Some Notes on Metadata Interchange

Ian A. Young

V3, 12 October 2008

1. Scope

These notes describe my position on the issue of inter-federation through metadata interchange between SAML federations. I try and lay out some terminology and options, but in addition there's a lot of weight given to the solutions I think are most beneficial at this point in federation development.

2. Aggregation and Interchange

In most cases, the smallest unit of metadata that I'll be dealing with here will be that for a single entity: an `<EntityDescriptor>` element, in other words. That unit might be transformed in various ways as it moves around, but in many ways we can treat it as indivisible.

There are many circumstances in which we want to perform metadata *aggregation:* for example, aggregating the metadata of entities belonging to their members is one of the major functions of SAML federations today. This simplifies deployment by the users of the aggregate both by providing a single point of acquisition for the metadata of many entities, and by having the federation act as a broker of technical trust. This is usually signified by the federation's signature on the metadata aggregate, resulting in a *trusted aggregate.*

I'm using the word "trust" here in the sense of technical trust, which is to say that the consumer of a trusted aggregate can verify the authorship and integrity of the aggregate by verifying the signature. Other kinds of trust, such as behavioural trust, or other semantics such as an assertion of "truth" or legal effect, are *not* implied for all signed aggregates. Such additional layers of meaning may be inferred for particular trusted aggregates only when those are intended by the signing entity.

Because the signature is part of the trusted aggregate itself rather than being bound to the transport mechanism, a trusted aggregate can be verified by any consumer no matter how they acquire it.

At present, federations generally only aggregate metadata associated with entities belonging to their federation members. In moving beyond the single-federation *status quo*, these notes are concerned with the case where metadata documents from multiple sources are aggregated.

My general model for this larger world is that of a directed graph:

- Nodes in the graph are aggregators, such as federations.

- Arcs in the graph represent transfers of aggregates from one aggregator to another. A bi-directional flow is of course represented by a pair of uni-directional flows between the partners.

Some previous work has tended to assume a hierarchical model of aggregation. I'm not particularly confident that such a model is viable across *all* the communities of interest, although of course it may be appropriate to some. I'm therefore using a more general model which can represent a hierarchical scheme as a subset without requiring an assumption of hierarchy.

3. Aggregate Publication

The SAML metadata specification defines a publication method for entity metadata documents involving placing the metadata at the location that is the same as the URI used as the entity name. I don't believe the trust issues around this method have yet been addressed, so for the purposes of these notes I'll assume that in practice individual entity metadata is always published as part of an aggregate by some registration authority, normally a federation.[1]

Most metadata aggregates are published by being placed at a known location, with integrity assured by a digital signature within the metadata document itself. The SAML metadata specification allows for `cacheDuration` and `validUntil` attributes to control freshness and prevent "old metadata" attacks, but not all publishers make use of these at present.

We should not rule out the possibility of other publication schemes, however. For example, as we move towards potentially very large aggregates of aggregates, an aggregator may wish to make individual entity metadata documents available independently through some request protocol.[2] In other words, I think we need to further separate the creation of a (logical) aggregate by an aggregator from its publication.

4. Home Federation and Multi-Homing

An entity has at least one *home federation,* by which I mean a federation which is prepared to register the entity under the federation's rules, using the federation's declared registration process.

Entities with more than one home federation may be said to be *multi-homed*. The way that higher education federations have grown up to date has been on a national basis; this has meant that many international service providers are necessarily multi-homed at present because there is no other way for them to service multiple federations than to register with each in turn. We might expect multi-homing to decrease in the longer term if inter-

[1] That's in practice. In theory, though, I think the model I'm discussing here generalises to individual metadata documents as well, and I'd regard it as a goal of the aggregation appliance described later to be able to handle this case when the trust issues are addressed.

[2] Something as simple as HTTP with a query parameter would probably suffice, as long as the resulting XML document is appropriately signed.

federation becomes a reality[3], but for now we must treat it as endemic and any inter-federation mechanism we deploy must therefore be able to deal with it well.

My preferred way of looking at this issue is to treat it not as a “multi-homing problem” as such (which might suggest that it should be dealt with by the home federations) but as a “metadata conflict” problem instead. In other words, we should treat any point in the system where aggregation of metadata from multiple sources occurs as potentially seeing a metadata conflict, in which metadata for the same entity appears more than once. If each such aggregation point possesses rules to resolve conflicts, multi-homing does not present a problem. For example, a federation accepting metadata from elsewhere might rely on a rule that always gives locally registered metadata precedence in the event of a conflict. An alternative would be to always give precedence to “newer” metadata from whatever source, although this would rely on a metadata extension being agreed for placing time-stamps on `EntityDescriptor` elements.

