IHE Work Item Proposal (Detailed)

# Proposed Work Item: XCA Deferred Response Option

Proposal Editor: Joe Lamy

Work Item Editor: Joe Lamy

Date: 15 November 2018

Version: 1.1

Domain: IT Infrastructure

**Summary**

Query and retrieval of documents across communities may in some cases require much more time than synchronous web services allow. We identify two such cases. Further, the existing asynchronous mechanisms available in XCA may not be a good match for some systems.

The XCPD profile has a Deferred Response Option, while the XCA profile does not. When ITI was designing the deferred mechanism for XCPD, it proposed and partially designed a draft mechanism for XCA as well. The only reason it was not originally completed was that there was not yet a pressing need. We now have one, so we are proposing to add a Deferred Response option to XCA.

The US Social Security Administration and a partnering Release Of Information vendor have indicated strong interest in solving this problem, and we believe others will have similar interest.

IHE ITI is an appropriate venue to solve this, so that the solution may be standardized and not left to individual implementers.

# The Problem

We identify two long-latency XCA use cases:

* There are a huge number of legacy paper clinical documents for which there is currently no compelling business case to make available via IHE Document Sharing profiles. The problem is the chicken and the egg – since most will never be requested electronically, it doesn't make business sense to scan, parse and register all documents, and for those that will be requested, the latency to do so is too long for synchronous web services.
* Responding Gateways that utilize the on-demand mechanism to aggregate data from multiple sources may take significant time to do so, causing HTTP timeouts.

The existing asynchronous mechanisms available for XCA may not be optimal for some implementers:

* WS-Addressing-style async is handled within the web stack and has known limitations as typically implemented: it is non-blocking for the requester but not the responder; it doesn’t handle long-latencies (e.g. days) well, and does not allow management of acknowledgements at the application layer. See the US CONNECT team [analysis](https://connectopensource.atlassian.net/wiki/spaces/CONNECTWIKI/pages/8585329/Asynchronous+Messaging+Engineering+Analysis), as well as the original IHE [white paper](ftp://ftp.ihe.net/IT_Infrastructure/iheitiyr8-2010-2011/Technical_Cmte/Profile_Work/DeferredMsging/IHE_ITI_WhitePaper_Async.0810.doc) that justified the deferred mechanism.
* AS4-style async is also handled within the web stack, and with its significantly distinct usage of the SOAP header is better suited for an environment in which all web services operate in this way.

These documents are often made available through other, slower, means by Release Of Information vendors. A Deferred XCA option would remove the technical barrier and allow a compelling business case to be made to use IHE Document Sharing profiles.

# Use Cases

Case 1: A document source has ownership / access to a very large number of paper-based clinical documents, and wishes to provide these documents electronically using IHE Document Sharing profiles. However, it does not know a priori which of these documents will be requested, and it is not cost effective to proactively scan, parse and register all documents. As each step in the workflow potentially involves this manual process, which could take hours to days, the system may need to make use of deferred responses for XCPD, XCA Query, and XCA Retrieve:

* Initiating Gateway sends deferred XCPD request
* User at responding system searches for patient records, determines matches, and returns response
* Initiating Gateway sends deferred XCA Query request
* User at responding system searches for appropriate clinical documents, scans, parses or otherwise generates document metadata, registers document entries, potentially stores documents, and returns response
* Initiating Gateway sends deferred XCA Retrieve request
* User at responding system performs any remaining tasks to make documents available, and returns response

Case 2: A Responding Gateway uses an on-demand document entry to generate a comprehensive longitudinal record (e.g. a CCD) for a patient from a number of sources. In some cases, the time needed to generate this document runs into the minutes, causing Initiating Gateways to time out. This system would make use of deferred responses for XCA Retrieve:

* Initiating Gateway sends deferred XCA Retrieve request for on-demand entry
* Responding Gateway compiles the on-demand document from various sources and sends response

# Standards & Systems

The proposal expects to implement changes to the IHE XCA profile. The following is a list of standards referenced for this proposal:

* OASIS Standard - ebXML Registry Services and Protocols Version 3.0 (2 May, 2005)
* IHE IT Infrastructure Technical Framework, Volume 1 (ITI TF-1): Integration Profiles (Rev. 13.0 Final Text - 2016-09-09)
* IHE IT Infrastructure Technical Framework, Volume 2a (ITI TF-2a): Transactions Part A (Rev. 13.0 Final Text - 2016-09-09)
* IHE IT Infrastructure Technical Framework, Volume 2b (ITI TF-2b): Transactions Part B (Rev. 15.0 Final Text – 2018-07-24)

# Technical Approach

**New actors**

No new IHE actors are currently anticipated as the result of this work item.

