Methods of Metadata Distribution

New and Emerging Methods of SAML Metadata Distribution

In September 2016, the Per-Entity Metadata Working Group, led by Scott Koranda, submitted an interim report to the InCommon Technical Advisory Committee with the following recommendation:

> *The working group finds and recommends that InCommon Operations proceed immediately with the design, creation, and delivery of a new InCommon metadata aggregate that contains only the metadata for identity providers (IdPs). The new IdP-only aggregate will help relieve issues some SPs face as the size of the existing InCommon metadata aggregates continues to grow.*

Subsequently, on October 4, InCommon Ops announced the general availability of a production-quality IdP-only metadata aggregate for SP deployments. See the IdP-only Aggregate wiki page for details.

Using the IdP-only metadata aggregate

The new IdP-only metadata aggregate is intended for use by SP deployments only. IdP deployments should continue to use the main production aggregate described on the Metadata Aggregates wiki page.

Characteristics of the IdP-only Aggregate

The IdP-only metadata aggregate is approximately 16MB, which is about 42% of the size of the full InCommon aggregate.

Determining the size of the IdP-only aggregate

```
$ MD_LOCATION1=http://md.incommon.org/InCommon/InCommon-metadata.xml
$ MD_LOCATION2=http://md.incommon.org/InCommon/InCommon-metadata-idp-only.xml
$ curl --silent --head $MD_LOCATION1 | grep -F Content-Length
Content-Length: 38623782
$ curl --silent --head $MD_LOCATION2 | grep -F Content-Length
Content-Length: 16438778
```

As of October 10, the IdP-only metadata aggregate contains 2231 entity descriptors, of which 447 are registered by InCommon. Each entity descriptor contains an `<md:IDPSSODescriptor>` child element. Seven (7) of the entities contain an `<md:SPSSODescriptor>` child element as well.

Determining the number of entities in the IdP-only aggregate

```
# For a description of the output of the count_entity_roles.xsl script, see:
# https://gist.github.com/trscavo/f766a88ff5feb5937e5be5a16a1ff0c0
$ curl --silent $MD_LOCATION1 | xsltproc ./count_entity_roles.xsl -
6546,4319,3245,2231,447
$ curl --silent $MD_LOCATION2 | xsltproc ./count_entity_roles.xsl -
2231,7,0,2231,447
```

For a complete up-to-date list of IdPs in InCommon metadata, see the List of IdP Display Names wiki page.

Benefits and Risks of the IdP-only Aggregate

Since the IdP-only metadata aggregate is significantly smaller than the full aggregate, the former buys valuable time for service provider deployments—especially modestly provisioned deployments—until per-entity metadata becomes readily available. For one particular class of service providers, the IdP-only aggregate will continue to be essential infrastructure long after other InCommon deployments have migrated to per-entity metadata. The rest of this section explains why this is so.

The vast majority of SPs do not have a dynamic discovery interface (i.e., a discovery interface that depends on published metadata) and so these SPs will be able to leverage per-entity metadata without delay. In fact, many of these SPs depend on a small number of fixed IdPs so the migration to per-entity metadata will be straightforward for them.

On the other hand, for the relatively few SPs that implement a dynamic discovery interface, the benefit of per-entity metadata is less clear since these SPs currently require an aggregate for IdP discovery. We expect these SPs to consume the IdP-only aggregate until the community addresses the IdP discovery issue brought about by per-entity metadata.

Be aware that there is no fallback aggregate of IdP-only metadata. In that sense, there is some risk associated with the use of the IdP-only aggregate. If you must fall back, you will have no choice but to fall back to the full Fallback Aggregate described on the Metadata Aggregates wiki page.

The Future is Per-entity Metadata

The Per-Entity Metadata Working Group is expected to submit its final report to the InCommon TAC by November 2016, after the community has reviewed the report. We anticipate that the working group will recommend that InCommon Operations deploy a production-quality metadata query server and that all InCommon SAML deployments (except those SPs that implement a dynamic discovery interface as discussed above) migrate to per-entity metadata as soon as possible.

Eventually all SAML deployments will benefit from per-entity metadata. IdP deployers, in particular, are anxiously awaiting the arrival of a metadata query server, and we expect many IdPs will be among the first deployments to realize the benefits of per-entity metadata.

What is per-entity metadata?

The SAML specification defines two entities: the *Identity Provider* (a producer of SAML assertions) and the *Service Provider* (a consumer of SAML assertions). A Service Provider requires the "metadata" of the Identity Provider (and vice versa). The metadata describe a SAML deployment, providing security, privacy, and interoperability to the relying party.

As a practical matter, SAML metadata is batch distributed as an aggregate of entity descriptors. With the proliferation of global aggregation services such as eduGAIN, the size of aggregates has grown dramatically, which is causing federations to re-examine existing methods of metadata distribution.

The term "per-entity metadata" refers to a single entity descriptor. The Metadata Query Protocol is an emerging standard that describes how to obtain per-entity metadata from a trusted oracle. Since the entity descriptor is the basic unit of policy and interoperability, this method of metadata distribution is both logical and efficient.

Two SAML implementations are known to support the Metadata Query Protocol: simpleSAMLphp and Shibboleth. (See the MDQ Client Software wiki page for more information.) In particular, support for the Metadata Query Protocol was introduced in version 3 of the Shibboleth IdP software. Shibboleth IdP deployments that have upgraded to Shibboleth IdP V3 will be among the first to migrate to per-entity metadata.

Shibboleth IdP V2 End-of-Life

Shibboleth IdP V2 reached end-of-life on July 31, 2016. In the future, no bug fixes, not even security-related bug fixes, will be issued. Upgrade to Shibboleth IdP V3 now!

Other SAML software will benefit from per-entity metadata as well. For example, Microsoft AD FS can be configured to retrieve a single entity descriptor from a metadata query server, which is a huge step in the right direction. The hope is that AD FS and other SAML implementations will eventually support the RESTful Metadata Query Protocol like simpleSAMLphp and Shibboleth.