We're Moving Your Metadata (But Not Your Cheese)

InCommon Forum
July 10, 2019

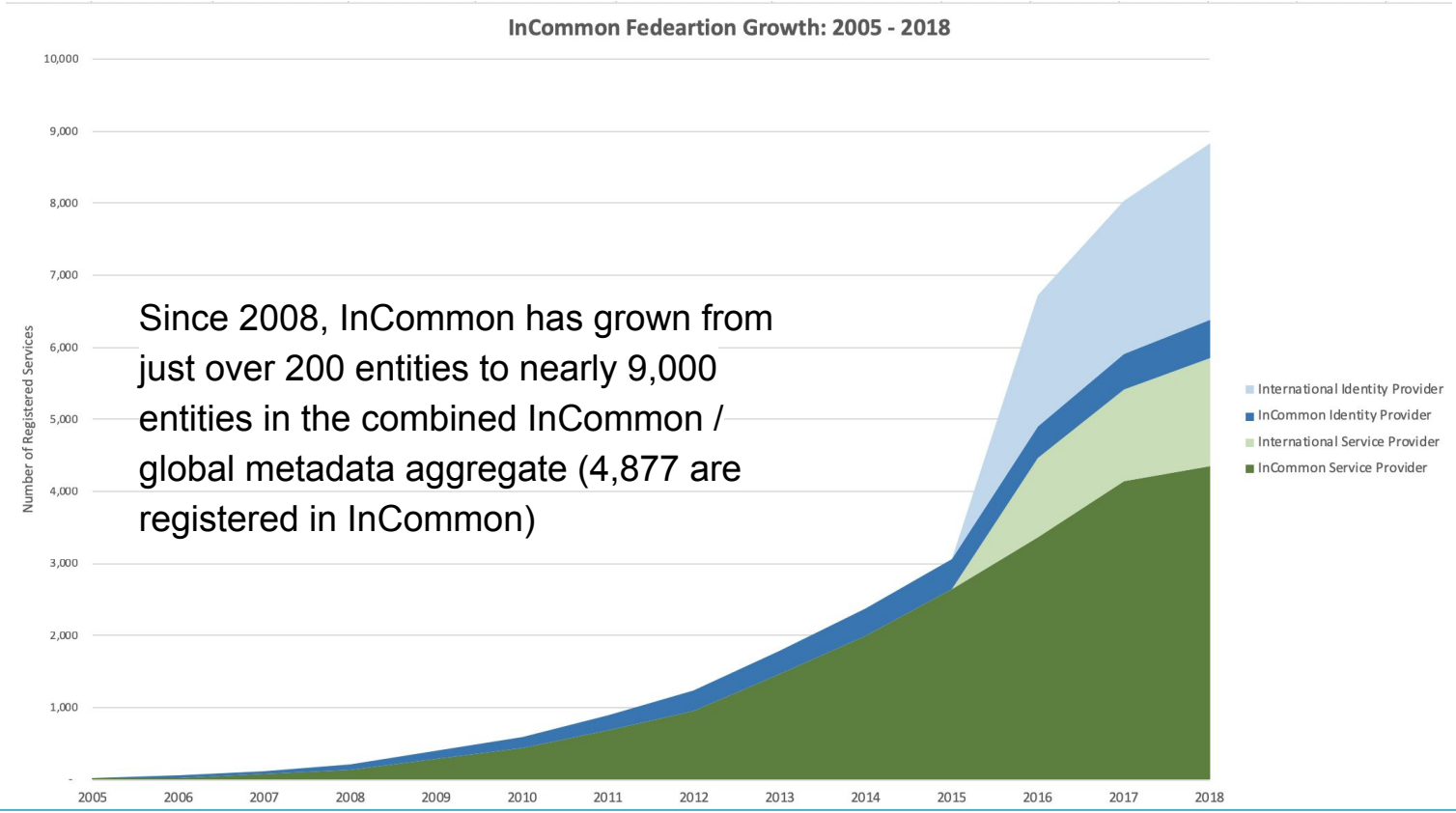
Albert Wu, Federation Service Manager, InCommon
Nick Roy, Director of Technology and Strategy, InCommon

New Metadata Distribution Service

Today's InCommon Forum explores InCommon's new Metadata Distribution Service, what it means to you, and what you will need to do to get ready for it.

