Web Services Reliability Options

A Comparison of Web Services Reliable Messaging Specifications

OASIS WSRM TC

May 7, 2004

Acknowledgements

- We thank the OASIS board of directors for this opportunity to respond to the IBM presentation made during the OASIS Reliable Infrastructures for XML Symposium messaging session
- We thank everyone who have provided comments or otherwise made their mark on the OASIS WS-Reliability Specification

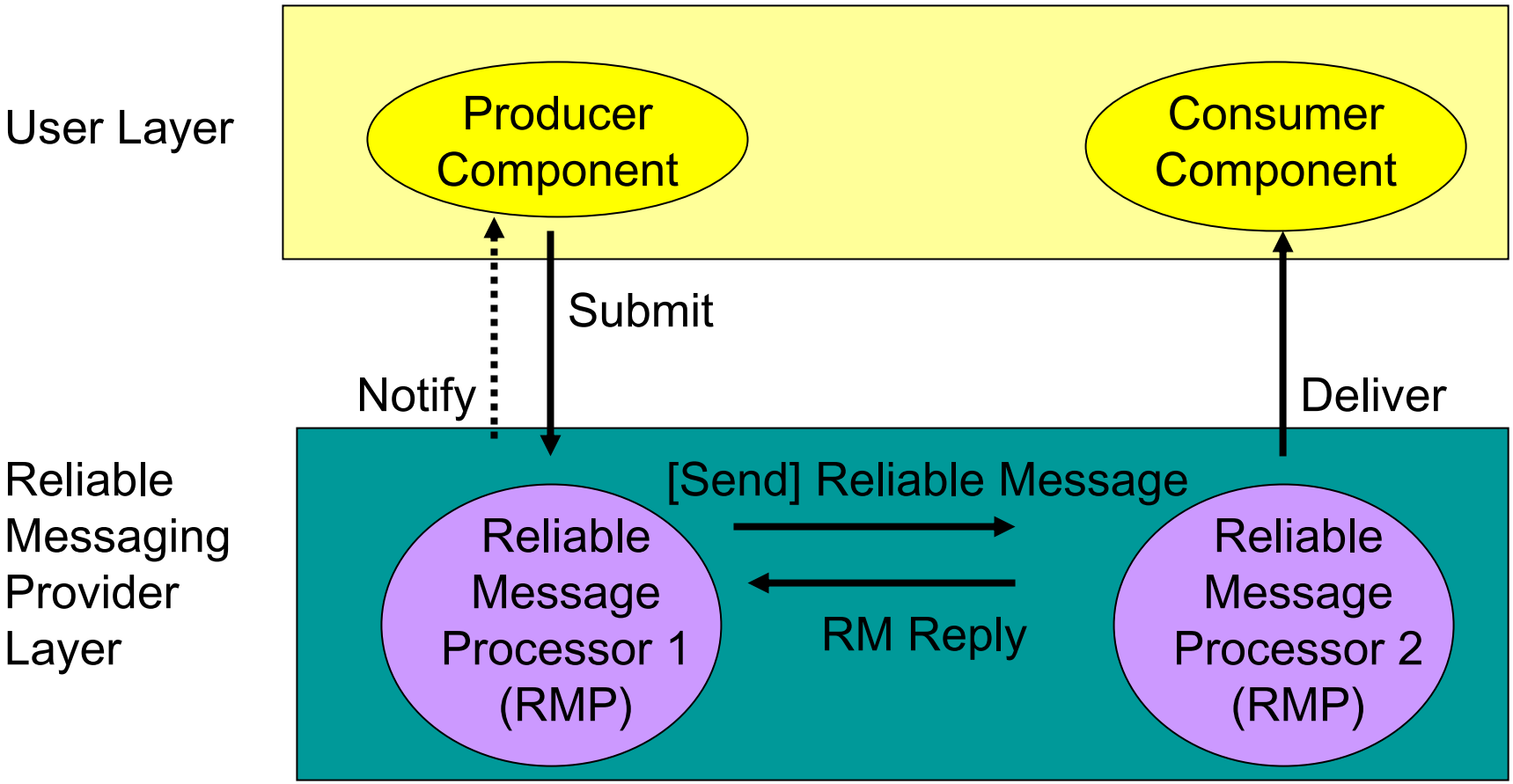
Background: Web Services Reliable Message Delivery Options

- Currently there are two choices
- Open Standards:
  - OASIS WSRM TC developed WS-Reliability (“WS-R”)
  - First published 9 January 2003
  - TC publicly announced 13 February 2003
- Proprietary:
  - IBM/BEA/Microsoft/TIBCO authored WS-ReliableMessaging (“WS-RM”)
  - First published 13 March 2003

Background: Motivations for a Reliable Transport

- Underlying communications mechanism variety
  - Traditional (TCP/IP)
  - High latency variance
  - Wireless telephony
  - Other / “non traditional” mechanisms
- Potential for message loss, and message re-ordering
- Lower level TCP characteristics do not adequately protect large multi-message Web Services business interactions

Background: Messaging Model



Background: Enabling Mechanisms

- Guaranteed delivery
  - Transfer of responsibility is unambiguous from sending RMP to receiving RMP
- Duplicate elimination
  - Transmission integrity is not affected by loss of acknowledgement or accidental duplication
- Message re-ordering
  - Messages are delivered in the order sent
- Grouping
  - Related messages are collected into a coherent unit

Comparison: WS-R Supported Use Cases

- Request-Response (business message exchange)
- One way messaging (business message)
- Polled receiver (firewall or device constraints)
- Long running group (logging model)
- Lightweight devices (cell phone and smaller)

- All are supported by WS-R with a common protocol respectful of implementation choices and resources
- WS-RM does not support polling and we believe its support for WSDL Request-Response to be underspecified
- WS-RM cannot operate with producers protected by a firewall.

