
ORACLE®

ORACLE®

Developing XML Applications with Oracle XML DB and Oracle XML Developer's Kit

Tim Yu Senior Development Manager
Balu Sthanikam Principal Member of Technical Staff

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions.
The development, release, and timing of any features or functionality described for Oracle's products remain at the sole discretion of Oracle.

Agenda

- XML development platform
- XDK Java
- XDK C
- Demos
  - XML Diff and Patch
  - Scalability Demos
- Q&A

Oracle XML Developer's Kit (XDK)

- Standards-based XML dev packages and utilities
  - High-performance, highly scalable XML features
  - Supports all Oracle certified platforms
  - Available in Java, C, and C++
  - Supports multiple character encodings
- 24x7 professional support

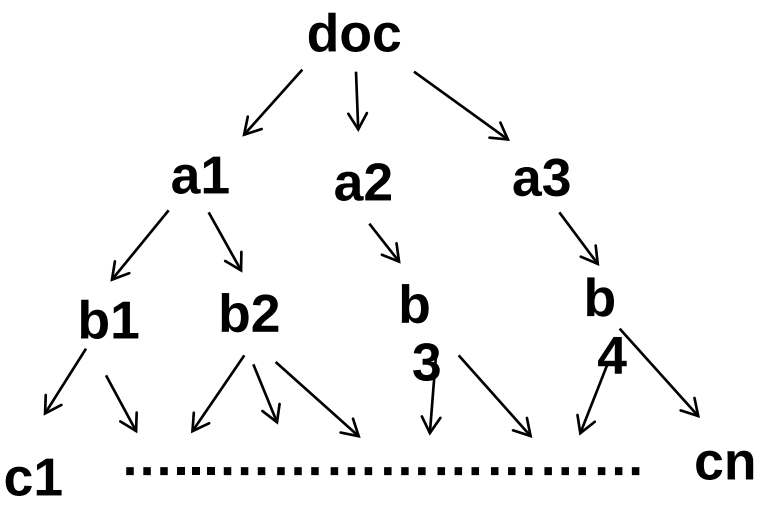
W3C/JSR Standards Support

Standard	Version	Language
XML	1.0	Java, C/C++
XML Namespaces	1.0	Java, C/C++
DOM	1.0/2.0/3.0	Java, C/C++
SAX + Extensions	1.0/2.0	Java, C/C++
XSLT + XPath	1.0/2.0	Java, C/C++
XML Schema	1.0	Java, C/C++
JAXP	1.2	Java
JAXB	1.0	Java
JSR 170	1.0	Java

Java XDK

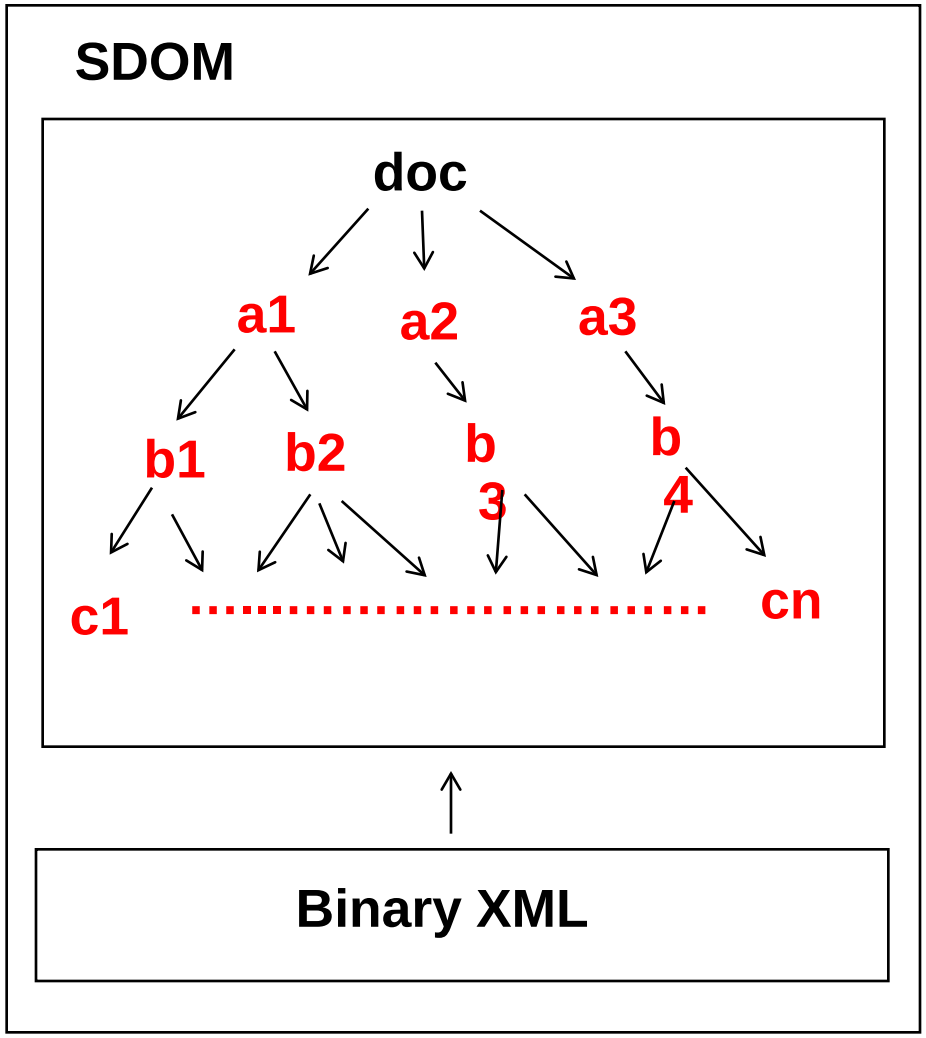
- Features
  - Scalable DOM
  - Streaming XPath
  - XSLT scalability improvements
  - XQJ Standard

