Fast Infoset & Fast Web Services

Paul Sandoz
Staff Engineer
Sun Microsystems

New standards on the way

- Two new specifications will go for Consent to Last Call in Moscow
- Fast Infoset
  - ITU-T Rec. X.891 | ISO/IEC 24824-1
- Fast Web Services
  - ITU-T Rec. X.892 | ISO/IEC 24824-2

Standards for optimizing the use of XML and Web services

- Generic Applications of ASN.1 (new category)
  - Fast Infoset
  - Fast Web Services

Standards for optimizing the use of XML and Web services

- Generic Applications of ASN.1 (new category)
  - Fast Infoset
  - Fast Web Services
- ASN.1 Encoding Rules
  - Mapping W3C XML Schema Definitions into ASN.1
    - ITU-T Rec. X.694 | ISO/IEC 8825-5
    - International standard

Overview

- Optimization of XML documents
- XML Information Set and Fast Infoset
- XML Schema and X.694
- Web services and Fast Web Services

Optimization of XML documents: Why?

- XML documents can be larger than binary equivalents
- XML documents can take longer to serialize and parse than binary equivalents
- Size, parsing and serializing characteristics of XML documents cause issues in certain environments

Redundancy compression does not always cut it

- In general the relationship between XML serialization and parsing is asymmetric
  - Takes longer to parse than serialize
- Compression (like GZIP) can flip this relationship
  - Increases serializing and parsing costs to reduce size

Use cases

- Military
  - Unify disparate systems
  - Bandwidth and processing constraints
- Telecoms
  - Wireless systems: mobile and satellite
  - Backend systems: high-throughput
- 3D graphics

People want to have their cake and eat it!

- Want to use XML
  - Reuse concepts, specifications and source code
  - Interoperate
- But don't want the size, parsing and serializing costs associated with XML documents
  - Mainly for the transmission of XML documents

People can! With the ASN.1 standards

- Optimizes size, serialization and parsing
- Integrate into XML “system” or “stack” without bifurcation of applications
- Binary documents replace XML documents

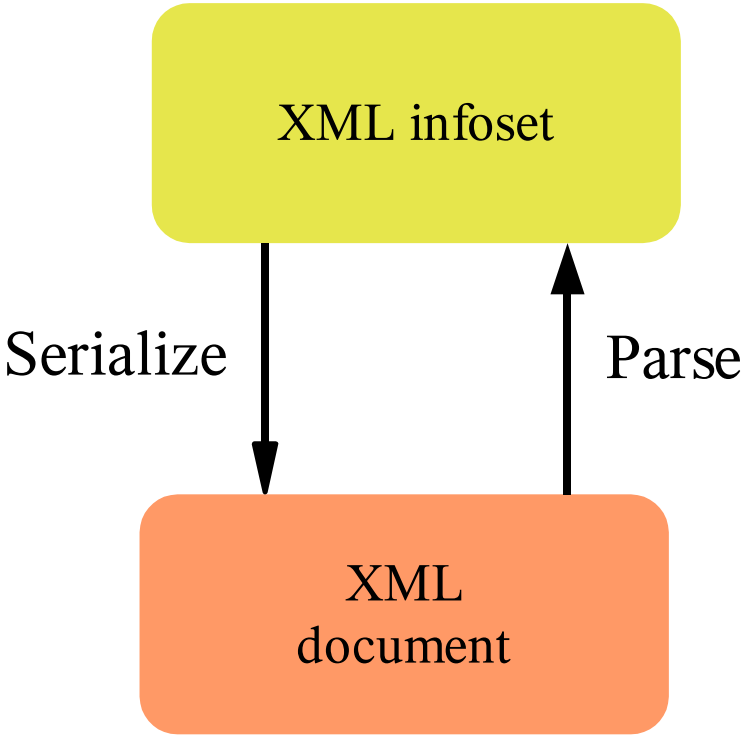
XML Information Set and Fast Infoset

Paul.Sandoz@Sun.Com Sun Microsystems

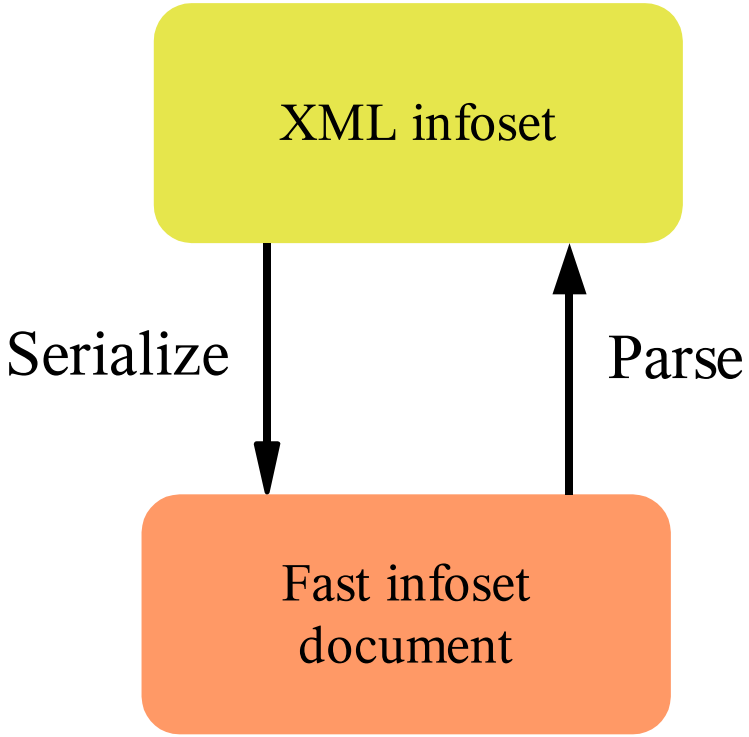
XML Information Set

- Specifies the result of parsing an XML document
- Consists of “glossary” of terms for talking about the result
  - Information items and properties
- XML Schema and SOAP 1.2 are based on the Information Set
- Simple data model for XML

