XML Processing on z/OS

Bill Carey
System z Strategy and Technology
Poughkeepsie, NY
wgcarey@us.ibm.com
Topics

- Z/OS XML Parsing technologies with focus on XML System Services (z/OS XML)
  - Exploitation by higher level support
  - Performance
  - zAAP & zIIP support
- Middleware/products processing XML
- DataPower positioning
Primary IBM XML parsing building blocks on z/OS

XML Support in Java  JDK
- XML4J and XSLT4J support (based on Apache Xerces and Xalan)
- High performance optimizations for
  - Parsing (XLXP-J)
  - Transformation / query processing (XL TXE-J)
- XLXP-J implements SAX and StAX
- Exploited by a variety of IBM, customer, and ISV apps (e.g. WebSphere)
- Additional support in WebSphere XML Feature Pack
- All Java can be directed to zAAP processors

XML Toolkit for z/OS
- C++ language bindings and common SAX/DOM interfaces
- Based on IBM Toronto XML4C/XSLT4C (aka Apache Xerces/Xalan) technologies
- Also exploited by a variety of IBM, customer & ISV apps

Enterprise COBOL and PL/I Compilers
- High performance, built-in XML parsing and creation/generation support

XML System Services (z/OS XML)
- A z/OS system component introduced in V1.8 with rollback to V1.7
- High performance, XML parsing support usable in all system environments (e.g. Task or SRB mode, cross-memory)
- Can process arbitrarily large documents
- More basic functionality than Apache parsers
- Strong underlying parsing technology usable directly or indirectly via other XML processing support

Function  Performance
What is z/OS XML System Services?

- An XML parser that is an integrated component of the z/OS base
  - Supports unique z/OS environments where minimum overhead is key
    - SRB and TCB modes
    - Cross-memory mode
    - No Language Environment® dependencies
  - Non-validating and validating (with V1.10) parser with well-formedness checking
  - No XML generation or XPath or XSLT processing capability
  - Assembler interface (V1.8), C/C++ interface (V1.9)

- Simple call model that avoids event-driven interface overhead

- Ability to handle very large documents ( >1 GB)

- XML documents parsed to a form readily usable by the invoking app

- Parsing operations are redirected to zAAP/zIIP specialty engines, where available

- Intended for z/OS system environments, middleware, and applications that need to handle XML very efficiently
What are XML System Services?

- An XML parser that is an integrated component of the z/OS base
  - High performance (minimal pathlength)
  - Supports unique z/OS environments where minimum overhead is key
    - SRB mode
    - Cross-memory mode
    - No LE dependencies
  - Well-formedness checking (in z/OS 1.7+), validation (in z/OS V1.9+)
  - No XML generation or XPath or XSLT processing capability currently
  - Output encoding = input encoding

- Simple call model that avoids event driven interface overhead

- Ability to handle very large documents (multi-Gigabytes)

- XML documents parsed to a self-defining binary form

- Intended for z/OS system environments, middleware, and applications that need to handle XML very efficiently

- Offloadable to zAAPs (task mode) and zIIPs (enclave SRB mode)

- A core XML technology base which, when optimized (by SW or HW acceleration), can benefit all exploiters

XML System Services = z/OS XML
Z/OS XML Design Overview

XML document

system service (exit) vector

@ get storage
@ free storage
@ string ID

Caller

Parser

Native parser interface

Core parser services

parsed data stream

<person>
  <name>
    John Doe
  </name>
</person>

START_ELEM person
START_ELEM name
CHAR_DATA John Doe
END_ELEM
END_ELEM

@ get storage
@ free storage
@ string ID

call to exits as needed

call to exits as needed
Z/OS XML Validating Parsing

XML DOC(S)

XML xsd Schemas

Parse

OSR Generator

Binary OSR

z/OS XML API

Non-Validating Parser

Validating Parser

Parsed Data Stream

OSR = Optimized Schema Representation
Z/OS XML zAAP/zIIP Enablement and XML Toolkit Exploitation

### Z/OS XML

- **V1.9:**
  - zAAP enablement of non-validating parsing capability for direct exploiters (Task Mode)
  - Rollback to V1.7, 1.8