Regular DOM



- Regular DOM
  - All nodes in memory
  - DOM size much bigger than input xml

Scalable DOM



- Scalable DOM
  - Nodes lazily created
  - Nodes can be garbage collected when not used
  - Less memory usage
  - Allows processing of larger documents
  - User configurable for more efficient access

Scalable DOM Creation

- SDOM creation from parser
- SDOM creation with existing binary xml

Creating SDOM from XML

```
DOMParser parser = new DOMParser();

parser.setAttribute(PARTIAL_DOM, TRUE);

parser.setAttribute(PAGE_MANAGER,
       new FilePageManager(“pageFile”));
… … // DOMParser other configuration

parser.parse(fileURL);

XMLDocument doc = parser.getDocument();
```

Generating Binary XML

```
// Step 1: get encoder
BinXMLProcessor proc =
  BinXMLProcessorFactory.createProcessor();

File bfile = new File(output_binary_file_name);
BinXMLStream bstr = proc.createBinXMLStream(bfile);
BinXMLEncoder enc = bstr.getEncoder();
```

Generating Binary XML

```
// Step 2: Set up parser

SAXParser parser = new SAXParser();
parser.setPreserveWhitespace(false);
parser.setContentHandler(enc.getContentHandler());
parser.setErrorHandler(enc.getErrorHandler());
parser.setProperty("http://xml.org/sax/properties/lexical-
  handler", enc.getLexicalHandler());
parser.setProperty("http://xml.org/sax/properties/declaration
  -handler", enc.getDeclHandler());

// Step 3: Encode
parser.parse(input_xml_file);
```

SDOM Using Existing Binary XML

```
// Create Scalable DOM from pre-encoded binary data
XMLDOMImplementation domi = new
  XMLDOMImplementation();

domi.setAttribute(XMLDocument.SCALABLE_DOM, true);

BinXMLProcessor proc =
  BinXMLProcessorFactory.createProcessor();

BinXMLStream stream =
  proc.createBinXMLStream(binaryXMLFile);

InfosetReader reader =
  stream.getDecoder().getReader();

XMLDocument  doc =
  (XMLDocument)domi.createDocument(reader);
```

Streaming XPath

- Streaming processing for expressions containing only following axes:
  - child
  - descendant
  - attribute
  - text
  - parent
  - ancestor

Streaming XPath

- Jaxp 1.3 API

```
// 1. Instantiate an XPathFactory.
  System.setProperty(
              XPathFactory.DEFAULT_PROPERTY_NAME
                        + ":" + XPathFactory.DEFAULT_OBJECT_MODEL_URI,
              "oracle.xml.xpath.JXPathFactory");
   javax.xml.xpath.XPathFactory factory =
                  javax.xml.xpath.XPathFactory.newInstance();

// 2. Use the XPathFactory to create a new XPath object
  javax.xml.xpath.XPath xpath = factory.newXPath();
```

Streaming XPath

```
// 3. Compile an XPath string into an XPathExpression
  javax.xml.xpath.XPathExpression expression = xpath.compile("/parent/child");

// 4. Evaluate the XPath expression on an input document
 String result = expression.evaluate(
                                    new org.xml.sax.InputSource("file:///dir/input.xml"));
```

XSLT Scalability Improvements

- XSLT optimizations
  - Uses streaming XPath evaluation
  - Improved SDOM usage
    - Minimizes holding of nodes
    - Allows SDOM nodes to be freed
    - Lower memory usage resulting in higher scalability

XQuery API for Java (XQJ)

- Standard, also known as JSR225. Reference implementation from Oracle.
- Similar to JDBC, but for XQuery /XML data rather than SQL/Relational data
- Typical XQJ program consists of:
  - Obtaining XQDataSource and XQConnection objects
  - Preparing and executing XQuery
  - Consuming XQuery result
  - Releasing resources