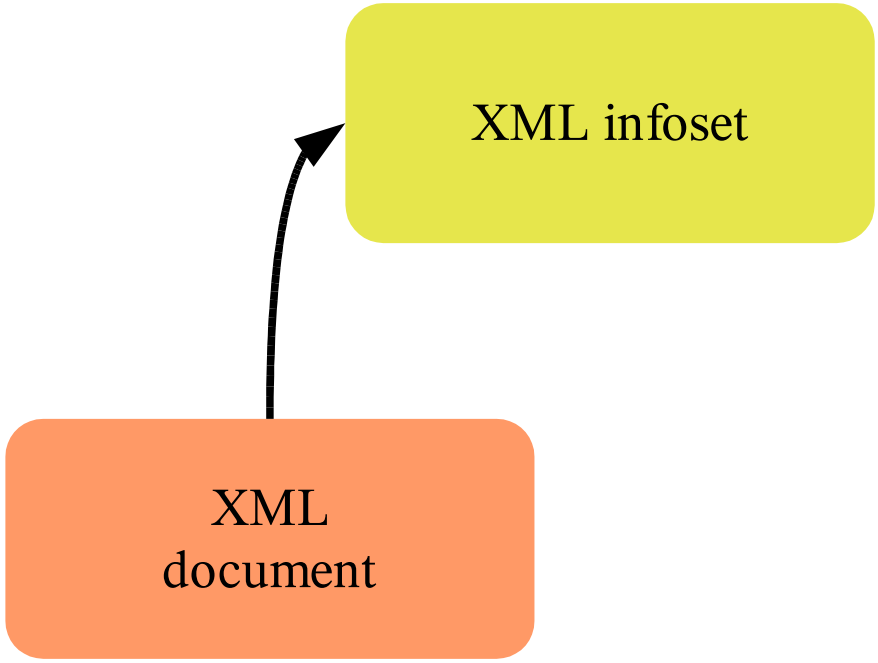
XML Information Set parsing and serializing



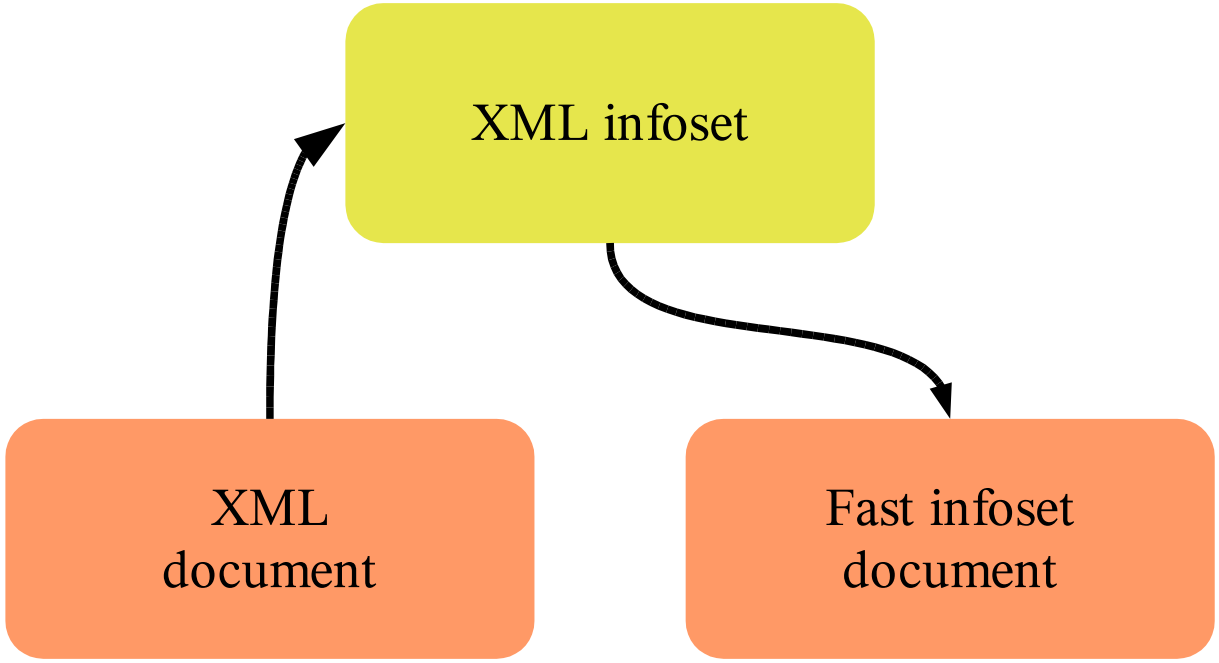
Fast Infoset Set parsing and serializing