- Origin and Introduction
- Project Timeline and Status
- Impacts on Deployers
- Getting Started With The Service
- Providing Your Feedback

InCommon has grown up



Since 2008, InCommon has grown from just over 200 entities to nearly 9,000 entities in the combined InCommon / global metadata aggregate (4,877 are registered in InCommon)

The explosive growth exposed scaling issues

As we scale, major drawbacks surfaced in the current metata aggregate distribution strategy:

- An error in a single entity descriptor in the metadata aggregate can cause widespread outage for services depending on the metadata aggregate.
- Large metadata aggregate slows system startup time for IDPs and SPs
- Dramatically increased memory requirements wastes system resources and strains resource-constrained deployers' ability to fully participate in federation.

Finding a better way to distribute federation metadata

- Allow runtime, need-based entity metadata retrieval - based on the SAML profile for the Metadata Query (MDQ) Protocol https://datatracker.ietf.org/doc/draft-young-md-query-saml/
- No need to preload the entire metadata aggregate
- Secure, scalable, and fault tolerant
- Already supported in Shibboleth, including smart caching to guard against service interruptions
- Backward compatible: you can still download and consume an aggregate

Paving a Path to a new Metadata Distribution Service

2014

Metadata Distribution WG: Recommendations:

- Expand use of multiple metadata aggregates
- Conduct pilot study to explore feasibility of per-entity metadata
- Conduct landscape study of needs and uses of hardware security modules
- Participate in the samlbits.org project
- https://spaces.at.internet2.edu/x/F4G8Ag

2016 - 2018

Per-Entity Metadata WG Report (2016):

- Defined requirements for a Per-Entity Metadata Distribution Service based on the MDQ protocol
- Outlined implementation and operational considerations.
- http://doi.org/10.26869/TI.5.1

In parallel, InCommon Operations:

- Conducted MDQ pilot study
- Developed strategy to scale to the cloud
- Designed solution to align with Per-Entity Metadata WG requirements

2019

Spring 2019:

Launch Technology Preview of Metadata Distribution Service (MDQ Service)

Summer 2019:

Metadata Distribution Service goes live.

- Release Candidate
- It’s live: you can use it in production
- Adopt now if:
  - you are comfortable with early adoption
  - you are experiencing issues loading current metadata aggregate
- https://spaces.at.internet2.edu/display/mdq

The InCommon Metadata Distribution Service is available **today**.


45th ANNIVERSARY
QUALITY AUTHENTIC INTEGRITY FAMILY
Cacique®
KEEP REFRIGERATED
Made in U.S.A.
TWIN PACK
Ranchero® BRAND
QUESO FRESCO
Part Skim Milk Cheese
Made with Grade A milk.
Made with milk from cows not treated with rBST*
REAL CALIFORNIA MILK
Recipes and Usage www.caciqueinc.com
24 OZ (1 1/2 LB) (680G)
PROUDLY MADE BY CACIQUE INC. INDUSTRY, CA 91744
Nutrition Facts
24 servings per container
Serving size 1oz (28g)
Amount per serving
Calories 80
% Daily Value*
Total Fat 6g 8%
Saturated Fat 4g 20%
Trans Fat 0g
Cholesterol 20g 7%
Sodium 200mg 9%
Total Carbohydrate 0g 0%
Dietary Fiber 0g 0%
Total Sugars 0g
Includes 0g Added Sugars 0%
Protein 6g
Vitamin D 0mcg 0%
Calcium 117mg 10%
Iron 1mg 6%
Potassium 28mg 0%
INGREDIENTS: CULTURED PASTEURIZED GRADE A MILK AND SKIM MILK, SEA SALT, AND ENZYMES. CONTAINS MILK.
0 74562 22112 8
FM2211216
*No significant difference has been shown between milk derived from rBST-treated and non-rBST treated cows.