XQJ requirements and concepts

- Support for static and dynamic context concepts of XQuery
- Deferred variable binding mode for binding an XML stream as external variable value
- XQuery Data Model (XQDM) specific factory methods allowing creation and destruction of XQuery item and sequence objects and XQuery Sequence Type objects
- Mappings for conversion of Java objects to XQDM instances in the case of binding XQuery external variables and context item, and XQDM instances to Java objects in the case of consumption of XQuery results
- Fine-grain exception classes for better diagnosability

Oracle Reference Implementation and References

- The Oracle Reference Implementation as well as the XQJ specification can be downloaded from:

  http://jcp.org/aboutJava/communityprocess/final/jsr225/index.html

XDK-C

- XML Developer Kit for C
  - A Standalone Library for C/C++ applications
  - A set of command line tools
  - SQL functions for use from Oracle RDBMS
  - Designed for enterprise XML use both in MT and DB tiers
    - Functionality
    - Performance
    - Scalability

Why use XDK-C

- Value add compared to other Vendors
  - Designed for use in Mid Tier and/or DB Tier
    - Provides special optimizations against Oracle XML DB
      - XMLType (Native XML type) can be operated on in both MT and DB tier
      - User can leverage advanced XML storage, processing and query capabilities of XML DB
    - Core API is unified for both MT and DB tier
      - Application can run in both tiers with minimal changes
    - High performance and scalability
      - Streaming, pipelined processing with XML Parser and XML Schema Validator
      - Optimized algorithms in XML Diff and Patch
      - High performance XSLT and XPath processor
  - Extensive character set support
    - UTF-8, UTF-16, EBCDIC based character sets.
    - National multi byte character sets like Japanese, Korean etc.

XDK-C components

- XML Parser
  - SAX, DOM and Pull parsers
- XML Schema Validator
- XML Diff and Patch
- XMLType processing – integration with Oracle XML DB
- XSLT and XPath processors

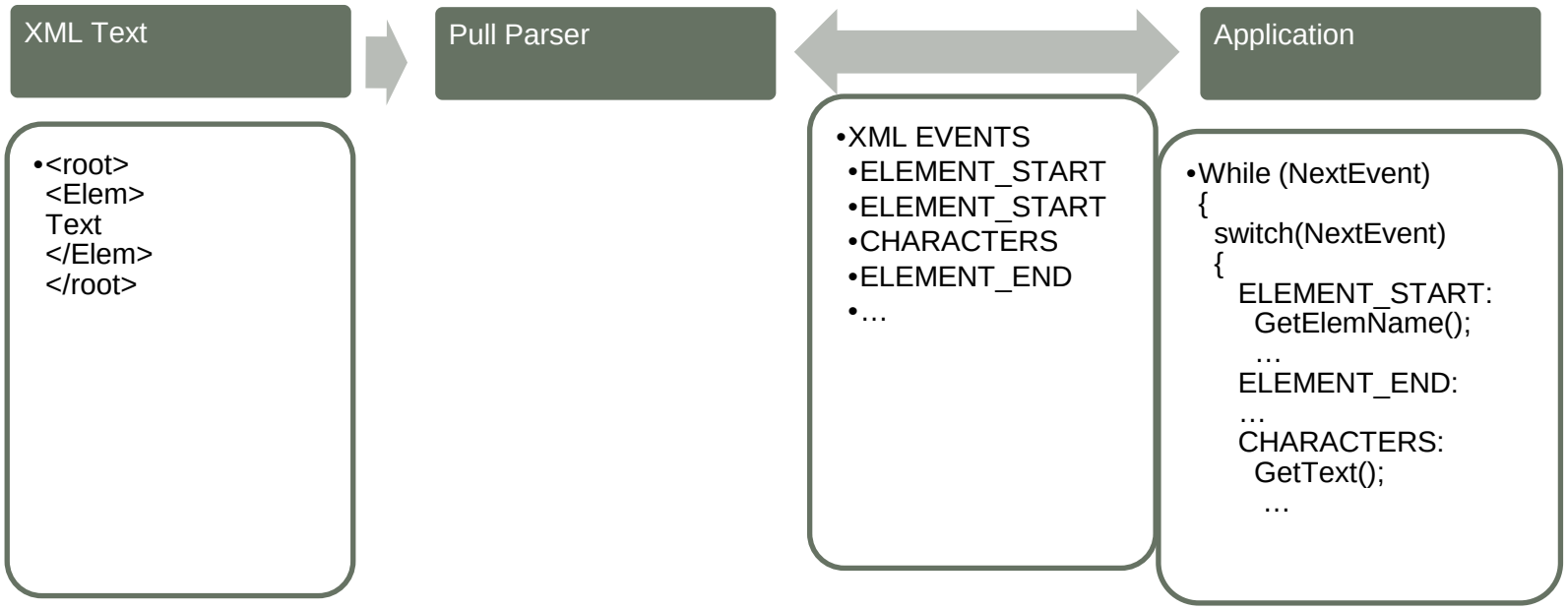
XDK–C components

- Components available as
  - C (most with C++ too) library functions
  - Command line tools
  - SQL functions
- Support multiple character sets
  - Unicode, multi byte national character sets etc.
- Can be used either inside or outside Database
  - Have unified API for both usages

High performance XML Parser

- Pull parser
  - Provides an Event representation (StAX like)
  - Does not build a tree in memory – less memory usage, faster
  - Much better application control over processing (pull vs. push)
- SAX and DOM parsers
- DTD validation in all parsers
- Can operate on different inputs - file, URL, database document, stream etc.

