

# *Enabling Information Sharing thru Common Services*

## **Efficient XML**

Presented To: Air Transportation Information  
Exchange Conference

Presented By: Aaron Braeckel

Date: Aug 31, 2011



Federal Aviation  
Administration

## **Air Transportation Information Exchange Conference - (featuring AIXM, WXXM and FIXM)**

August 30, 2011 - September 1, 2011  
NOAA Science Center & Auditorium  
Silver Spring, Maryland

# XML Benefits



Air Transportation Information  
Exchange Conference - (featuring  
AIXM, WXXM and FIXM)

Extensible

Human-readable

Self-describing

Hardware, software, platform-independent

Expressive data model (trees, graphs, etc.)

Validatable

Namespaces

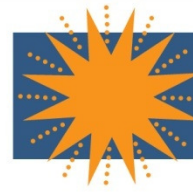
...



EUROCONTROL



Federal Aviation  
Administration



# XML Issues

## Processing Efficiency

- Encoding/decoding time
  - Latency
- Message throughput
- Battery life
  - mobile devices

## Compactness

- Bandwidth/transmission\*
  - Mobile devices
  - Next-generation aircraft
  - Latency
- Storage
- Archival

## Memory Use

Highly dependent on parsing technique (SAX, DOM, StAX)

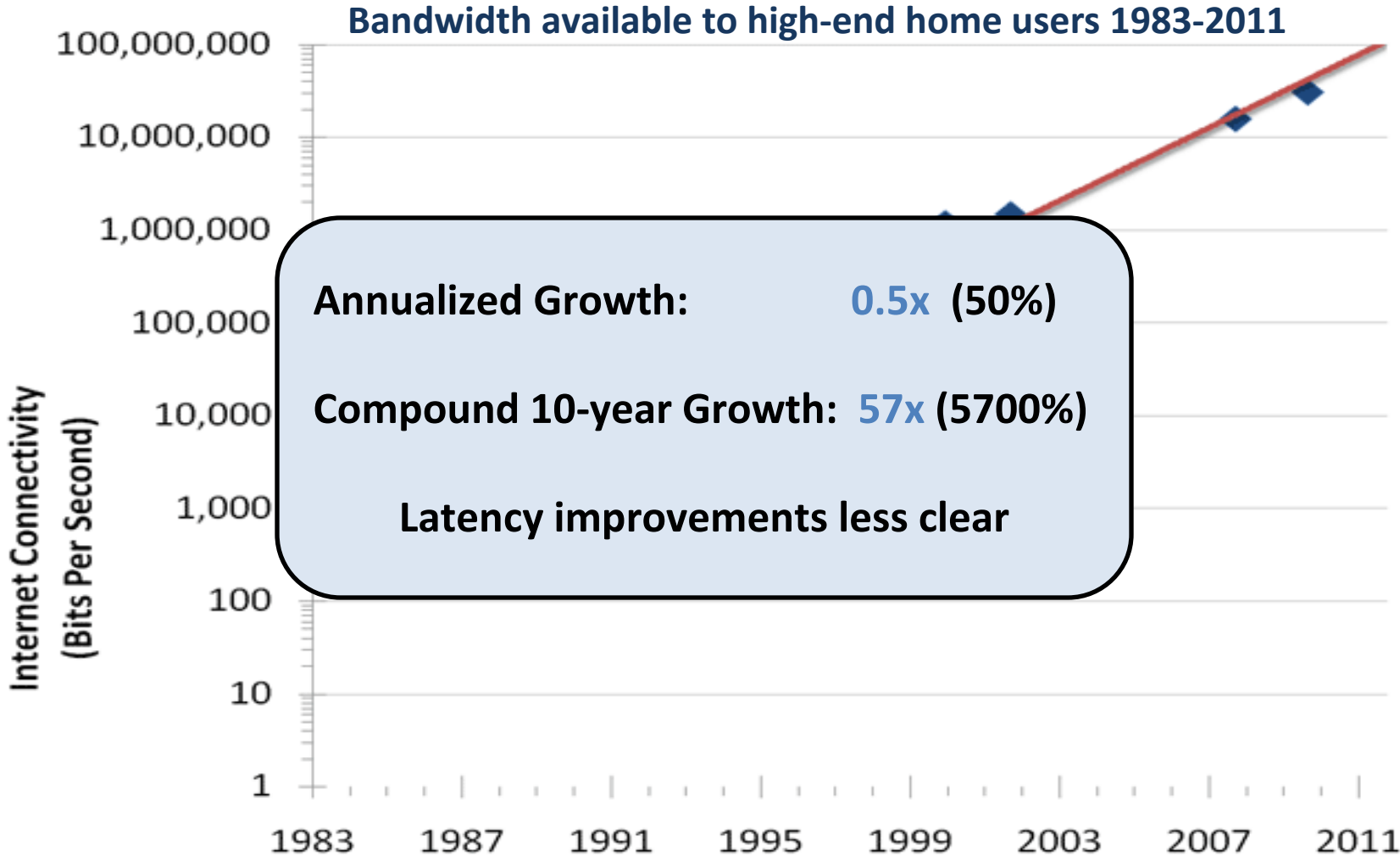
\*(non-trivial) recurring cost in some cases