Queso Fresco

Nick Roy
Director of Technology and Strategy, InCommon

Why Are We Moving Your Metadata?

- The old metadata signing/publishing process is dependent on a lot of manual processes
- The new metadata service supports per-entity metadata (“MDQ”)
  - WAY less memory usage and faster load times for participant IdPs and SPs
- The old signing key had been in use for 15 years, time for a new key
- New service features required radically different infrastructure
- New infrastructure and new key means we need new metadata endpoint locations

Service Operational Plans

Requirements from the Final Report of the 2016 InCommon Per-Entity Metadata Working Group (summarized):

Full report: http://doi.org/10.26869/TI.5.1

- Mitigate risk of man-in-the-middle attacks by maintaining document signature and using TLS
- Service uptime of 99.99% on a monthly basis
- Latency of no more than 200ms for 99% of queries ***from the Internet2 network***
- Monitor performance and availability from geographically disparate locations, make publicly available

Security

Metadata is the root of trust for technical interoperation in a federation

- Both legacy and new metadata services support XML signature verification and TLS
- New metadata service uses a hardware security module for signing metadata

Availability

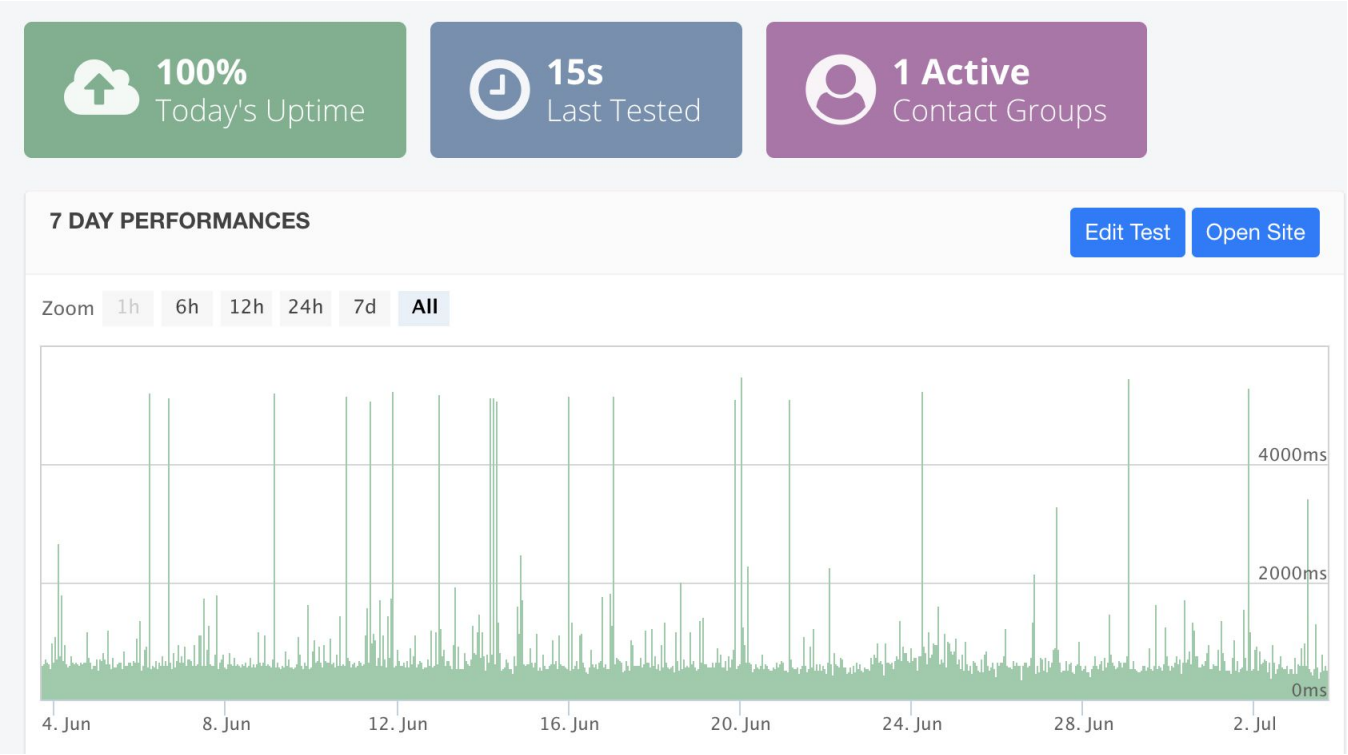
Because the newly-supported MDQ protocol involves real-time queries for metadata, availability is crucial