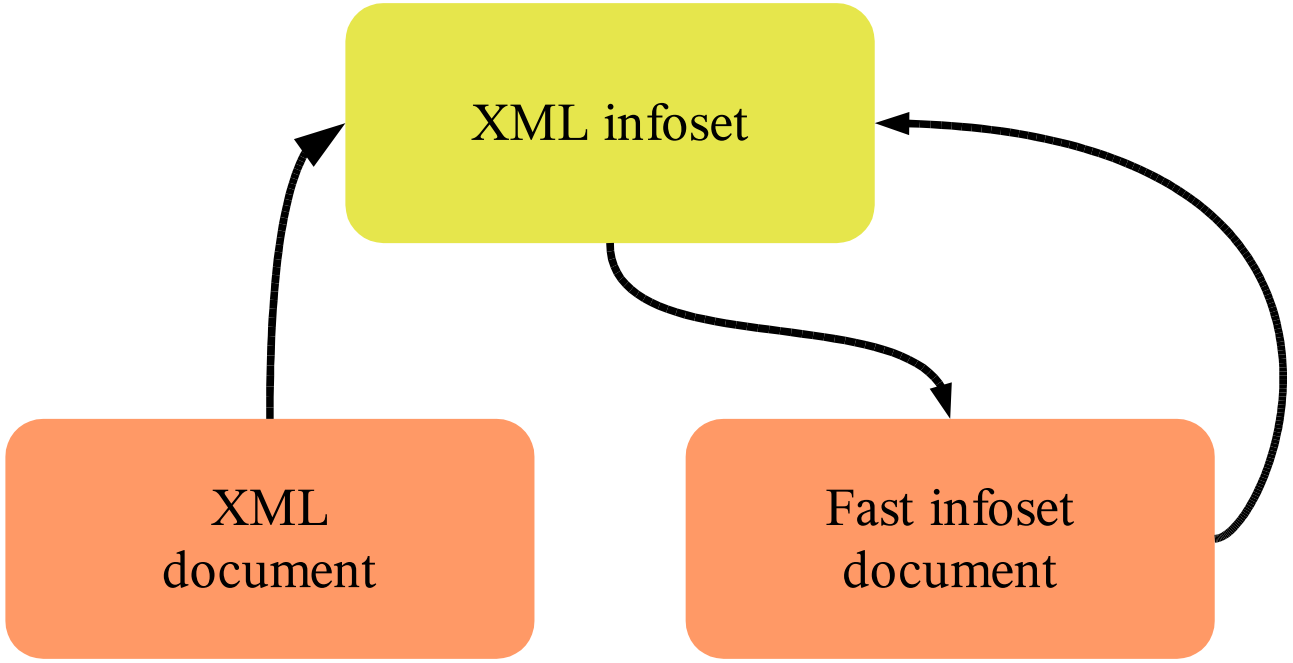
Round tripping



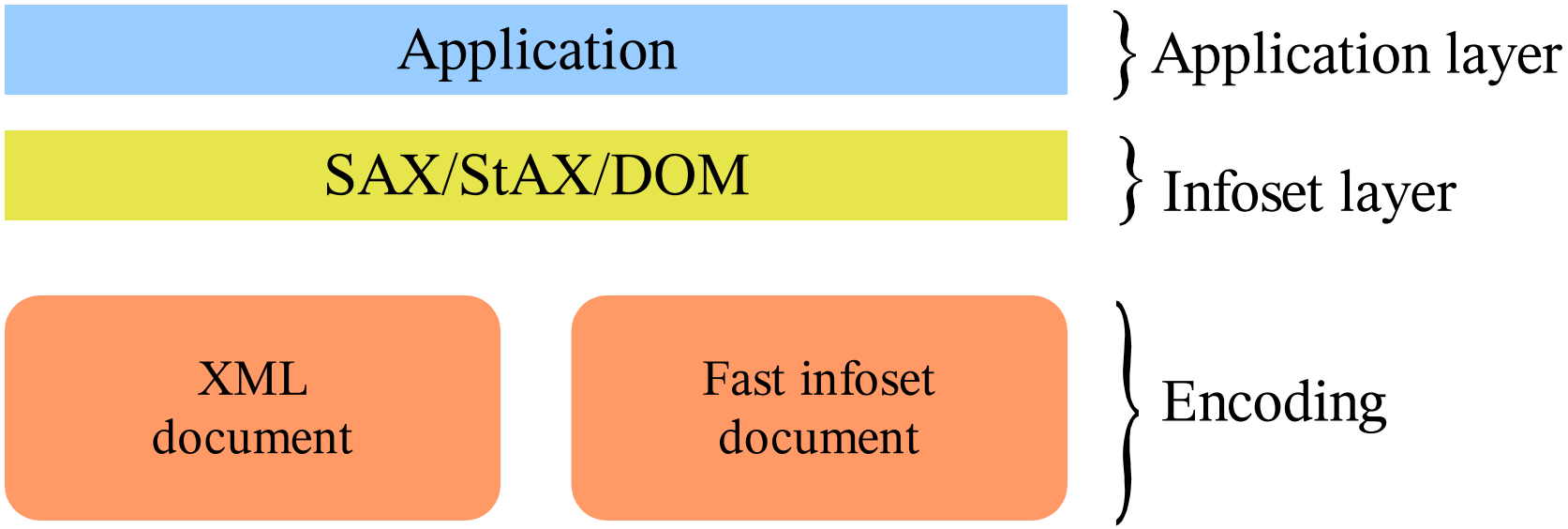
Round tripping



Round tripping



Integration



Fast Infoset features

- Retains self-description or self-structuring property of XML
- Compaction using tokenization
- Vocabularies
- Restricted Alphabets
- Encoding algorithms
- XTech 2005 paper will go into more details

Fast Infoset optimization

- Serialization speed comparable to XML
- SAX parsing speed 3x to 5x faster than Apache Xerces XML parser
- Fast infoset documents 30% to 70% smaller than XML documents
- See http://fi.dev.java.net for detailed results

Fast Infoset adoption

- Adopted by Web3D organization
  - Part 3: Binary Encoding: ISO/IEC 19776, Extensible 3D (X3D) encodings
- Open source implementation at http://fi.dev.java.net
  - Join the project and contribute!
- Sun is integrating open source implementation into next release of the Java Web Services Developer Pack

