

# XML Programming: JAXP

Andy Clark



# Java API for XML Processing

- Standard Java API for loading, creating, accessing, and transforming XML documents
  - JAXP: parsing
    - Supports DOM and SAX
  - TrAX: transformation
    - Generic transformation framework
- Java standard extension (i.e. javax package)
  - Defined by Sun JCP

# JAXP Implementation

- JAXP is a set of defined interfaces
  - Parsers distribute implementation configured to return appropriate parser instances
- Search method for finding implementation
  - System property: “javax.xml.parsers.DocumentBuilderFactory”
  - JAXP properties: “%JAVA\_HOME%/lib/jaxp.properties”
  - Jar services:  
“META-INF/services/javax.xml.parsers.DocumentBuilderFactory”

# javax.xml.parsers

- DOM
  - `DocumentBuilder`, `DocumentBuilderFactory`
- SAX
  - `SAXParser`, `SAXParserFactory`
- General
  - `FactoryConfigurationError`
  - `ParserConfigurationException`

# DOM: Parsing Document

- Get instance of document builder factory
  - `DocumentBuilderFactory factory = DocumentBuilderFactory.newInstance();`
  - `factory.setNamespaceAware(true);`
  - `factory.setValidating(false);`
- Create document builder
  - `DocumentBuilder builder = factory.newDocumentBuilder();`
- Parse document
  - `Document document = builder.parse("document.xml");`

# DOM: Creating Empty Document

- Get instance of document builder factory
  - `DocumentBuilderFactory factory = DocumentBuilderFactory.newInstance();`
- Create document builder
  - `DocumentBuilder builder = factory.newDocumentBuilder();`
- Create document
  - `Document document = builder.newDocument();`

# SAX: Parsing Document

- Get instance of SAX parser factory
  - `SAXParserFactory factory = SAXParserFactory.newInstance();`
  - `factory.setNamespaceAware(true);`
  - `factory.setValidating(false);`
- Create SAX parser
  - `SAXParser parser = factory.newSAXParser();`
- Parse document
  - `parser.parse("document.xml", new MyHandler());`

# Configuration

- DocumentBuilderFactory
  - Allows setting of “attributes”
    - String identifier
    - Object value
- SAXParserFactory
  - Allows setting of “features”
    - String identifier
    - boolean state
  - SAXParser allows setting of “properties”
    - String identifier, Object value



# Error Handling and Entity Resolution

- DocumentBuilder
  - Can set a SAX ErrorHandler
  - Can set a SAX EntityResolver
- SAXParser
  - Must query XMLReader first, then
    - Can set ErrorHandler
    - Can set EntityResolver

# Useful Links

- JAXP
  - <http://java.sun.com/xml/>
- Apache Xerces XML Parser
  - <http://xml.apache.org/xerces2-j/>

# **XML Programming: JAXP**

Andy Clark