- **V1.10:**
  - 100% zIIP enablement (Enclave SRB mode)
    - Rollback to V1.8, 1.9
  - Addition of validation capability (with zAAP/zIIP re-direction)
    - Rollback to V1.9

### XML Toolkit

- R9 SPE: Adds hooks to allow XML4C (SAX2 and DOM, non-validating) to utilize z/OS XML as underlying scanner where feature/option match exists (May 2008)

- R10: Adds hooks to allow XML4C (SAX2 and DOM, validating) to utilize z/OS XML validation support where feature/option match exists (Dec 2008)

  - **Support provided by z-specific classes that closely mimic existing XML4C cross-platform classes**
  - **Restrictions exist on the functions supported**
XML Toolkit vs. z/OS XML Non-validation Performance Comparison

Smaller is better

Toolkit 31-bit SAX Non-Validation vs Toolkit<->XMLSS R12 vs XMLSS R12 Non-Validation

All docs

IBM Laboratory results, a controlled measurement environment, your results may vary.
XML Toolkit vs. z/OS XML Validation Performance Comparison

All docs

Toolkit 31-bit SAX Validation vs Toolkit<->XMLSS R12 vs XMLSS R12

Relative Cycles Per ByteParsed

<table>
<thead>
<tr>
<th></th>
<th>Toolkit MIPS</th>
<th>XMLSS Offloadable MIPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>XM4C5.7</td>
<td>1.00</td>
<td>0.79</td>
</tr>
<tr>
<td>XML4C5.7 Toolkit w/namespace changes &lt;- &gt;XMLSS R12</td>
<td>0.60</td>
<td>0.486</td>
</tr>
<tr>
<td>XML4C5.7 Toolkit &lt;-&gt;XMLSS R12 original Goal</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>XMLSS R12</td>
<td>0.486</td>
<td></td>
</tr>
</tbody>
</table>

IBM Laboratory results, a controlled measurement environment, your results may vary.

Smaller is better
Enterprise COBOL V4 Compiler XML enhancements on z/OS

- Includes optional use of z/OS XML System Services as base underlying technology for significant new XML-related enhancements, e.g.
  - Namespace support
  - UTF-8 code page support
  - Improved well formedness checking
  - Multiple buffer support
- Gains zAAP enablement for XML parsing via z/OS XML exploitation
- Via compiler option (since 100% compatibility with existing parsing support not feasible)
  - XMLPARSE (COMPAT | XMLSS)
- COBOL V4.1 also provided additional XML GENERATE statement and general COBOL performance and functionality improvements as well
- As of COBOL V4.2 (Aug 2009), now also provides XML schema validation via use of z/OS XML

Note: COBOL apps may also directly use z/OS XML Services
IBM Enterprise PL/I XML enhancements

- Enterprise PL/I V3.8 (5655-H31) provided a new PLISAXC built-in function which can allow PL/I programs to use z/OS XML System Services for XML processing
  - Uses z/OS XML System Services as the base underlying technology for significant new XML-related enhancements:
    - Name space support
    - Large document (greater than 2GB) support
    - UTF-8 support, etc.
  - Gains zAAP enablement for XML parsing via z/OS XML System Services usage

- As of Enterprise PL/I V4.1 (Sept 2010), now provides validating parser support via use of z/OS XML
  - Via new PLISAXD built-in function

Note: PL/I apps may also directly use z/OS XML Services
IBM CICS TS V4 exploitation of z/OS XML

- The performance of XML parsing in CICS is improved with the exploitation of IBM z/OS XML System Services parser, which is used in portions of the processing of inbound SOAP messages in CICS.

- Faster XML processing:
  - CICS TS V4.1 showed a reduction in the CPU time used to parse XML messages, due to CICS now utilizing the z/OS XML parser.