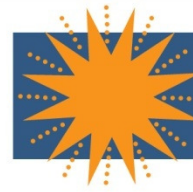
# Nielsen's Law



Air Transportation Information Exchange Conference - (featuring AIXM, WXXM and FIXM)



Federal Aviation Administration



# Case Studies

## Legacy Binary METAR:



Product	Legacy Size	Unformatted XML Size	Formatted XML Size
<b>METAR (24hrs)</b>	25468818	301227422 <b>(12x)</b>	409222080 <b>(16x)</b>
<b>TAF (24hrs)</b>	10434486	125747574 <b>(12x)</b>	116830236 <b>(11x)</b>

## Legacy DoD Product Comparison: <sup>1</sup>

10x, 17x, 22x, and 121x larger with XML

<sup>1</sup> Efficient XML - Taking Net-Centric Operations to the Edge. John Schneider



# Solution Classes



## General-purpose Compression - BZIP2, GZIP, etc.

- Additional processing requirements
- Still requires processing on uncompressed XML (size-only solution)

## XML Appliances/Hardware



- System deployment – often expensive, not omnipresent throughout any system
- Typically address limited processing requirements (XPath, etc.)

```
<?xml version="1.0" encoding="UTF-8"?>
<myXML>
  <Identifier>XML data</Identifier>
  <ProducedBy type="agency">
    National Center for Atmospheric Research
  </ProducedBy>
  <Data>
    ... binary contents here ...
  </Data>
</myXML>
```

## XML-wrapped Binary - base64, etc. XML for “metadata”, binary for data

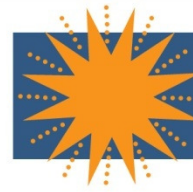
- Binary portions are “opaque”, no XML compatibility
- Base64 introduces additional data size (~30%)



## “Binary” XML Encodings - Fast Infoset, EXI, etc. Alternative (binary) encoding of the XML data model

- Human readability lost

# Binary XML



Air Transportation Information  
Exchange Conference - (featuring  
AIXM, WXXM and FIXM)

- Can be used at the lowest level of the XML stack
- Transparent to developers and systems
- Omnipresent (software)
- Lossless conversion
- Increased efficiency
- Open source/commercial libraries
- Supports all types of processing currently in use (DOM, SAX, etc.)



EUROCONTROL



Federal Aviation  
Administration



# Binary XML Solutions



Air Transportation Information  
Exchange Conference - (featuring  
AIXM, WXXM and FIXM)

Data Format	Standards Bodies	W3C EXI Characteristics	Notes
Fast Infoset	ITU-T ISO	Was not considered to satisfy: •Compactness •Generality	
EXI	W3C	Meets all characteristics	W3C Recommendation
BiM	ISO (MPEG WG)	Not Measured	
BXML	OGC	Not Measured	OGC Best Practice, not a standard
WBXML	Open Mobile Alliance W3C	Not Measured	Proposed W3C Standard (inactive?)

