Introducing per-entity metadata service

Jump to:

What is per-entity metadata? | Why use per-entity metadata? | System availability and performance | Further reading

The InCommon Metadata Service employs a style of metadata distribution referred to as "per-entity metadata". This page explains what is per-entity metadata.

What is per-entity metadata?

Per-entity metadata allows your identity provider(IdP) or service provider(SP) to retrieve metadata for individual entities in a as-needed, real-time manner. It eliminates the need to download and pre-load the entire aggregate. Metadata is delivered through a protocol called MDQ ("Metadata Query," details at: https://datatracker.ietf.org/doc/draft-young-md-query/.

TIP: Per-entity metadata is often referred to as "MDQ", named after the protocol it is based on, in conversation. When you hear "MDQ" in a InCommon metadata -related discussion, it is synonymous with "per-entity metadata".

Why use per-entity metadata?

With per-entity metadata (MDQ), your SP or IdP only retrieves the metadata for a specific entity as it is needed. For example, your IdP, may only regularly interact with a fraction of the registered SPs in InCommon. With per-entity metadata, it will no longer need to load every InCommon SPs metadata in to memory. It only retrieves and loads the metadata needed for SSO integration. This results in a significantly lower memory footprint. It also leads to quicker system start up time. Using MDQ, the UK Federation reports that it is possible to run the Shibboleth IdP with as little as 500MB heap size compared to the current recommendation of a minimum heap size of 1.5GB.

Additionally, because metadata for a specific entity is available at a specific URL, it is possible to pre-fetch metadata for entities you may consider high-value, such as commonly-used entities. For example, an SP in InCommon that primarily interacts with a single IdP could pre-fetch and cache that one IdP on startup and refresh it on a regular basis (just as you would with the aggregate). This setup can help minimize risks specified below.

System availability and performance

In a per-entity metadata service, an entity's metadata is fetched on-demand when federation software needs it. This typically occurs when a user attempts to access a service using his/her federated credential. This means that if the distribution service becomes unavailable and the metadata client has not cached the relevant metadata, a user would not be able to log in to the service he/she is trying to access.

> *Related: Amazon CloudFront's Service Level Agreement.*

InCommon has engineered its metadata service to ensure high availability and performance by employing technologies such as deploying the service via a commercial CDN (Amazon CloudFront). In addition, we are also developing automatic service monitoring to automatically sent alerts when there is anomaly. Stay tuned for updates on information about how to subscribe to these monitoring alerts.

Further reading

Related content

- Migrating to the MDQ Service
- Metadata Distribution Service Documentation
- Introducing per-entity metadata service
- Configure Shibboleth service provider
- Configure other software
- Prefetch an entity with Shibboleth
- Locating the production metadata
- Metadata signing key for the Production environment
- Metadata Service
- Configure Shibboleth identity provider

Get help

Can't find what you are looking for?

help Ask the community