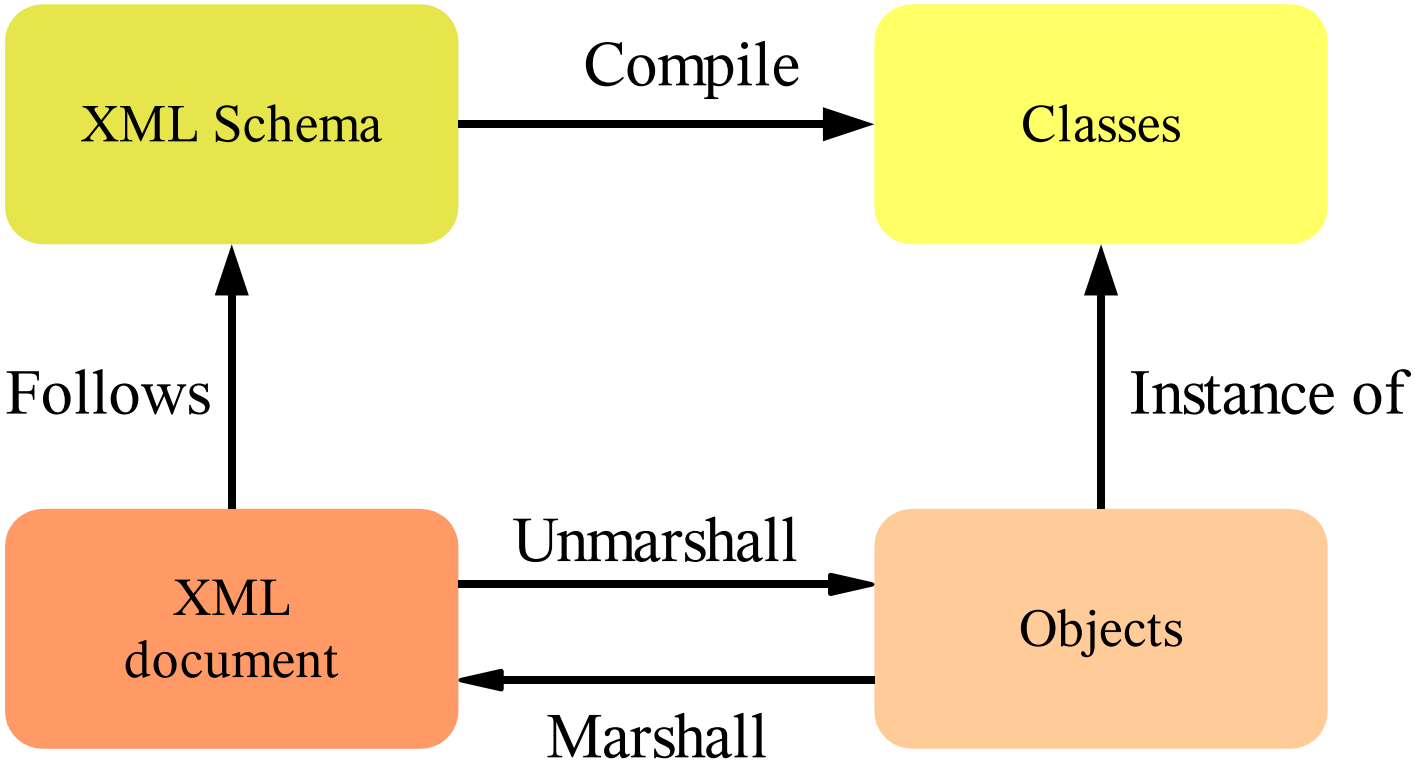
(W3C) XML Schema and X.694

Paul.Sandoz@Sun.Com Sun Microsystems

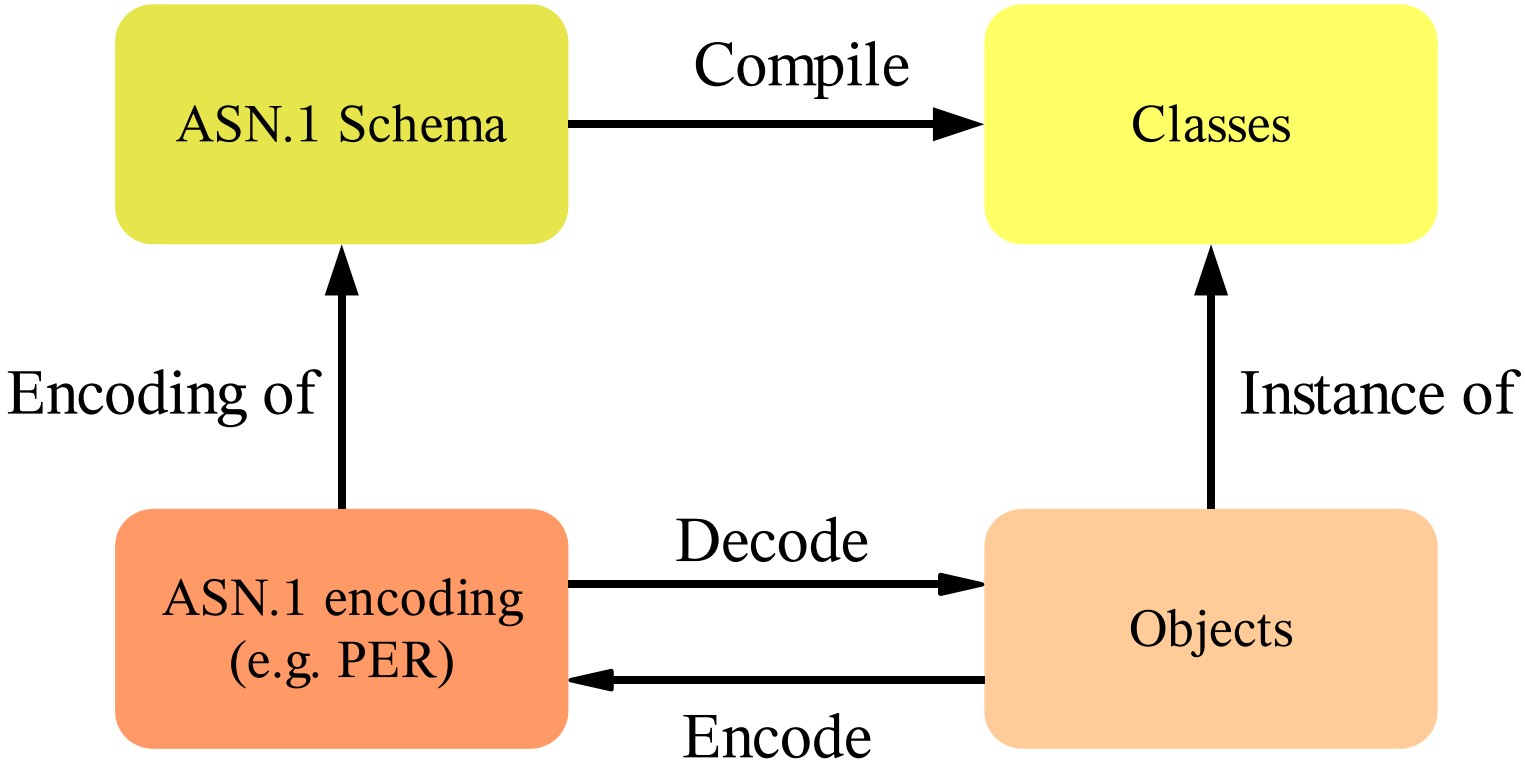
XML Schema

- Defines structure and content of an XML document
- Validation of XML documents
- Enables binding of XML documents to programmatic types