5. Entity Mobility

An entity is defined as *mobile* if its registration in one of its home federations has the effect of making the entity's metadata available to entities in other, non-home, federations. I don't mean this in the trivial sense that someone could configure an entity registered in federation B to consume federation A's metadata as well as federation B's: this is always possible, and requires no work on the part of either federation A or federation B. I mean it in the specific sense that metadata originating from an entity's registration in federation A becomes available to members of federation B as part of the latter federation's published metadata. In other words, federation B accepts that federation A can act as a *registration authority* for federation A's mobile entities within federation B.

6. Mobility Mechanisms

There are two obvious classes of mechanism by which entity mobility could be achieved:

- by means of a home federation's normal published metadata

- by means of a separate metadata aggregate

Although the first alternative is probably the simplest to implement, I believe we should lean towards the second because (amongst other advantages, described later) it allows the metadata being made available to other federations to be different to the metadata published to the federation's own members. In particular, this means that it can be made to conform to a multilateral profile developed for the purpose, where a federation's own metadata might not do so in a variety of ways.

[3] Indeed, some may have concerns that entities would tend to register with the least expensive federation available to them, adversely affecting the business models of federations with higher charges and forcing down prices. That has certainly been the experience in the DNS and SSL certificate arenas, which are analogous in many ways.

7. Bilateral vs. Multilateral Mobility

At present, we're actively considering a small number of bilateral arrangements: federation-A with federation-B, federation-A with federation-C, let's say. This might potentially result in federation-A providing a different mobile metadata aggregate to each of its partners. It should certainly be possible for a federation to publish more than one mobile metadata aggregate should the need arise.

When we reach a large number of inter-operating federations, however, this mechanism will be hard to scale. I propose, therefore, that we regard such “targeted” aggregates as a last resort, and use as our baseline assumption a scenario in which each federation publishes a single aggregate of mobile entity metadata to all interested parties. Similarly, each federation would subscribe to the (single) mobile entity metadata aggregate for each federation which it regards as reliable enough for its purposes.

This might or might not match the legal perspective, in which we're more likely — at least initially — to want to negotiate multiple bilateral relationships. My own feeling, however, is that even if that is the case for now that we'd be unwise to assume that it will always be necessary. For example, I can easily believe that many federations will be perfectly happy to accept the mobile entity metadata aggregate of, for example, the UK federation or InCommon without any legal agreement backing it up at all. To help foster this kind of use case, I think we could and should look more closely into Leif's suggestion that federations should make it explicit that this kind of ad hoc, un-negotiated use is acceptable to them, for example by attaching Creative Commons or similar licences allowing the creation and publication of derivative works. Of course, we'd also want to make the absence of liability in such use cases clear as well.

8. Universal vs. Selective Mobility

It would be possible to consider a world of *universal mobility:* registration of an entity in any federation would result in its appearance in the metadata of other participating federations. In other words, each federation would publish metadata which was the union of the metadata of entities it had registered combined with the metadata of all entities registered by other partner federations.

A variation on this theme is the *regional universal aggregation* model, in which a meta-federation is formed from the complete metadata aggregates of a number of smaller federations; the metadata aggregate of the meta-federation would then be made available to members of all the individual federations. This might be a workable solution in a situation where all federation operators in a region trusted each other completely.

On the contrary, however, I currently believe that — certainly in the initial stages — not all entities registered by a federation should be assumed to be mobile. In particular, I suggest that we work under the assumption that making an entity mobile should require an affirmative opt-in step by the entity's owner to the home federation operator.

Selective mobility is another reason to prefer a separate mobile metadata aggregate rather than making use of federations' existing aggregates: although it would be possible to label mobile entities within a larger aggregate, using a separate aggregate means not having to agree on a labelling standard.

I'd also note that a *selective* regional aggregation model, in which a meta-federation aggregates the *mobile* metadata aggregates of other federations before presenting it back to them, makes a lot more sense to me than the universal variant.

9. How Much Mobility is Required?

Let IdP-A have federation-A as its home federation, and SP-B have federation-B as its home federation. In order to communicate, each entity must have access to the metadata for the other entity. If each entity consumed only the metadata for its own home federation, this would mean that both entities would need to be mobile: federation-A's registration for IdP-A would need to be available within federation-B, and federation-B's registration for SP-B would need to make metadata available within federation-A. A good name for this might be *symmetric mobility.*

The alternative — likewise, reasonably called *asymmetric mobility* — would involve only one of the two entities becoming mobile. In principle, either IdP-A or SP-B might be mobilised, with the other's metadata being published only within the metadata for its home federation. In this case, the mobile entity must consume the published metadata for both federations in order for communication to be possible.

There are several reasons not to assume that we always want symmetric mobility, and in fact my proposal is that at least for now we perform only asymmetric mobility, restricted to service provider entities. This preserves the normally assumed function of federations as identity-centric communities as opposed to service-centric ones, and avoids some (behavioural) trust issues arising from the observation that identity provider entity metadata is likely to contain federation-specific *policy* metadata (scopes and identity assurance policy assertions) where service provider metadata does not usually do so at present.