**Existing actors**

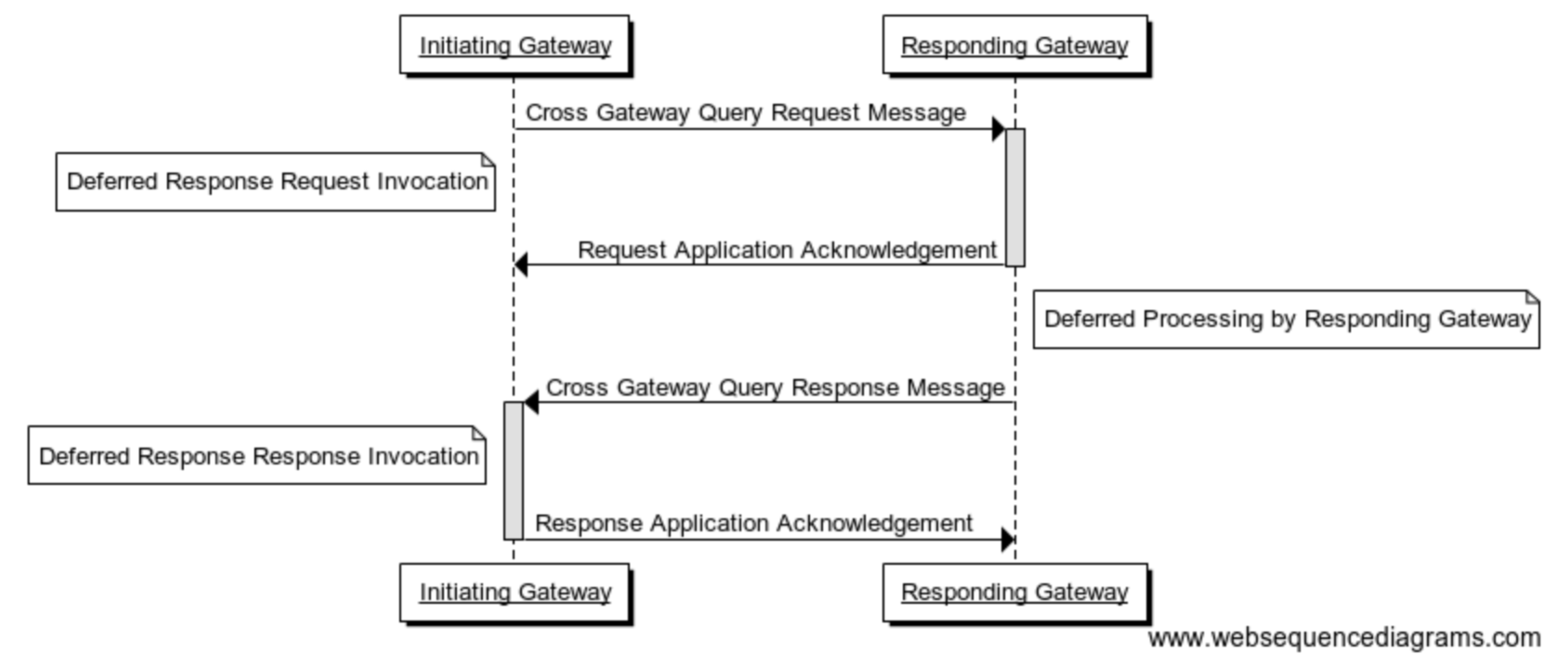
IHE actors involved in this change will be the XCA Initiating Gateway and XCA Responding Gateway. Depending on grouping requirements, most likely the Document Consumer and On-Demand Document Source, and potentially the Document Registry and Document Repository may be affected.

**New transactions (standards used)**

The ITI-38 and ITI-39 transactions would receive a new “Deferred Response option”.

The ITI-18 and ITI-43 transactions may receive a new “Deferred Response option”.