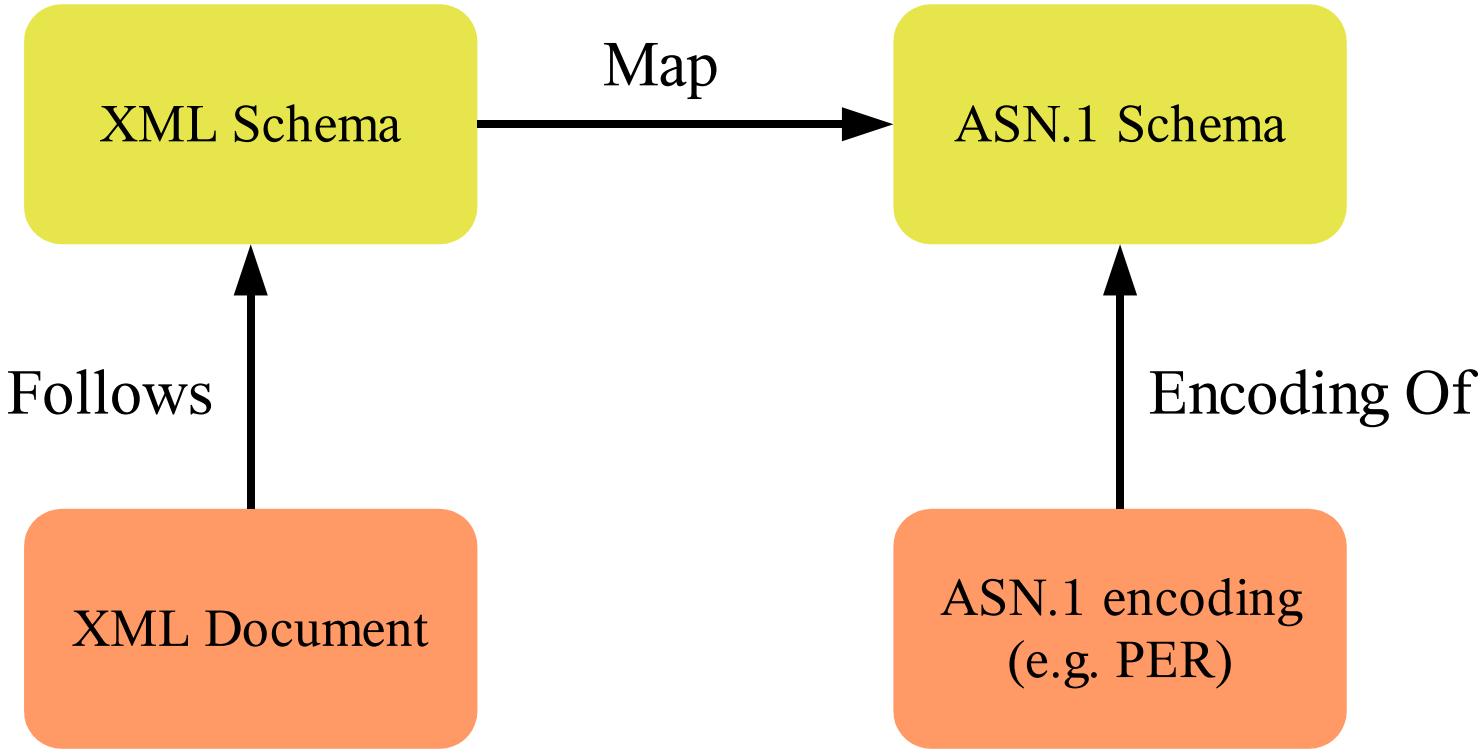
XML Schema binding



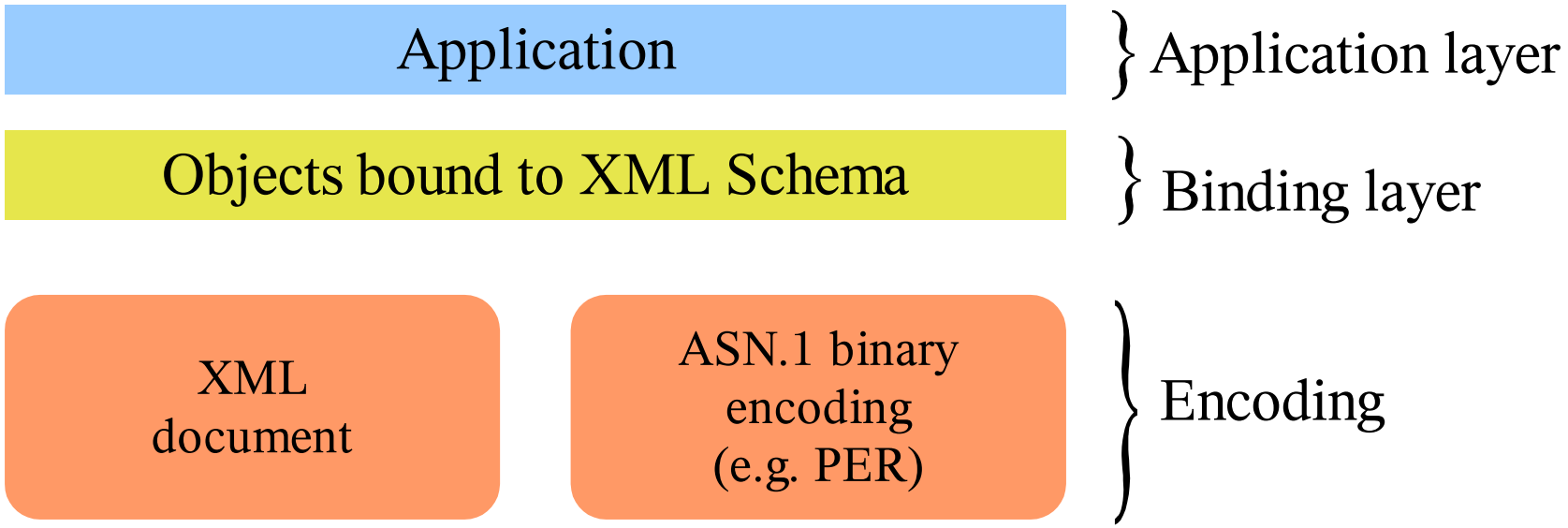
ASN.1 Schema binding



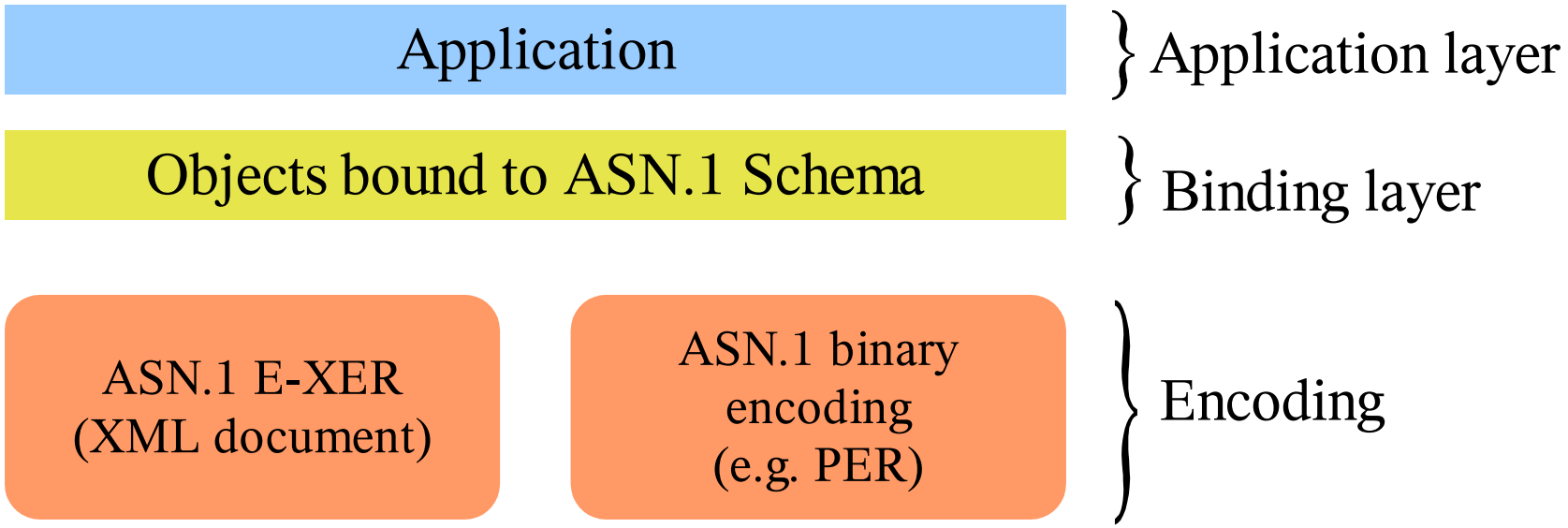
Mapping of XML Schema to ASN.1 Schema



Integration



Integration starting from ASN.1 (ASN.1 is an XML schema too!)



X.694 features

- Does not retain self-description or self-structuring property of XML
  - Creator and processor require same schema to interoperate
- 100% XML Schema support
- Compact documents that are fast to serialize and parse

X.694 optimization

- Round trip encoding/decoding of PER encoded documents 4x to 10x faster than serializing/parsing XML documents
- PER encoded documents 80% smaller than XML documents
- For small XML documents equivalent PER documents can be smaller than GZIPed XML