Other formats/approaches (such as XMill) not considered due to lack of standardization



Federal Aviation  
Administration

# Prior Work



Air Transportation Information  
Exchange Conference - (featuring  
AIXM, WXXM and FIXM)

## W3C XML Binary Characterization Working Group

- Collected requirements
- Collected use cases
- Identified essential requirements

**“Binary XML is needed”**

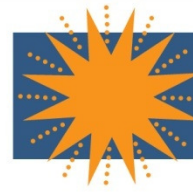
## W3C Efficient XML Interchange Working Group

- Performed measurements
- Made recommendations
- Produced a measurement framework
- Authored the EXI data format

**“it is possible to achieve substantial gains over  
XML... in a wide variety of use cases”**



Federal Aviation  
Administration



# W3C Binary Characterization Working Group

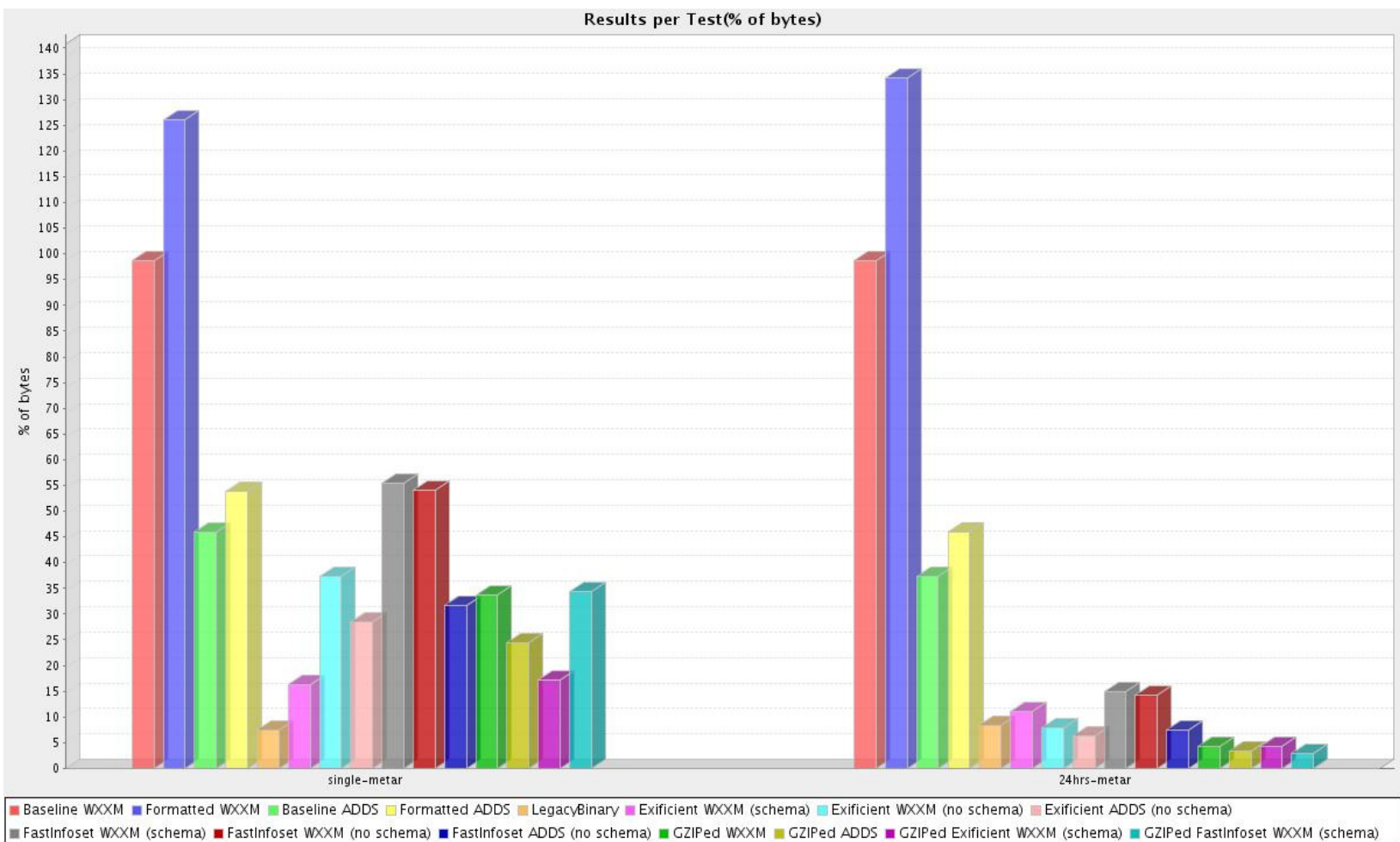
- Directly Readable & Writable
- Transport Independence
- Compactness
- Human Language Neutral
- Platform Neutrality
- Integratable into XML Stack
- Royalty Free
- Fragmentable
- Streamable
- Roundtrip Support
- Generality
- Schema Extensions and Deviations
- Format Version Identifier
- Content Type Management
- Self-Contained