- We have chosen a commercial content distribution network, Amazon CloudFront, to achieve at least 99.9% uptime
  - It was infeasible from a cost perspective to achieve a ‘four nines’ guarantee

- We are doing geographically distributed monitoring using StatusCake
  - Service status URL is included in our service documentation, and covers more than just the metadata service

- Testing indicates that metadata in the edge cache achieves <= 100ms latency, for > 99% of queries ***from a location on or near the Internet2 network***

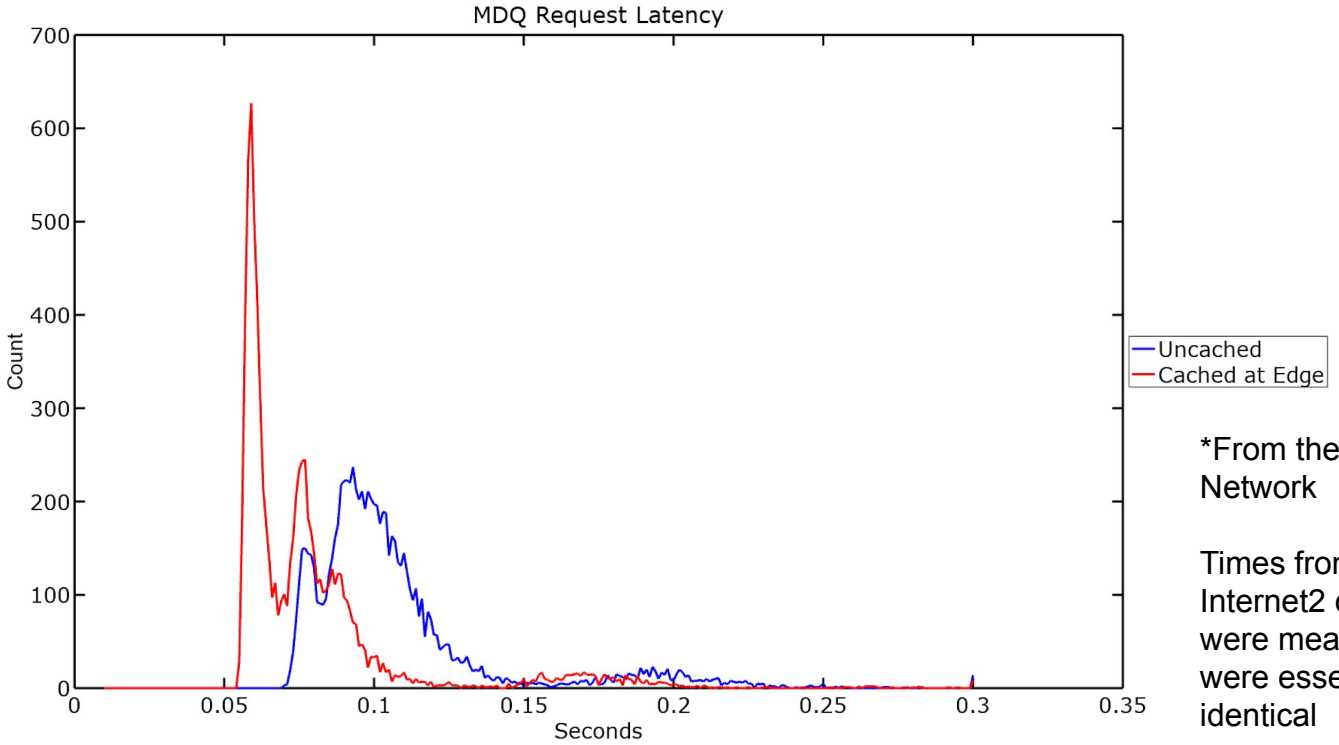
Availability



A Note About Performance Measurement

- StatusCake shows significantly higher response times than our measurement from on the Internet2 network, and from an I2 network connector
- Likely due to a combination of timing (cache expired), network latency, spreading out their checks across different locations over a long period of time (each check hitting from a different location so edge cache takes a cache hit)
- We are hitting our availability goals

Cached vs. Non-Cached Metadata Response Times*



*From the Internet2 Network

Times from an Internet2 connector were measured and were essentially identical

Security

Retiring Old Signing Key

- The original InCommon metadata signing key was created in a secure ceremony on March 30, 2004
- It has been in use since then
- It is time for a new, larger key
- We're taking the opportunity of changing where all InCommon SAML deployments get their metadata from to switch to a new signing key, 3072 bits in length, to provide some additional breathing room
- The new key was generated in a secure, documented process on April 8, 2019 in the Internet2 Denver office, with five Internet2 staff present, including one corporate officer
- We have a key recovery process based on n of m with a number of "shard stewards" who need to come together to agree to recover the key in the case of a DR scenario
- The new key was securely imported into an Amazon HSM and steps have been taken to ensure it cannot be exported

When you start using the new metadata service, you must configure your deployment to use the new public key for metadata signature verification