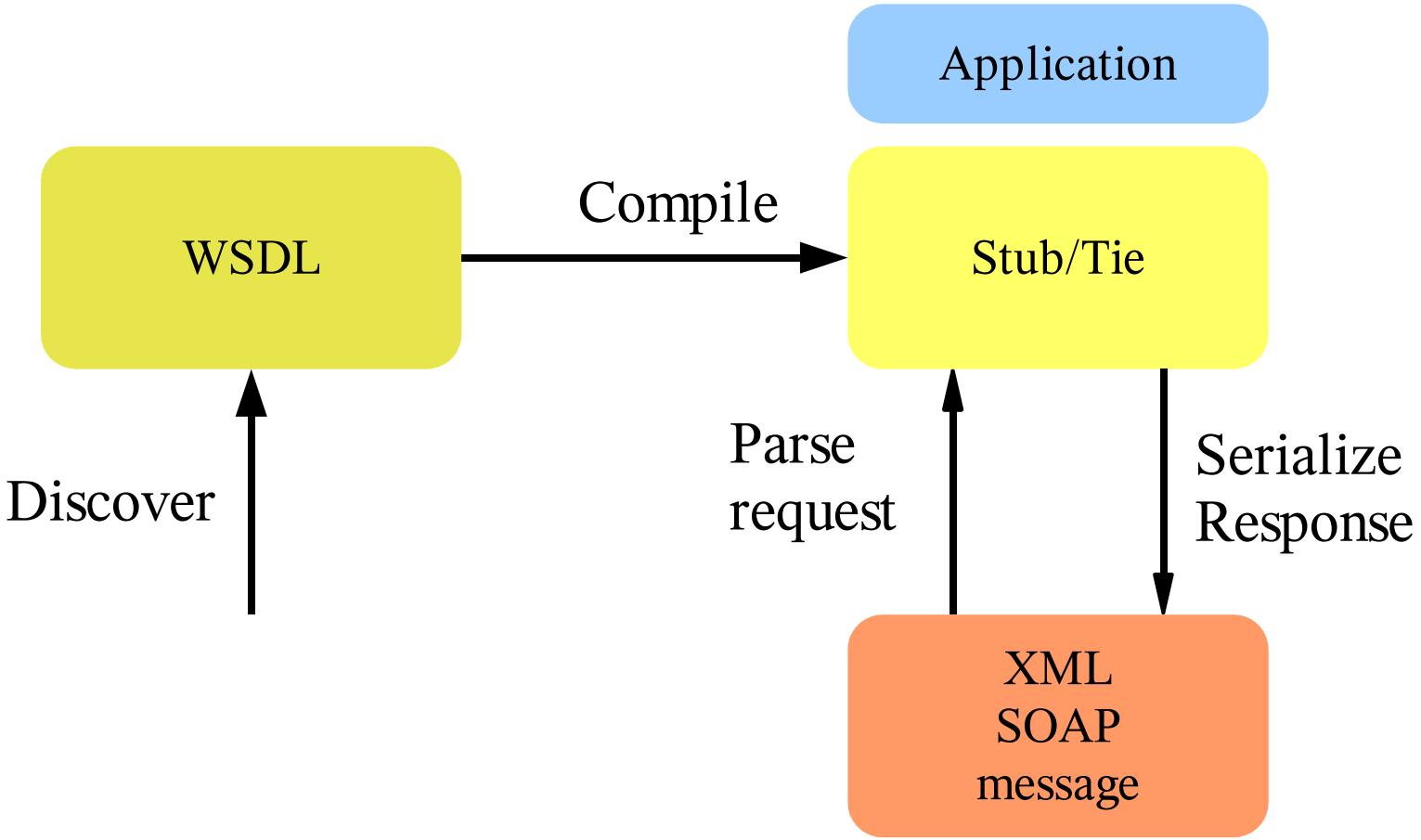
X.694 adoption

- Fundamental part of Fast Web Services
- OSS Nokalva has implementation
  - OSS XML Schema Tools