# Efficient XML Report 1.0

## - METARs



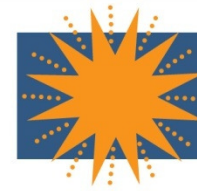
Air Transportation Information Exchange Conference - (featuring AIXM, WXXM and FIXM)



Federal Aviation Administration

# Efficient XML Report 1.0

## – all results, 24 hrs



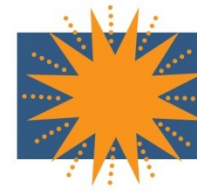
Air Transportation Information  
Exchange Conference - (featuring  
AIXM, WXXM and FIXM)

Data Format	Average Compaction (relative to baseline WXXM)
Formatted <b>WXXM</b>	<b>1.36</b>
Baseline <b>ADDS</b>	<b>0.38</b>
Exificient <b>WXXM</b> (with schema)	<b>0.11</b>
Exificient <b>WXXM</b> (without schema)	<b>0.08</b>
Exificient <b>ADDS</b> (without schema)	<b>0.06</b>
Sun's Fast Infoset <b>WXXM</b> (with schema)	<b>0.15</b>
Sun's Fast Infoset <b>WXXM</b> (without schema)	<b>0.14</b>
Sun's Fast Infoset <b>ADDS</b> (without schema)	<b>0.08</b>
GZIP <b>WXXM</b>	<b>0.04</b>
GZIP <b>ADDS</b>	<b>0.03</b>
GZIP Exificient <b>WXXM</b> (with schema)	<b>0.04</b>
GZIP Sun's Fast Infoset <b>WXXM</b> (without schema)	<b>0.03</b>
Legacy Binary (METARs and TAFs only)	<b>0.15</b>