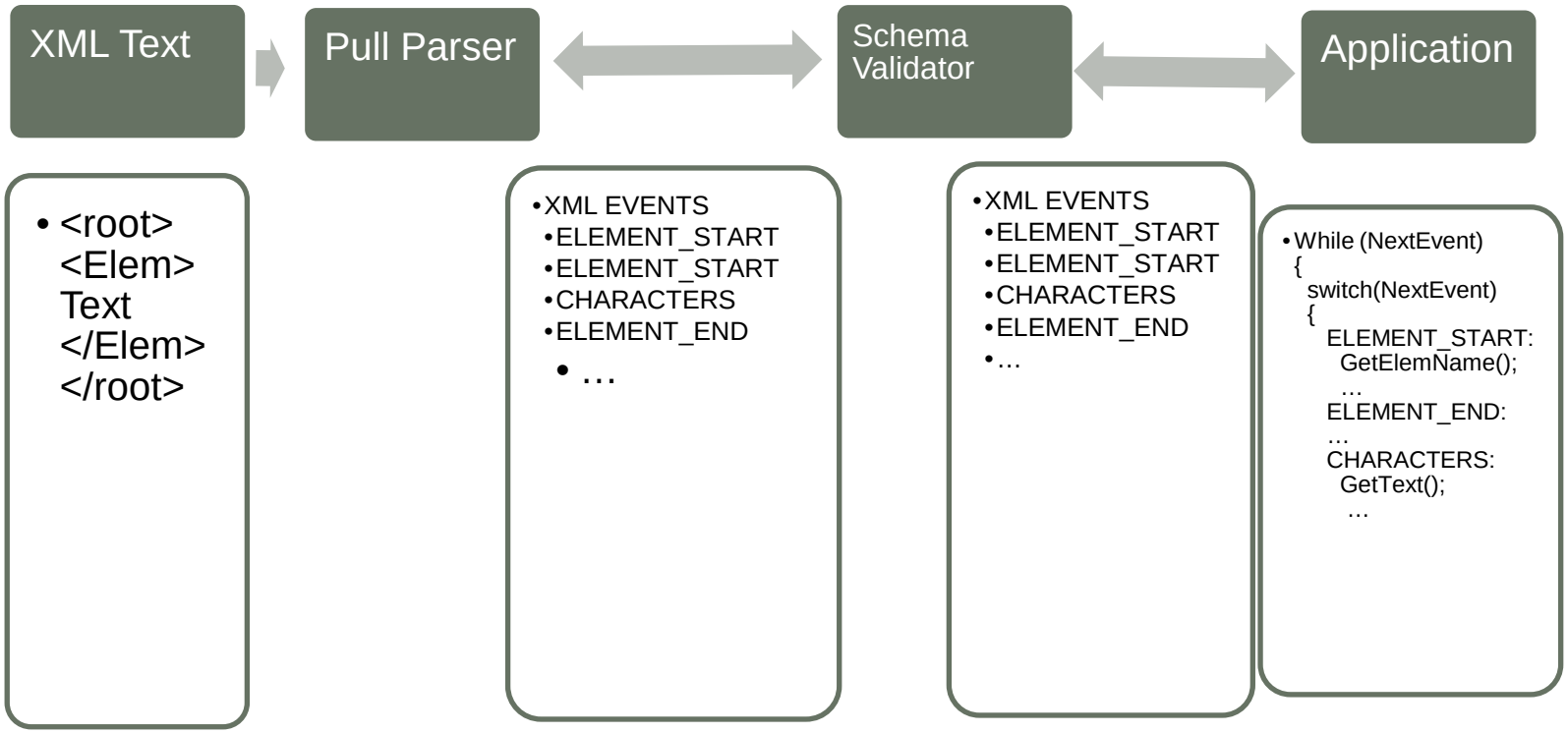
Pull Parser


XML Text
Pull Parser
Application
•<root>
<Elem>
Text
</Elem>
</root>
•XML EVENTS
•ELEMENT_START
•ELEMENT_START
•CHARACTERS
•ELEMENT_END
•…
•While (NextEvent)
{
switch(NextEvent)
{
ELEMENT_START:
GetElemName();
…
ELEMENT_END:
…
CHARACTERS:
GetText();
…

XML Schema Validator

- Full support of XML Schema 1.0 plus “XML Schema 1.0 specification Errata”
- Streaming capability with ‘pull’ processing
  - Schema Validates with out building a tree representation
    - Uses less memory, much faster
  - Output of Pull parser can be pipelined to combine parsing and validation
    - Faster processing for both operations