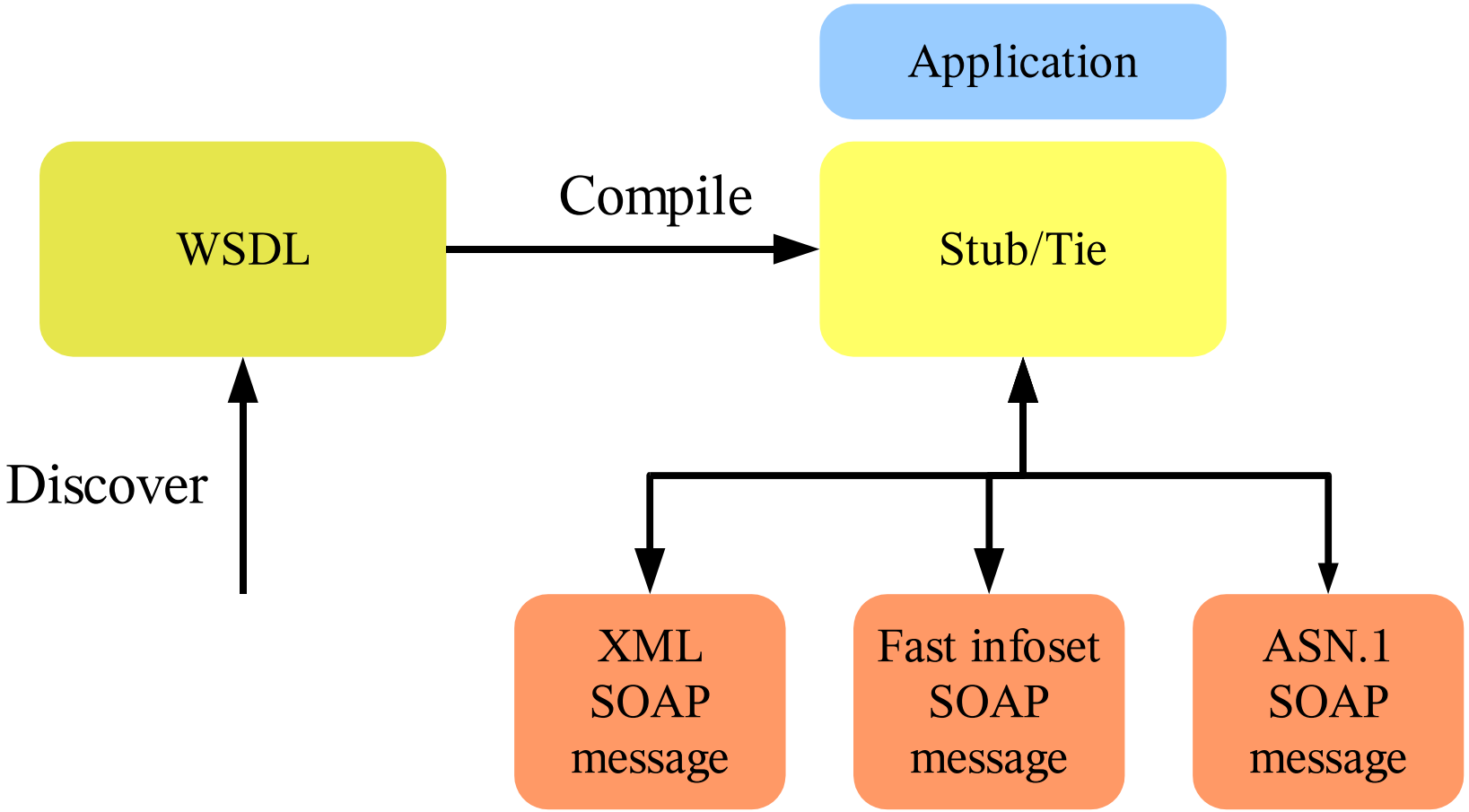
Web services and Fast Web Services

Paul.Sandoz@Sun.Com Sun Microsystems

Web services



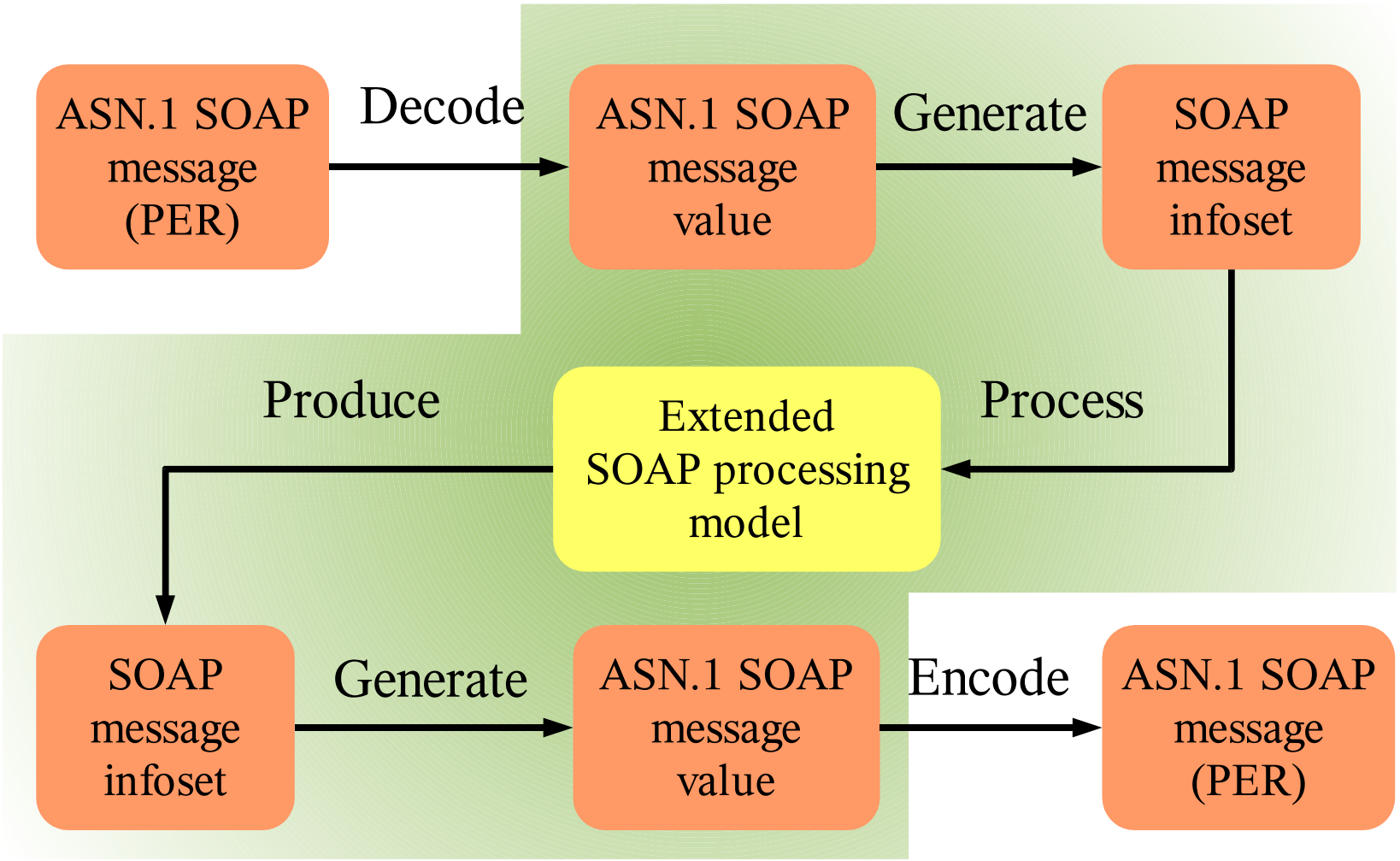
Fast Web Services



ASN.1 SOAP messages

- Binary representation of SOAP 1.2 message infoset
- Use X.694 and PER for the content of SOAP messages
- SOAP processing model extended to process ASN.1 SOAP messages

Conceptual processing



Fast Web Services features

- Supports WS-I conforming WSDL with no modifications
- Annotations to WSDL
  - Declare Fast Web Services on WSDL port
  - Declare use of OIDs and ROIDs instead of URIs
- ASN.1 SOAP message content
  - X.694 encoded PER
  - Fast infoset documents

Fast Web Services adoption

- OSS Nokalva will have product in FYQ1
- Sun integrating Fast Web Services support for Fast Infoset into Java Web Services Developer Pack
- Web Service for Module database
  - Fast Web Services support for Fast Infoset

Summary

- Fast Infoset
  - Self-describing or self-structured messages
- X.694
  - Highly compact and efficient non-self-describing messages
- Fast Web Services
  - SOAP message infosets as fast infoset documents
  - SOAP message infosets as ASN.1 SOAP messages

And in the future...

- Fast Infoset and with X.694?
- Fast Web Services with Fast Infoset and with X.694?

Further information

- For all things ASN.1 and XML related
  - First hit for google “ASN.1 XML”
- Open source Fast Infoset
  - First hit for google “Fast Infoset”

Fast Infoset & Fast Web Services

Paul.Sandoz@sun.com