



Handling Binary Content and Large Documents in Oracle SOA Suite 11g and OSB



Ramkumar Menon
Oracle Corporation



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 remains at the sole discretion of Oracle.



Program Agenda

- Overview
- Common usage patterns
- Data flavors and Processing patterns
- Relevant Product Configurations
- Management and Monitoring



What does “Handling” mean?



- “Gracefully” process payloads of larger than "usual" message sizes, volumes and concurrencies within the current server capacity and resources.
- Servers are still responsive and processing other requests.

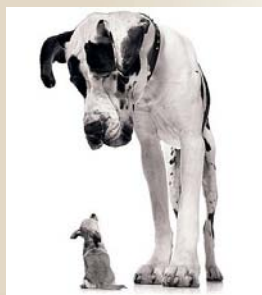


How large is “Large”?

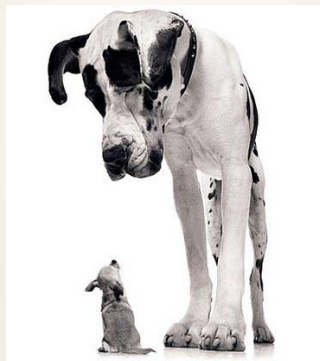
1 MB



10 MB



500 MB



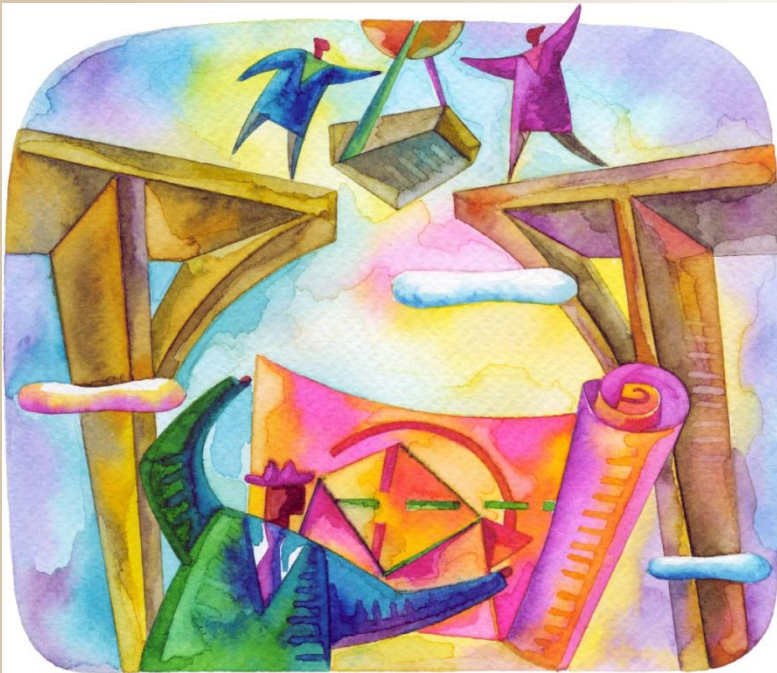
2 GB



There is no hard and fast rule on what is considered a “large payload”!



Planning for Large Message Processing



- Sizing and Capacity Planning
- Application Design and Architecture
- Performance Tuning Cycles
- Configuration

Program Agenda

- Overview
- **Common usage patterns**
- Data flavors and Processing patterns
- Relevant Product Configurations
- Management and Monitoring

Common use-cases

- Moving GIS data from a remote FTP Server to a local file system
- Receiving Purchase Orders via SOAP/HTTP, with additional documents and artifacts coming in as attachments.
- Receiving a Bulk submission of Automobile orders (1000s in one)
- Moving Media between storage and transcoding FTP Servers.



Program Agenda

- Overview
- Common usage patterns
- **Data sources, flavors and processing patterns**
- Relevant Product Configurations
- Management and Monitoring



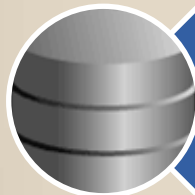
Data Sources



Over HTTP



From Local and Remote File servers



Database repositories



Messaging servers

Data Flavors

- Binary Data
- Complex XML Document
- XML Documents with repeating sections
- XML + Binary
- Native delimited or positional text file



Data Usage Patterns

- Pass through processing
- Partial Data Extraction (For XML and native Messages)
 - Query for Content based Routing
 - Smaller target documents populated from few source nodes
- Full Message Transformation (for XML and native messages)