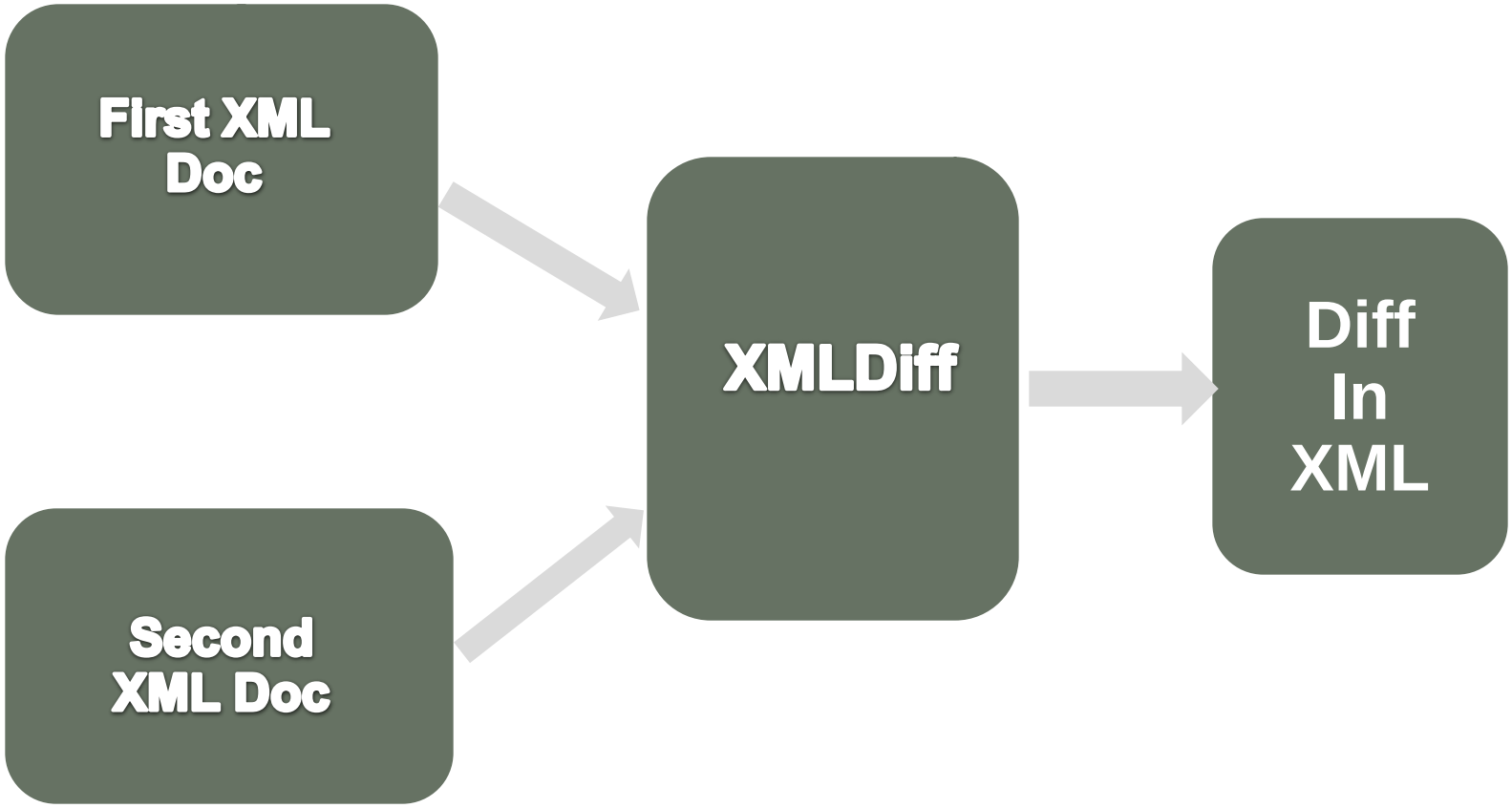
XML Schema Validator – Pipelined & Streaming processing



XML Diff

- Analogous to UNIX Diff but designed for XML
- Captures the diff between two XML documents in XML
- Diff conforms to Oracle's Xdiff XML Schema
- The diff can be used in several ways
  - Determine if two docs differ and if so where
  - Diff can be stored or transmitted in stead of storing/transmitting the whole doc
  - Facilitates applying the same change on multiple docs