Testing and Feedback

Technology Preview

- The metadata technology preview has had no reports of availability or performance issues. It was made available to users of the old MDQ-beta service in February
  - A couple deployers ran into issues with metadata filtering, and have provided work-arounds which are included in our documentation
- Legacy MDQ-beta service retired shortly after the technology preview was made available
- The technology preview replaces the old “preview” metadata aggregate
- It is where we will do public-facing testing of new features and changes
- It uses a separate key from the new metadata service release candidate

Service documentation landing page: https://spaces.at.internet2.edu/x/2wR0C

Testing and Feedback

Release Candidate

- We are providing the new production service in a “release candidate” mode for the next few months
- This service is effectively production-ready, with some finishing touches yet to apply
- The service should meet our performance and availability goals from here on out
- The key for the release candidate is long-lived and stable

We need your help! Please try out the new service

- Provide feedback using the “Get Help” link on the new InCommon web site: https://incommon.org/help/

Service documentation landing page: https://spaces.at.internet2.edu/x/2wR0C

Migration Plans

Migration from Existing Metadata Distribution Service to New MDS

- “Aggregates” live on in the new service
- We will seek “official” community acceptance of this service, soon
- Then begins a lengthy migration campaign to get ***everyone*** over to the new service
- We’ll monitor use of the old metadata distribution endpoints and nag people to move
- We will set a deadline for this move (TBD)
    - ***Eventually…***
      If you don’t move by the deadline, you will stop getting updated metadata

Thank You and Q & A

Contact Info:

Albert: awu@internet2.edu
Nick: nroy@internet2.edu

General InCommon-related questions: help@incommon.org

New InCommon website: https://incommon.org

Upcoming Events

...

BaseCAMP - Training and Learning Opportunity - InCommon Federation
August 13-15 - Milwaukee, Wisconsin
https://meetings.internet2.edu/2019-basecamp/