Use-cases

- Large XML Messages (SOAP, XML, File, DB)
 - Complex Non-Repeating
 - Repeating, Batch-able
- Large Native/Delimited Messages (File System)
- Large Binary Messages
 - File System
 - Arriving on SOAP over HTTP (XML + Binary)
 - Data read from a DB table/AQ queue

Case 1 – Large XML – Complex Non-Repeating

- Leverage streaming features of File/FTP/DB/AQ adapters
- Leverage content streaming feature of OSB
- Read as Attachment if pass-through for file/ftp
- Streaming is optimal for small queries, but ineffective for full XML Transformations.
- If you are using XSLT, ensure you use “streamResultToTempFile” while generating large Documents.
- Appropriately scope all BPEL Variables **#Kscope**



Case 1 – Large XML – Complex Non-Repeating - Screenshots

The screenshot displays the Oracle Service Bus configuration interface. At the top, a code snippet shows a WSDL definition: `<wsdl:definitions name="Read" targetNamespace="http://xmlns.oracle.com/ohb1/adapter/file/Consolidat...`

The main configuration window is titled "Create a Proxy Service (test/)" and includes a "General Configuration" section. The "Service Name" is set to "StreamingProxy". Under "Service Type", the "Any XML Service" option is selected. The "Content Streaming" section is highlighted with a black circle, showing "Enabled" checked, with "Memory Buffer" selected as the buffer type.

Below the main configuration is the "Static Routing" section. It shows a routing rule with a "Filter Expression" field, a "Writer::Write" operation, and a "Sequential" dropdown. The "Assign Values" dropdown is also highlighted with a black circle.

On the left side, a partial XML snippet is visible: `<adapter- <connec <endpoi <acti <pr <pr <pr </act </endpo </adapter </endpoi </adapter-`

On the right side, a "Writer" component is shown with the operation "Write".



Terms to understand

- Streaming
- Scalable DOM
- Debatching
- Chunked Read



Scalable DOM

- Feature of XDK 11g – Scalable and Pluggable DOM
- Pluggable DOM Support
 - DOM API split from data (plug-in/internal)
- Lazy Materialization
- Configurable DOM Settings
 - PARTIAL_DOM
 - ACCESS_MODE
 - STREAMING, FORWARD_READ



Case 2 – Large XML –Repeating Structure

- Streaming transformations now available from PS3 for processing large XML Documents.
- Debatching
- Chunked Read Interactions



Case 3 – Large Native Data – Repeating Structure

- Streaming and Non-Streaming Translator Xpath functions available.
- Chunked Read for Native Data

Case 4 –Large Binary Files – File System

- “Read as Attachment” option on inbound File/FTP adapter.
- Mediator can pass through these attachments without materializing the file. BPEL will dehydrate attachments.
- “Use streaming” option enables partial to more complex data access and transformations within the large file at reduced cost by virtue of Scalable DOM feature.



Processing the Large Binary Files

- Xpath Functions
- Pass-through
- Retrieving Attachment content
- Sending attachment content to other Composites



Case 5 - Large Binary Attachments over SOAP

- SwA - SOAP with Attachments
 - Useful for larger-sized payloads
 - Does not materialize the message like MTOM does
 - SOA Mediator can pass through attachments
 - SOA 11g supports streaming of attachments on inbound and outbound. Streaming of attachments is not supported for MTOM.
 - Not supported for use with Spring component today.



Case 6 - Large Binary Attachments over SOAP

- MTOM
 - Useful for mid to large sized payloads
 - Defined as a simple policy assertion on the service or reference.
 - Available only on WS Bindings.



Notes

- Usage of Opaque Translation
- Random access updates to XML Documents on Streaming transformations.



Program Agenda

- Overview
- Common usage patterns
- Data flavors and Processing patterns
- **Relevant Product Configurations**
- Management and Monitoring



Server Configurations

- Dedicated Domain?
- Audit Levels
- JVM
 - Heap 32 v/s 64 bit
 - Garbage Collection
- Max Message Size
- Transaction Timeouts
- Stuck Thread Max Time
- Max Test Wait Seconds (For dehyd store)



Program Agenda

- Overview
- Common usage patterns
- Data flavors and Processing patterns
- Relevant Product Configurations
- **Management and Monitoring**



Administrative Tasks

- Tablespace and Data files for large Attachments in Dehydration Store
- Data Retention and Purge Policy
- Configure Rejection of very large messages



Q & A