XMLDiff



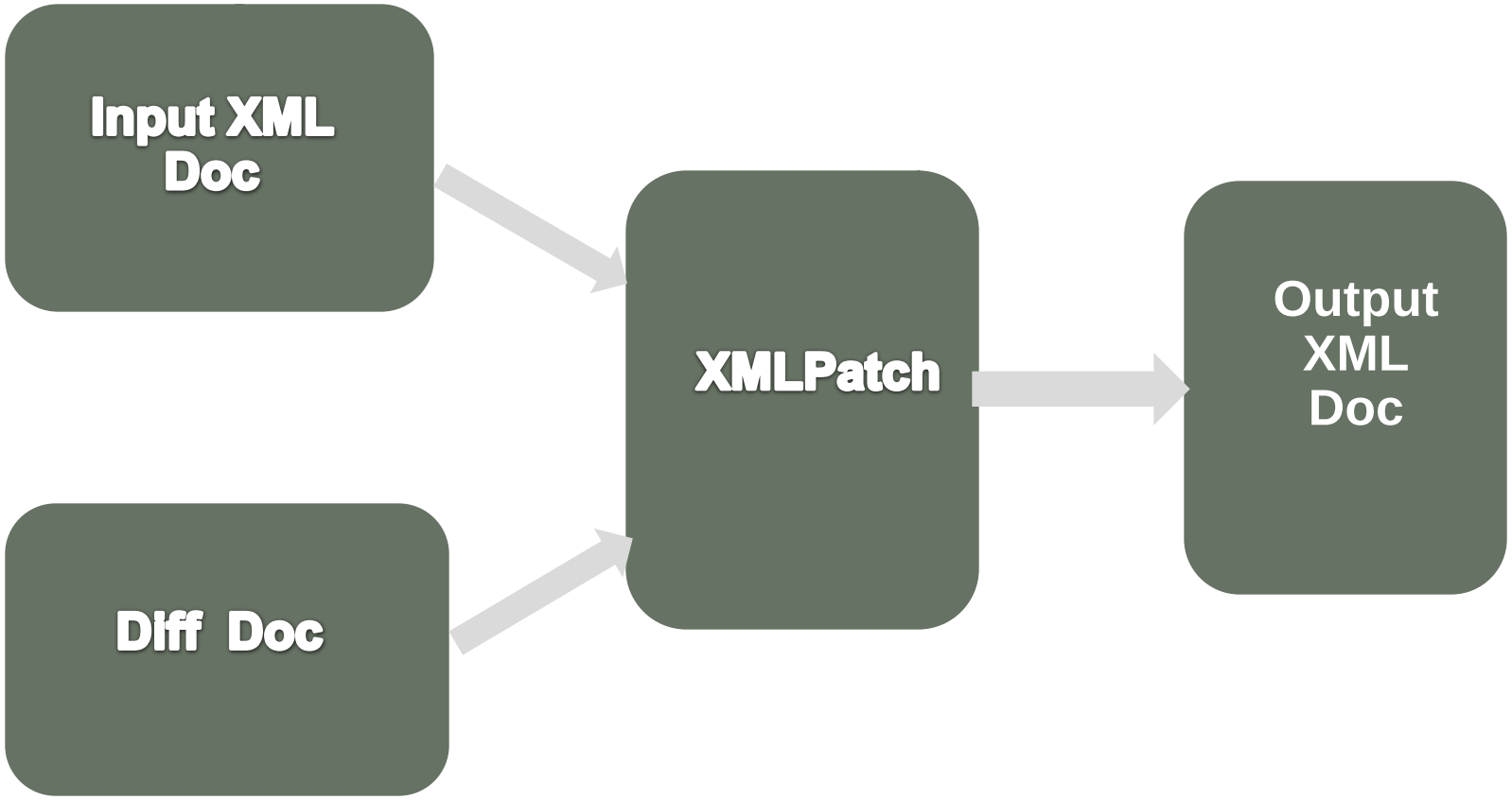
XML Diff Continued

- Optimized algorithm to find the minimal sub tree change
- Diff algorithm is customizable
  - Can work on hash values at specified depth
  - Hashing speeds up the diff
- XML Diff Can also be used for
  - XML digest (hash) generation
  - Associate a hash value for an XML document
  - Can be stored along with the doc for a very fast doc comparison

XML Patch

- Analogous to UNIX Patch but designed for XML
- Applies changes (diff) in XML (Xdiff Schema) to input document
- Diff can be either generated by XMLDiff or hand coded
- Useful for
  - applying same changes to several XML docs
  - The client can send the diff alone over N/W and server applies the diff to doc in repository