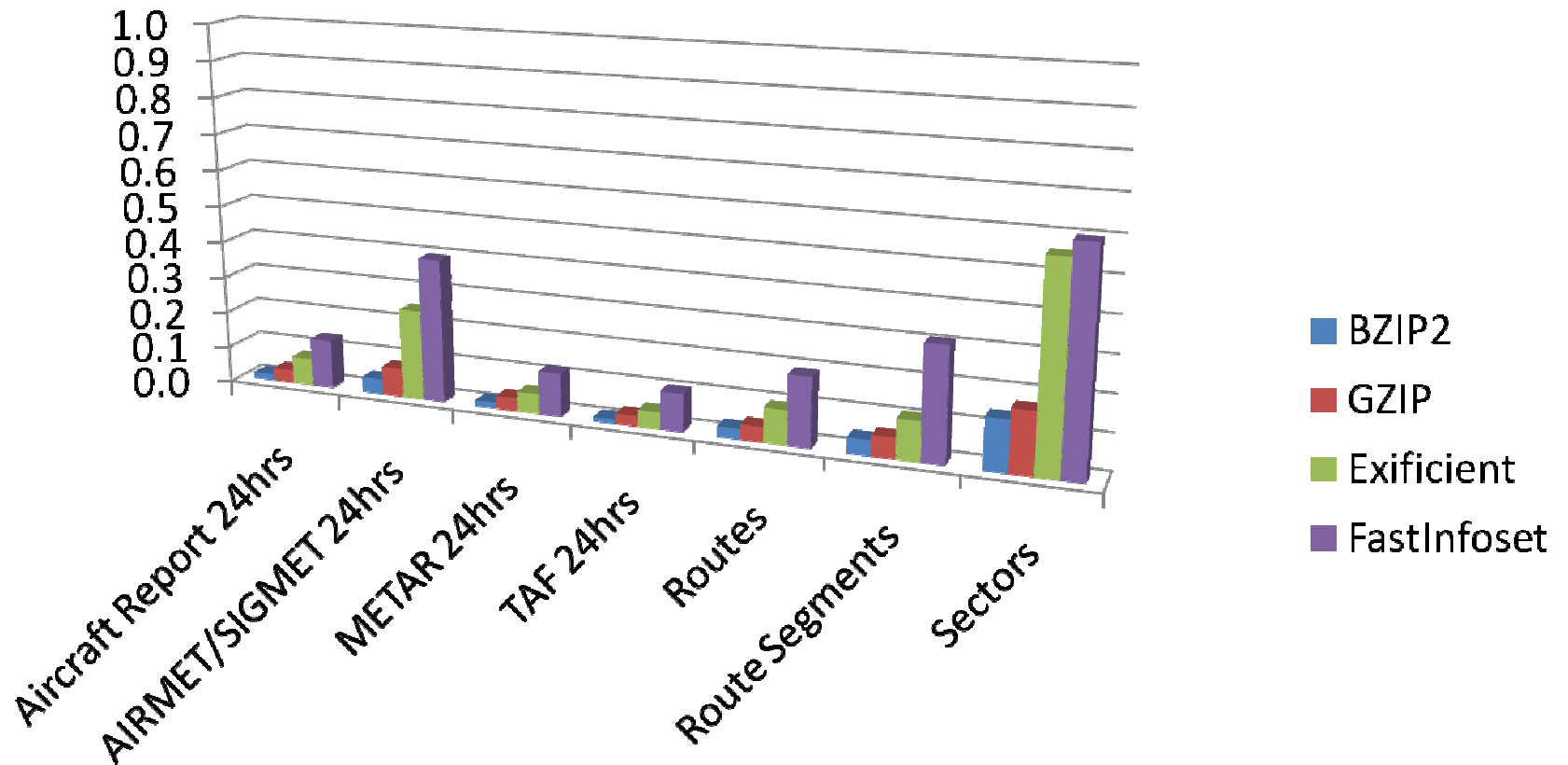


Federal Aviation  
Administration

# AIXM and WXXM



Air Transportation Information Exchange Conference - (featuring AIXM, WXXM and FIXM)



**Exifcient and FastInfoset were run schema-less**



Federal Aviation Administration

# Bandwidth and Latency



Air Transportation Information  
Exchange Conference - (featuring  
AIXM, WXXM and FIXM)

24 hrs of (XML) METARs	426Mb
GZIPed	15.3Mb

ideal bits per second = TCP window size / Round trip latency in seconds

Assume **30ms** ping time and TCP window size of **64k**:

524288 bits / 0.030 seconds = 17476266 bits per second

24 hrs of (XML) METARs	25.5 sec
GZIPed	0.9 sec

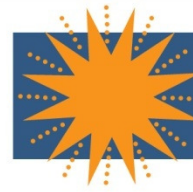


EUROCONTROL



Federal Aviation  
Administration

# Efficient XML Report 2.0



Air Transportation Information  
Exchange Conference - (featuring  
AIXM, WXXM and FIXM)

## Processing Analysis

- Encode
- Decode

## Network

- Two machines
- Fat/Skinny pipes
- Transactions per second

## Expanded Compaction Analysis

- Commercial Libraries
- With/without schema



Federal Aviation  
Administration



# Conclusions



Air Transportation Information  
Exchange Conference - (featuring  
AIXM, WXXM and FIXM)

## XML Has Performance Challenges

- ...along with its benefits

## Favor SAX/StAX over DOM

- “Streaming” memory use

## Allow compression

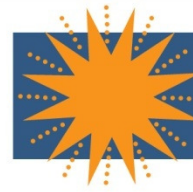
- “Worst” technique offers 60% improvement to compactness
- Non-trivial latency reductions

## Try Binary XML/EXI

- Compatible with entire XML stack (XSLT, Xpath, etc.)
- Favorable compactness
- Improved processing likely
- Commercial and open source libraries
- Official W3C Recommendation, meets all desirable characteristics



Federal Aviation  
Administration



## More Information / Contacts

- <http://www.w3.org/XML/EXI/>
- [https://wiki.ucar.edu/download/attachments/23364539/EfficientXMLReport\\_v1+0.doc?version=1&modificationDate=1271765646000](https://wiki.ucar.edu/download/attachments/23364539/EfficientXMLReport_v1+0.doc?version=1&modificationDate=1271765646000)