The cost of asymmetric mobility of service provider entities only will of course in some sense fall on the service providers, who will have to:

- configure their system to accept metadata from multiple sources
- manage multi-federation discovery rather than devolving responsibility for discovery to their home federation's WAYF service

10. Transitive Mobility

If entity-A is registered by federation-A and made mobile, it may appear in the federation metadata of federation-B if federation-B accepts that federation-A's procedures and policies makes its published mobile metadata sufficiently reliable for federation B's purposes. Is this transitive? In other words, can federation-C sensibly accept federation-B's assertion of entity-A's metadata as sufficient for its purposes, a situation we might call *transitive mobility?*

My feeling is that this situation suffers from the same “introducer problem” that we see in web-of-trust systems like PGP. For example, as federation-C, I may vet federation-B’s registration procedures and accept that if followed that they meet my own standards. That does not necessarily mean that I trust federation-B to make that same judgement on my behalf when it comes to federation-A’s procedures. In other words, we need to separate trust in someone acting as a registration authority from trust in that same person as an introducer of registration authorities.

There may be relationships in which transitive mobility is possible, particularly in situations where it is an explicit goal such as a regional meta-federation. However, I would suggest that we avoid this case for the foreseeable future, and assume that federations only mobilise entities they themselves have registered, and not any which they have obtained metadata from through subscription to another federation's mobile entity metadata aggregate.

11. Operational Aspects of Mobile Aggregate Publication

Logically, this can be broken down into *selection, transformation* and *publication* phases.

Selection: A federation should only include opted-in entities in its mobile entity aggregate. In addition, it should exclude any entities whose metadata can't be transformed in such a way as to meet the agreed inter-federation metadata profile.

Transformation: the metadata for the selected entities is put in a form which meets the inter-federation metadata profile, for example by removing elements meaningful only to the home federation (e.g., federation-specific ID attributes, custom extensions).

Publication: the resulting aggregate should be signed by the home federation's normal metadata signing key and published at an announced location.

12. Operational Aspects of Mobile Aggregate Subscription

Logically, this can be broken down into *subscription, transformation* and *selection* phases.

Subscription: the metadata aggregate from the other federation is periodically fetched and its signature verified.

Transformation: probably mostly a repeat of the publication transformation phase, to allow robustness if the other federation deviates slightly from the inter-federation metadata profile.[4] In addition, elements specific to the destination federation may be added, for example to indicate the source of the metadata.

Selection: some entities may have to be discarded at this point, for example to provide conflict resolution.

13. Inter-Federation Metadata Profile

We need to develop this as a formal specification, I believe. It should be layered on top of Scott's current Oasis SAML TC draft profile for interoperable SAML metadata. In particular, this means that key material must be embedded rather than referenced by

[4] Detecting and rejecting non-conforming metadata here would be another option.

KeyName, which can't be assumed to have the same semantics in an arbitrary destination federation. This means that entities for which the home federation does not have embeddable key material can't be mobilised.

I don't think that we want this profile to say that the simple SAML 2 profile Andreas is developing is mandatory, but I would want to say that it is strongly recommended at least for mobile entities that support SAML 2 at all.

14. Straw Man: Aggregation Appliance

One thing that is clear from discussions is that different federations will have different needs when it comes to metadata interchange with their peers, not least because their internal procedures and infrastructure have grown up independently. We might address this variation by building custom technical solutions for each pair of partner federations, but I believe that it is both possible and preferable to think instead in terms of building a common technology which is sufficiently modular and flexible to deployed in many situations. Leif coined the term “aggregation appliance” for this concept in a recent call; I’d like to commandeer the term as a seed from which a long-term deployment strategy for multi-party metadata interchange might be grown.

I’ll define an *aggregation appliance* as a hardware or virtual appliance containing amongst other things an *aggregation engine*; an application which performs the functions associated with aggregation of an arbitrary number of *inbound aggregates.* For each of these, the engine would be configured with:

- publication location and mechanism
- rules about refresh periods
- trust models for that aggregate
- white-list and black-list by entity ID or perhaps by general XPath expression
- possibly, generic transformation chains expressed in XSLT
- precedence rules allowing one inbound aggregate to win over another when the same entity appears in more than one place

In turn, the aggregation engine would be capable of signing and publishing the result of periodically processing the inbound aggregates into one or more *outbound aggregates* of its own.

To the above definition, I’d add the notion that if a federation ran such an appliance to generate the metadata aggregate published to its members, one inbound aggregate would necessarily be what we currently regard as “the federation’s metadata”: the aggregated metadata for entities registered by federation members. One result of this is that such an aggregation appliance could be plugged into existing federation production infrastructures without major rework being required to registration processes and databases.

Similarly, it’s likely that we can define the aggregation engine such that it can derive the mobile entity metadata aggregate from the federation’s registered entity metadata aggregate and publish that as well.