XMLPatch



t

```
[bsthanik_xdk_main]:tklocal/xdkdemo >cat first.xml
<?xml version="1.0" encoding="UTF-8"?>
<PurchaseOrder>
    <Reference>ALLEN-20011021102116O2PDT</Reference>
    <Requestor>Michael L. Scott</Requestor>
    <User>SCOTT</User>
    <CostCenter>S30</CostCenter>
    <ShippingInstructions>
        <name>Michale L. Scott</name>
        <address>
          300 Oracle Parkway
          Redwood Shores
          CA 94065, USA
        </address>
        <telephone>650 506 7300</telephone>
    </ShippingInstructions>
    <SpecialInstructions>
      <DeliveryMethod>Courier</DeliveryMethod>
      <TimeFrame>5 Days.</TimeFrame>
      <Signature>Yes</Signature>
    </SpecialInstructions>
    <LineItems>
        <LineItem ItemNumber="1">
            <Description>Vagabond</Description>
            <Part Id="037429148921" UnitPrice="29.95" Quantity="4"/>
        </LineItem>
        <LineItem ItemNumber="2">
            <Description>Samurai Two: Duel at Ichijoji Temple</Description>
            <Part Id="037429125526" UnitPrice="29.95" Quantity="1"/>
        </LineItem>
        <LineItem ItemNumber="3">
            <Description>Rififi</Description>
            <Part Id="037429155622" UnitPrice="29.95" Quantity="3"/>
        </LineItem>
        <LineItem ItemNumber="4">
            <Description>Spartacus</Description>
            <Part Id="715515011723" UnitPrice="39.96" Quantity="3"/>
        </LineItem>
        <LineItem ItemNumber="5">
            <Description>Kwaidan</Description>
            <Part Id="037429152027" UnitPrice="29.95" Quantity="2"/>
        </LineItem>
        <LineItem ItemNumber="6">
            <Description>Black Orpheus</Description>
            <Part Id="037429138328" UnitPrice="29.95" Quantity="2"/>
        </LineItem>
        <LineItem ItemNumber="7">
            <Description>The Shop on Main Street</Description>
            <Part Id="037429156124" UnitPrice="29.95" Quantity="3"/>
        </LineItem>
        <LineItem ItemNumber="8">
            <Description>The 39 Steps</Description>
            <Part Id="037429135228" UnitPrice="39.95" Quantity="3"/>
        </LineItem>
    </LineItems>
</PurchaseOrder>
[bsthanik_xdk_main]:tklocal/xdkdemo >
[bsthanik_xdk_main]:tklocal/xdkdemo >
[bsthanik_xdk_main]:tklocal/xdkdemo >
```