Comparison: Benefits of WS-R over WS-RM: Group Management

- WS-R does not require a message exchange for group establishment or termination
  - Benefit: All group establishment is implicit and low overhead
- WS-RM supports an optional sequence establishment message exchange
  - When used: adds latency, and dependency on other protocol messages that may not be reliable themselves
  - To prevent a late arriving duplicate message from causing a new sequence to be automatically started, the sender must either use createSequence explicitly, or must send the expiry time with every message in that sequence
- In WS-RM, the choice seems to be either additional latency, or specification of expiry time

Comparison: Benefits of WS-R over WS-RM: negative acknowledgement

- WS-R has no “nack” (negative acknowledgement)
  - Comment: The feature is an optimization that assumes receiver properly distinguishes the difference between a delayed message and a missing message. Correct implementation requires Extra Special Programming
  - Hazard: If overused, especially in conjunction with retry, will promote network congestion failures
  - Benefit: WS-R will not cause congested network failure on missing message recovery

Comparison: WS-R is less Dependant on other Specifications

- WS-R does not rely on proprietary policy and addressing protocols to configure mandatory sender and receiver options
  - Benefits:
    - WS-R is self contained
    - WS-R receiver does not need to be pre-configured prior to message exchange
    - WS-R requires no pre-requisite proprietary protocols

Specific responses to IBM's assertions

- Each of the following slides responds to an assertion made during the IBM Presentation
- WS-R has been open for public comment, and IBM has not submitted any comments to the TC
- IBM as were the other authors of WS-RM were invited to participate in the OASIS TC and are still welcome should they desire constructive participation

IBM's Assertion: Two Schemas and namespaces are unnecessary

- Good point
- Initially two schemas were intended to accommodate SOAP 1.1 to 1.2 differences
- Since SOAP 1.2 was in process at the start and since SOAP 1.2 has been final since June 2003, it is clear that two schemas are unnecessary
- The TC has agreed to define one schema for use with both SOAP 1.1 and SOAP 1.2

IBM's Assertion: Why are Soap Faults not used for RM-Fault?

- SOAP faults are used when the error cannot be hidden from the user layer
- SOAP fault model does not provide for batching of faults and acknowledgement indications
- Although possible to send a SOAP fault in an HTTP request, it is unusual to send a SOAP Fault in a request
- Not mapping RM-Fault to SOAP fault allows piggy-backing of RM-Faults on business messages

IBM's Assertion: Holding an Ack until application delivery causes delay

- Ack on receipt is not reliable and gives the sender false assurance due to gap between receipt and delivery
- Example of this failure mode is a power failure between ack and persistence or ultimate message usage
- WS-R defines delivery as the point where the receiving RMP has accepted responsibility for the message and potentially made it available to the consumer
- The TC will clarify the text

IBM's Assertion: Unclear if WS-R composes with WS-Addressing or WS-MessageDelivery.

- TC desires composability with many other mechanisms, however the TC will not specify a proprietary mechanism nor will it specify one mechanism at the exclusion of others
- The TC will review the spec for extensibility in this regard

IBM’s Assertion: Persistence model precludes use on devices lacking non-volatile storage

- Both WS-R and WS-RM require equivalent levels of state storage during operation
- Guaranteed delivery requires RMP functionality
- Non-volatile queues can be used to enhance reliability
- WS-R does not require non-volatile storage

IBM's Assertion: Mandatory expiry time requires synchronization of clocks

- Expiry is not a tight tolerance parameter
- The producer determines expiry time to meet business need, system configuration, and network conditions, and should be set large enough to allow for expected clock skews
- Resource reclamation is thus based on producer need or system configuration
- Mandatory expiry time significantly simplifies the protocol.

IBM's Assertion: WS-R Spec does not state that receiver must ack all delivered messages with each ack indication

- WS-R protocol sends RM-replies (acks or RM-fault indications) only as required, there is no requirement to provide entire group history with each rm-response.
- WS-R Response reply pattern places RM-Reply in Soap Response for the single message in the request.
- WS-R callback reply pattern includes RM-Replies for all messages not already acknowledged in each callback
- Acknowledgement and fault indications can be requested for all messages sent in a group by sending a WS-R poll request including that groupID.

IBM's Assertion: Unnecessary implementation details in spec

- WS-R does not contain details of any particular implementation, but does provide hints and guidance
- A description of bits-on-the-wire alone does not adequately describe end point behavior; procedural description improves clarity
- Many correspondents have expressed appreciation for such guidance
- The TC will clearly label this useful implementation guidance from "normative" specification any may publish it as a separate implementation guide

IBM's Assertion: WS-R is a complex spec with many occurrences of the word "if"

- Most "ifs" in WS-R are used to describe behavior not alternative implementations
- The use of the word "if" does not indicate complexity as there are many alternative expressions
- At some point it may be useful to compare state diagrams as a more meaningful test

IBM's Assertion: WS-R has too big a "THUNK" factor

- This is a silly issue. The spec needs to be big enough to be clear and complete
- THUNK units relate to weight, not completeness, complexity or clarity.
- Including the page count of the referenced specifications not common to WS-R grows the WS-RM page count from 40 pages (IBM version) to over 117. vs. 68 pages in WS-R v0.996
- WS-R does not use 8 point type ;-)

Conclusion

- We thank all participants for their input and efforts in the creation of WS-R
- The OASIS WSRM TC is finalizing the WS-R spec taking all comments into account
- Please direct comments about the WS-R specification or this presentation to wsrm-comment@lists.oasis-open.org