- zAAP eligibility.
  - The portion of XML processing of inbound SOAP messages in CICS TS V4 performed by z/OS XML System Services is eligible for zAAP
z/OS XML Support and Potential for Exploitation

APPLICATION

DB2, CICS, RDz, .... (potential IBM and ISV exploiters)

COBOL/PLI XML Parsers

XML Toolkit –XML4C
SAX / DOM

Validating parser

XercesJ/XML4J Parser (zAAP offloaded since it is Java)

Non-Validating Parser

XML System Services

zAAP/zIIP-enabled

Other XML Processors
## Roll-out of zAAP and zIIP support in z/OS XML from a DB2 9 z/OS PureXML Support perspective

<table>
<thead>
<tr>
<th>Timeframe</th>
<th>Pre-z/OS R9</th>
<th>SEPT 2007 – z/OS R9</th>
<th>SEPT 2008 – z/OS R10</th>
<th>SEPT 2010 – z/OS R12</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Task (TCB) Mode</strong></td>
<td><img src="image1.png" alt="Diagram" /></td>
<td><img src="image2.png" alt="Diagram" /></td>
<td><img src="image3.png" alt="Diagram" /></td>
<td><img src="image4.png" alt="Diagram" /></td>
</tr>
<tr>
<td><strong>DRDA - Enclave SRB Mode</strong></td>
<td><img src="image5.png" alt="Diagram" /></td>
<td><img src="image6.png" alt="Diagram" /></td>
<td><img src="image7.png" alt="Diagram" /></td>
<td><img src="image8.png" alt="Diagram" /></td>
</tr>
<tr>
<td>XML System Services</td>
<td>XML System Services</td>
<td>XML System Services</td>
<td>XML System Services</td>
<td>XML System Services</td>
</tr>
<tr>
<td>Runs on standard CP</td>
<td>Runs on standard CP</td>
<td>100% zAAP-able</td>
<td>100% zAAP-able</td>
<td>Runs on standard CP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100% zAAP-able (no change)</td>
<td>100% zAAP-able (no change)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>zAAP-able at DB2 dialed in %</td>
<td>zAAP-able at DB2 dialed in %</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(no change)</td>
<td>(no change)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sub-system, Middleware, Development Tooling XML Support

- **General suggestion:** Utilize higher level XML processing services when advantages can be realized
  - Improved productivity
  - Faster time to market
  - Potentially better performance (not always)
  - Concentrate on addressing your business problem - not doing plumbing

- **Significant XML processing support in many areas:**
  - DB2 9: PureXML
  - WebSphere, WebSphere ESB, WebSphere Process Server
  - Rational Developer for z (RDz)
  - WebSphere Message Broker
  - CICS: XML, Web Services enablement
  - IMS: XML, Web Services enablement for IMS; IMS DB XML support
  - WebSphere Transformation Extender (WTX)
  - WebSphere DataPower appliances
  - Vendor products, e.g. from HostBridge, SOA Software, DataDirect, and others
  - Etc etc etc
IBM WebSphere DataPower SOA Appliance

z/OS XML and DataPower®: Complementary Capabilities

- XML processing on System z:
  - Provide fundamental platform XML processing support for use by z/OS, IBM/ISV middleware and products, as well as customer applications.
  - Focus: On-platform XML performance and price/performance

- DataPower:
  - A network proxy SOA / Web Services appliance, optionally supplements on-platform solutions based on functionality, performance, cost, ...
    - Securing External Web Services
    - Legacy Integration
    - Hub Mediation
    - Web Services Management
    - Portal Acceleration

z/OS V1.10 - Communication Server Network Security Services (NSS) can provide ID authentication and access checks to DataPower appliance.

z/OS V1.11 – Comm Server Sysplex Distributor can manage workloads to DataPower

XI50z – DataPower appliance in blade form factor integrated into Zenterprise zBX
Summary ... What we have discussed

- XML Parsing technologies on z/OS with focus on XML System Services
  - Exploitation by higher level support
  - Performance
    - zAAP & zIIP support – providing additional price/performance benefits
- Added value middleware/products processing XML
- Positioning of DataPower versus on-platform XML processing support
Reference Information

- Toolkit FMIDs: HXML190 for Toolkit V1.9, HXML1A0 for Toolkit 1A0
- Toolkit Program ID: 5655-J51 (5655-I30 for S&S) - both unpriced for z/OS Licensees (same for V1.8, V1.9, V1.10)
  - See also related precursor papers:
    - http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/ WP101088 and
Questions?
Thank You!