Comments from Tom Scavo

Comments on the Final Report of the Per-Entity Metadata Working Group

> *Availability of the MDQ service should be 99.99% to 99.999%. [1. Executive Summary]*

This should read: "Availability of the MDQ service should be at least 99.99%."

> *Weaknesses in XML digital signature implementations have been found in the past, however, and will plausibly be discovered in the future. Malicious actors could exploit such weaknesses and prepare a rogue file to distribute to their target. [5.4 Security Related Risks]*

I don't know what "weaknesses in XML digital signature implementations have been found in the past" but in any case this is independent of TLS on the metadata query server. These two topics require separate paragraphs at least.

> *Another possible attack is the substitution of old metadata that is no longer current, but still within its validity period, for the current metadata. The use of MDQ increases the frequency of metadata requests and thus potentially enlarges the opportunity for such attacks. [5.4 Security Related Risks]*

The "frequency of metadata requests" is not the only reason to deploy TLS on the metadata query server. A query occurs in-band, just-in-time, which means an attacker can induce a query on demand. That raises the bar for per-entity metadata.

> *It is assumed that a common technical and process infrastructure will be used for both per-entity and aggregate distribution of metadata.[5.5 Unavailability of the Publishing Infrastructure]*

I don't think that's true. In particular, the current infrastructure to distribute aggregate metadata is insufficient for per-entity metadata distribution.

> *The architecture for the MDQ service will share much of the infrastructure that already exists to produce the aggregates. [6. MDQ Service Architecture]*

Same here. I don't think is true.

> *The working group's consensus was that the MDQ service must provide very high availability to the federation's IdPs and SPs, on the order of 99.99% to 99.999% availability. [6.2. Adding Per-Entity Metadata to the Infrastructure]*

This should read "at least 99.99% availability."

> *In order to achieve 99.99% to 99.999% availability (4.3 to 0.43 minutes of downtime per month) [6.2. Adding Per-Entity Metadata to the Infrastructure]*

This should read "achieve at least 99.99% availability (at most 4.3 minutes of downtime per month)."

> *The security of the MDQ service must, as much as possible, be equivalent to existing aggregate service. As noted above, the nature of MDQ may increase the risk of a metadata consumer receiving out of date information that is still within its validity period, due to a man-in-the-middle attack between the distribution layer and the consumer; this should be mitigated through the use of TLS. [7.1 Security]*

InCommon does not serve aggregate metadata over TLS. I can't really tell if TLS on the metadata query server is a requirement or not.

Personally, I think TLS on the MDQ server should be a strong requirement since: 1) it prevents man-in-the-middle attacks, and 2) it allows AD FS to leverage per-entity metadata.

> *The InCommon MDQ service must realize a monthly service uptime percentage of 99.99% to 99.999%. [7.2 Availability]*

This should read "uptime percentage of at least 99.99%."

> *Any MDQ client operated by an InCommon Participant must find the service available and responding normally to MDQ queries. Any time a client queries the service and the service does not respond normally due to any issue with the service delivery infrastructure is known as an outage. [7.2 Availability]*

What does "respond normally" mean?

> *The working group does not recommend any changes to the processes and systems that produce the metadata that is distributed both as aggregates and via the MDQ service. If at some time, however, InCommon were to decide to leverage per-entity distribution to institute more frequent, or real-time, publishing of metadata updates, then the availability of the publishing infrastructure should be addressed in light of such new service commitments. [7.4 Metadata Production]*

The following phrase doesn't make any sense to me: "leverage per-entity distribution to institute more frequent, or real-time, publishing of metadata updates." I suspect there's some misunderstanding about the metadata production process. We should probably discuss this on a call if the opportunity presents itself.

> *A decision to publish metadata more frequently would require reconsideration of the metadata validity and caching intervals. [7.4 Metadata Production]*

I don't see the connection between metadata production and metadata validity.

I think metadata validity should be discussed independent of metadata production. As you know, the current validity window (`validUntil`) is two weeks. At least one WG participant called for a reduction in the validity window, on the order of one week if I recall. Even though this would be very difficult for Ops to realize, it should probably be called out in the report.

The caching interval (`cacheDuration`) is currently undefined in aggregate metadata. Is the WG recommending that this be included in per-entity metadata, and if so, what should its value be?

Perhaps there should be a separate section on "Metadata Validity and Caching." See below for some possible content.

> *[8.2 Medium Term (2–12 months)]*

TAC should commission a new WG to address the discovery issue.

> *Per-entity metadata distribution is in production during this period. [8.3 Long Term (12--24 months)]*

If that's true, then there needs to be a proof of concept in the medium term.

> *A solution for discovery that is analogous to existing discovery approaches is identified and deployed. [8.3 Long Term (12--24 months)]*

That depends on the findings of an independent WG, which is why TAC needs to commission that WG asap.

> *InCommon develops a plan to retire aggregate distribution in the 36-48 month time frame. [8.4 Longer Term (24+ months)]*

I suggest the quoted sentence be deleted. We don't need or want such a plan since we have no intention of phasing out the aggregate, at least not for the foreseeable future.

Metadata Validity and Caching

Every metadata aggregate published by InCommon has a two-week validity window, that is, the `validUntil` XML attribute on the `<md:EntitiesDescriptor>` element is set to 14 days in the future.

OTOH, there is no `cacheDuration` XML attribute on InCommon metadata. Deployers are advised to attempt a metadata refresh every hour but this is intended to be a software configuration, not a metadata-driven behavior.

This fact was discussed recently on a call of the Per-Entity Metadata WG. I believe the group concluded that the `validUntil` XML attribute on per-entity metadata could remain the same as it on the aggregate but it was recommended by the group that ultimately the value should be reduced to one week across the board, on both aggregates and entities. (It is thought that four days might be optimal.)

To my knowledge, the WG has not recommended (one way or the other) that a `cacheDuration` XML attribute be added to per-entity metadata.