C++ 2.4.0
10. Xalan for C++ 1.7
11. LTG RXP Parser 1.2.8
12. Trolltech QT 3.3.0
13. Oracle XML Developers Kit for C 9.2.0.6.0
14. Oracle XML Developers Kit for Java 9.2.0.6.0
15. Apache XML Security for C++ 08.02.2004
16. Apache XML Security for Java 1.0.5D2

4. XML Files

1. **XMLGEN** - autogenerated simple XML file filled with random values, have 3 levels of deep and very simple XSD schema. Examples and schema can be accessed in XML Benchmark source package.
2. **OPCGEN** – autogenerated sequence of OPC(<http://www.opcfoundation.org>) messages filled with random data. This messages have 1-5 levels of deep and complex xsd schema. There are 16 different type messages which dramatically differs from each other. Examples and schema can be accessed in XML Benchmark source package.
3. 12 MB Resource Description File (RDF) from Dmoz.org
4. 1 MB Library Catalog file

Parsing Benchmarks

	Executable Size	OPCGEN 4KB	XMLGEN 256KB	Dmoz RDF
Expat	28	9857	9,06	294,2
Expat + Arabica (DOM)	481	347561	453,1	7479,43
Expat + Arabica (SAX)	149	135863	201,1	3019
Expat + CSLXML (DOM)	373	137798	92,06	2456
Expat + CSLXML (SAX)	245	51356	28,4	1022
Expat + Sablotron	28	96871	96,28	2035
LibXML	27	43384	42,5	1645
LibXML + Gdome	27	45328	42,92	1734
LibXML (SAX)	27	12198	9,8	433
LibXML (no threads)	27	21069	25,16	1127
LibXML Push	28	44326	44,98	1654
Xerces (DOM)	41	83923	67,39	1750
Xerces (SAX)	55	32479	21,41	969
RXP Parser	257	22998	38,81	1060
QT XML Module (DOM)	33	208300	213,35	3810
QT XML Module (SAX)	46	218078	238,62	4390
Oracle XDK for C (DOM)	1341	80640	13,36	1487
Oracle XDK for C (SAX)	1341	78479	10,35	1475
Oracle XDK for C++ (DOM)	1304	84064	12,83	1541
Oracle XDK for C++ (SAX)	1304	75137	10,72	1473
Oracle XDK for Java (DOM)	1	68343	79,35	1630
Oracle XDK for Java (SAX)	1	40393	22,5	922,76
Sun JAXP (DOM)	1	92435	61,48	1367,44
Sun JAXP (SAX)	2	44400	22	787

Validation Benchmark

	Executable Size	OPCGEN 4KB	XMLGEN 4KB	XMLGEN 256KB
LibXML	30	513	0,2	15,14
LibXML (no threads)	30	327	0,15	11,02
Xerces	50	22777	2,95	57,7
Oracle XDK for C	1547	2237	0,44	15,18
Sun JAXP	3	34451	3,1	67,3

XSLT Transformation Benchmark

	Executable Size	XMLGEN 4KB	XMLGEN 256KB	XMLGEN 4096KB
Sablotron	30	14,6	613	10714
LibXSLT	28	2,53	108	1690
LibXSLT (no threads)	28	1,63	60,45	942
Xalan	53	4,79	202,3	4276
Oracle XDK for C	1342	3,31	113,1	6044
Sun JAXP	3	7,01	304,3	Failed

DOM Benchmark

	Executable Size	OPCMsg 5KB	OPCMsg 500KB	OPCMsg 9MB
Expat + CSLXML	256	2,86	366,58	Failed
Expat + Sablotron	20	2,87	293,6	9233
LibXML	18	0,76	95,17	1950,63
LibXML + LibGDome	19	2,99	348,38	7573
LibXML (no threads)	18	0,47	65,43	1338
Xerces	36	1,6	152,3	3019
QT XML Module	38	2,72	286,33	5815
Oracle XDK for C	1337	0,44	28,71	585,59
Oracle XDK for C++	1299	0,45	29,68	598,86
Oracle XDK for Java	4	1,26	256,75	Failed
Sun JAXP	5	0,66	173,34	Failed

XML Security Benchmark

	Executable Size	OPCGEN 4KB	XMLGEN 256KB	Dmoz RDF
LibXML + XMLSec (Signature)	35	20109	172,8	3163
LibXML + XMLSec (Encryption)	35	27166	184,3	7162
LibXML (no threads) + XMLSec (Signature)	35	17165	134,9	2511
LibXML (no threads) + XMLSec (Encryption)	35	24662	166,3	6928
Xerces + Apache XML Security (Signature)	35	172885	3890	Failed
Xerces + Apache XML Security (Encryption)	35	59580	747,3	Failed
JAXP + Apache XML Security (Signature)	4	1314979	Failed	Failed

Stress Test

Parsers able process 64MB file: expat, arabica (SAX), csllxml (DOM,SAX), libxml (DOM,SAX), oracle (DOM,SAX), QT (SAX), RXP, Xerces (DOM,SAX), xml4c (DOM,SAX)

Parsers able process 128MB file: expat, arabica (SAX), csllxml (DOM,SAX), libxml (SAX), oracle (SAX), QT (SAX), Xerces (SAX), xml4c (SAX)

Parsers able process 256MB file: expat, arabica (SAX), csllxml (SAX), libxml (SAX), oracle (SAX), Xerces (SAX), xml4c (SAX)

Parsers able process 512MB file: Xerces (SAX), xml4c (SAX)

Standards Violations Notes

- Oracle bugs in schema verification and can drop core while validating non-valid documents.
- Gnome XML 2.6.5 library have problems with validation of xsd:QName attributes
- Xerces for C 2.4 don't catch schema implementation (xsi) type violations