x

```
[bsthanik_xdk_main]:tklocal/xdkdemo >cat second.xml
<?xml version="1.0" encoding="UTF-8"?>
<PurchaseOrder>
    <Reference>ALLEN-20011021102116O2PDT</Reference>
    <Requestor>Michael L. Scott</Requestor>
    <User>SCOTT</User>
    <CostCenter>S50</CostCenter>
    <ShippingInstructions>
        <name>Michale L. Scott</name>
        <address>
          300 Oracle Parkway
          Redwood Shores
          CA 94065, USA
        </address>
        <telephone>650 506 7300</telephone>
    </ShippingInstructions>
    <LineItems>
        <LineItem ItemNumber="1">
            <Description>Vagabond</Description>
            <Part Id="037429148921" UnitPrice="29.95" Quantity="4"/>
        </LineItem>
        <LineItem ItemNumber="2">
            <Description>Samurai Two: Duel at Ichijoji Temple</Description>
            <Part Id="037429125526" UnitPrice="29.95" Quantity="1"/>
        </LineItem>
        <LineItem ItemNumber="3">
            <Description>Rififi</Description>
            <Part Id="037429155622" UnitPrice="29.95" Quantity="3"/>
        </LineItem>
        <LineItem ItemNumber="4">
            <Description>Spartacus</Description>
            <Part Id="715515011723" UnitPrice="39.96" Quantity="3"/>
        </LineItem>
        <LineItem ItemNumber="5">
            <Description>Kwaidan</Description>
            <Part Id="037429152027" UnitPrice="29.95" Quantity="2"/>
        </LineItem>
        <LineItem ItemNumber="6">
            <Description>Black Orpheus</Description>
            <Part Id="037429138328" UnitPrice="29.95" Quantity="2"/>
        </LineItem>
        <LineItem ItemNumber="7">
            <Description>The Shop on Main Street</Description>
            <Part Id="037429156124" UnitPrice="29.95" Quantity="3"/>
        </LineItem>
        <LineItem ItemNumber="7.5">
            <Description>New Lineitem inserted</Description>
            <Part Id="037429135999" UnitPrice="99999" Quantity="77"/>
        </LineItem>
        <LineItem ItemNumber="8">
            <Description>The 39 Steps</Description>
            <Part Id="037429135228" UnitPrice="39.95" Quantity="3"/>
        </LineItem>
    </LineItems>
</PurchaseOrder>
[bsthanik_xdk_main]:tklocal/xdkdemo >
[bsthanik_xdk_main]:tklocal/xdkdemo >
[bsthanik_xdk_main]:tklocal/xdkdemo >
[bsthanik_xdk_main]:tklocal/xdkdemo >
```

```
[bsthanik_xdk_main]:tklocal/xdkdemo >xmldiff first.xml second.xml
<?xml version="1.0" encoding="UTF-8"?>
<xd:xdiff xsi:schemaLocation="http://xmlns.oracle.com/xdb/xdiff.xsd http://xmlns.oracle.com/xdb/xdiff.xsd" xmlns:xd="http://xmlns.oracle.com/xdb/xdiff.xsd" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
    <?oracle-xmldiff operations-in-docorder="true" output-model="snapshot" diff-algorithm="global"?>
    <xd:update-node xd:node-type="text" xd:xpath="/PurchaseOrder[1]/CostCenter[1]/text()[1]">
        <xd:content>S50</xd:content>
    </xd:update-node>
    <xd:delete-node xd:node-type="element" xd:xpath="/PurchaseOrder[1]/SpecialInstructions[1]"/>
    <xd:insert-node-before xd:node-type="element" xd:xpath="/PurchaseOrder[1]/LineItems[1]/LineItem[8]">
        <xd:content>
            <LineItem ItemNumber="7.5">
                <Description>New Lineitem inserted</Description>
                <Part Id="037429135999" UnitPrice="99999" Quantity="77"/>
            </LineItem>
        </xd:content>
    </xd:insert-node-before>
</xd:xdiff>
[bsthanik_xdk_main]:tklocal/xdkdemo >xmldiff first.xml second.xml > diff.xml
[bsthanik_xdk_main]:tklocal/xdkdemo >
```

```
[bsthanik_xdk_main]:tklocal/xdkdemo >xmldiff first.xml second.xml > diff.xml
[bsthanik_xdk_main]:tklocal/xdkdemo >xmlpatch first.xml diff.xml
<?xml version="1.0" encoding="UTF-8"?>
<PurchaseOrder>
    <Reference>ALLEN-20011021102116O2PDT</Reference>
    <Requestor>Michael L. Scott</Requestor>
    <User>SCOTT</User>
    <CostCenter>S50</CostCenter>
    <ShippingInstructions>
        <name>Michale L. Scott</name>
        <address>
          300 Oracle Parkway
          Redwood Shores
          CA 94065, USA
        </address>
        <telephone>650 506 7300</telephone>
    </ShippingInstructions>
    <LineItems>
        <LineItem ItemNumber="1">
            <Description>Vagabond</Description>
            <Part Id="037429148921" UnitPrice="29.95" Quantity="4"/>
        </LineItem>
        <LineItem ItemNumber="2">
            <Description>Samurai Two: Duel at Ichijoji Temple</Description>
            <Part Id="037429125526" UnitPrice="29.95" Quantity="1"/>
        </LineItem>
        <LineItem ItemNumber="3">
            <Description>Rififi</Description>
            <Part Id="037429155622" UnitPrice="29.95" Quantity="3"/>
        </LineItem>
        <LineItem ItemNumber="4">
            <Description>Spartacus</Description>
            <Part Id="715515011723" UnitPrice="39.96" Quantity="3"/>
        </LineItem>
        <LineItem ItemNumber="5">
            <Description>Kwaidan</Description>
            <Part Id="037429152027" UnitPrice="29.95" Quantity="2"/>
        </LineItem>
        <LineItem ItemNumber="6">
            <Description>Black Orpheus</Description>
            <Part Id="037429138328" UnitPrice="29.95" Quantity="2"/>
        </LineItem>
        <LineItem ItemNumber="7">
            <Description>The Shop on Main Street</Description>
            <Part Id="037429156124" UnitPrice="29.95" Quantity="3"/>
        </LineItem>
        <LineItem ItemNumber="7.5">
            <Description>New Lineitem inserted</Description>
            <Part Id="037429135999" UnitPrice="99999" Quantity="77"/>
        </LineItem>
        <LineItem ItemNumber="8">
            <Description>The 39 Steps</Description>
            <Part Id="037429135228" UnitPrice="39.95" Quantity="3"/>
        </LineItem>
    </LineItems>
</PurchaseOrder>
[bsthanik_xdk_main]:tklocal/xdkdemo >
```

XMLType – integration with Oracle XML DB

- XDK-C is integrated to work with OCI (Oracle Call Interface) and Oracle XML DB
  - With Oracle XML DB user can make use of
    - Advanced XML Storage, processing and query capabilities
    - XML Query (XQuery), XPath, XSLT
    - Security, scalability and other enterprise features of Oracle RDBMS
- Client application can use combined capabilities of XDK-C library and OCI
  - Native XMLType processing for XML Data
  - XMLType can be bound to different kinds of repository in Oracle XMLDB
    - Binary XML
    - Object Relational
    - CLOB storages
    - XML DB Repository

XSLT and XPath processor

- XSLT compiler and XVM (XSLT Virtual Machine)
  - Compiler creates a bytecode representation of the stylesheet or XPath
  - Compiled version is subsequently supplied to XSLT VM for XSLT transformation or XPath processing
  - High performance
    - Much faster than the conventional ‘interpreter’ approach for XSLT
    - Uses less memory
  - Bytecode is platform independent and can be saved and reused

Oracle XDK Resources

- **Oracle Technology Network**
  - http://otn.oracle.com/tech/xml
  - Downloads, Demos, Samples, Papers
  - XML Forum

The preceding is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions.
The development, release, and timing of any features or functionality described for Oracle's products remain at the sole discretion of Oracle.

ORACLE IS THE INFORMATION COMPANY