The Deferred Response option would use two 2-way SOAP messages for the request and response, with additional application-level acknowledgements. XCA Query is shown below as an example:



Keeping with the patterns established by the XCPD Deferred Response option, we plan to do the following:

* Define separate WS-Addressing Actions for the deferred request and response
* Use WS-Addressing RelatesTo on the response to correlate it to the original request

There are a handful of technical questions that will need to be answered:

* What format should the acknowledgement messages be? Prior IHE work on this suggested ebXML RegistryResponse for both Query and Retrieve.
  + If ebXML RegistryResponse is used, what status values should be used? ebXML defines the value Unavailable for this case, while IHE tentatively defined the values RequestAccepted and ResponseAccepted.
* Will we need to explicitly indicate deferred processing in the body of the request, or will the Action be sufficient?
  + XCPD includes this explicitly in the responsePriorityCode field.
  + For the XCA Query, the ebXML AdhocQueryRequest has no analog, but it is extensible with Slots.
  + The XCA Retrieve message is defined by IHE, so it could be extended.
* How will the Responding Gateway know what endpoint to send the response to?
  + XCPD includes this explicitly in the respondTo field.
  + For the XCA Query, the ebXML AdhocQueryRequest has no analog, but it is extensible with Slots.
  + The XCA Retrieve message is defined by IHE, so it could be extended.
  + Should the Initiating Gateway specify an endpoint or a HCID?
    - Using a HCID that the Responding Gateway resolves via a directory might help with some dynamic routing cases.
  + As an alternate method, if the Initiating Gateway is grouped with an XUA X-Service User, its HomeCommunityID is included in the SOAP header, and the Responding Gateway could use that to look up a deferred response endpoint. This technique was used by the US eHealth Exchange for deferred XCPD, as use of the respondTo element was not yet defined by IHE.
    - Tech Committee did not think this was appropriate.
  + As an alternate method, we could define a SOAP header block for the response endpoint similar to WS-Addressing ReplyTo. NOTE: This would be a problem if combining deferred and AS4 mechanisms.
* What error cases will exist, and how will we handle them?
  + Sending a deferred request to a Responding Gateway that doesn’t support it
  + Existing error cases will be unchanged

**Impact on existing integration profiles**

The XCA profile would receive a new “Deferred Response option”.

The XDS.b profile may receive a new “Deferred Response option”.

**New integration profiles needed**

No new profiles are needed.

**Breakdown of tasks that need to be accomplished**

* Determine which technical solution to go with: DSUB, Deferred, etc.
* Specify a “Deferred Response option” for the XCA and potentially XDS.b profiles:
  + Determine if a “Deferred Response option” is necessary for XDS.b.
  + Define WS-Addressing requirements for Action and RelatesTo.
  + Determine format of the acknowledgement messages.
  + Determine if any changes to Vol 2x Appendix V are needed (using the 2018 revisions).
  + Enumerate error cases and define error messages and/or status codes.
* Produce complete draft for final review.
* XCA Deferred Response option prepared for public comments.
* XCA Deferred Response option prepared for trial implementation.

# Risks

For an ROI vendor to successfully field this profile, there may be additional considerations. Specifically, an initial sweep of patient demographic data may be practical, in order to prepopulate the MPI (Master Patient Index) and perhaps a centralized RLS (Record Locator Service) as well.

# Open Issues

* Will we include in scope the possibility of multiple responses to a single request, each response coming asynchronously?
  + Currently this is considered out of scope.
* Will we need to add a Deferred option to XDS.b?
  + As an analog, Vol 1 18.2.2: if Initiating Gateway supports async and XDS Affinity Domain, must support async on ITI-18 and ITI-43.
  + It wouldn't work just to have the Document Consumer use async on ITI-18 and ITI-43 while the Initiating Gateway uses deferred on ITI-38 and ITI-39, as the longer deferred timelines could break async.
* It may not be needed to justify this work, but it would be good to know if typical implementations of the AS4 asynchronous mechanism support latencies of days.
  + Closed Issue AS4-5 decided to “keep the Deferred Option as currently specified as it is an application (versus transport-level) mode of operation, longer term delays (hours and days) compared to the AS4 Asynchronous push and push typical delays (minutes).”

# Effort Estimates

<The technical committee will use this area to record details of the effort estimation.>