Agenda and Notes - 2016-09-07

Per-Entity Metadata Working Group - 2016-09-07
Agenda and Notes

[EtherPad used to create these notes: Agenda_and_Notes_-_2016-09-07.etherpad]

Dial in from a Phone:
Dial one of the following numbers:
+1.408.740.7256
+1.888.240.2560
+1.408.317.9253
195646158 #
Meeting URL (for VOIP and video): https://bluejeans.com/195646158
Wiki space: https://spaces.at.internet2.edu/x/T4PmBQ

Attendees

- Scott Koranda (LIGO)
- Michael Domingues (University of Iowa)
- David Walker (Internet2/InCommon)
- Tom Scavo, InCommon/Internet2
- Tom Mitchell (GENI)
- Ian Young
- IJ Kim, Internet2
- John Kazmerzak, University of Iowa
- Walter Hoehn, Memphis
- Rhys Smith, Jisc
- Phil Pishioneri, Penn State (leaving @14:25UTC)
- Chris Phillips, CANARIE
- Paul Caskey, Internet2
- Scott Cantor, tOSU

Agenda and Notes

1. NOTE WELL: All Internet2 Activities are governed by the Internet2 Intellectual Property Framework. - http://www.internet2.edu/policies/intellectual-property-framework/
2. NOTE WELL: The call is being recorded.
3. Agenda bash
4. Interim report/finding on IdP only aggregate
    a. Any feedback from TAC or Steering?
    b. What happens now?
    c. Anything further needed from the working group?
    d. TomS: The TAC accepted the interim report and will forward it to Steering as an FYI. (It doesn't require a Steering vote.) Nothing more is needed from the working group.
5. Update from UK fed MDQ rollout (Rhys)
    a. Moved UK federation infrastructure to Azure
    b. Ian has created Shibboleth MDA pipelines to take a single aggregate and output per-entity files, then sign them (performed with HSM)
    c. Symlinking SHA1 hash of entityID to per-entity file (and supports gzipped versions of each)
    d. Result is a MDQ server (Apache) that serves static files that are generated whenever a new aggregate is created (once it's in production).
    e. Not using commercial CDN at present (?)
    f. Pipeline performance (N.B.: Uses a "top-of-the-line" HSM -- Thales nShield Connect): Generated 3605 files in 00:01:20
    g. 10.45k is the average size of the per-enttiy metadata files
    h. If you want to hit it, it's at http://mdq-test.ukfederation.org.uk/entities/
        i. e.g. http://mdq-test.ukfederation.org.uk/entities/https:%2F%2Ftest-idp.ukfederation.org.uk%2Fidp%2Fshibboleth
6. Acceptable latency for our Requirements section
    a. Where do we measure latency?
    b. What numbers do we require?
    c. Strawman: "The 99th percentile of response times for queries to the distribution layer must be less than 500ms, as measured from [the Internet2 backbone]
        i. Do we have a measurement point on [the Internet2 backbone]?
    d. Over what period should we do these measurements?
    e. How often should the measurements be taken?
    f. The IdP could be instrumented to log response times. We could monitor selected IdPs.
        i. We should request that instrumentation from the TIER project.
    g. We'll change 500ms to 200ms at 99th percentile (David and Scott K will add further specifiers / decorations to this metric)
7. Responsiveness/Performance
    a. Ability to maintain the latency requirements over time -- should include target latency over rate of queries -- incorporate load for target metrics
    b. Understanding the initial load to be placed on the servers is hard. Function of net federation login activity (future) as opposed to net aggregate size (current). We'll put the issue on Ops's road map.
    c. The performance targets above will be measured on a monthly basis
        i. Scott C: In the education space (as opposed to the commercial sector) we tend to have different seasonal load peaks
8. HTTPS and the TLS trust model for InCommon MDQ service
    a. Further discussion of pros and cons
    b. Consensus for report?

        i. HTTPS is desirable, but not required in our specs (?)
            1. Perhaps a Phase II item?
        ii. We'll need to decide what the certificate should be.
        iii. ScottC: At some point, we'll need to address validity times
            1. These will probably be hours or days.
            2. TLS can help mitigate the risk of getting stale metadata from a spoofed server.
    c. To be clear, TLS is *not* a substitute for the signatures in the metadata.
    d. We'll continue this discussion via email and in next week's call.
9. Deployment architecture: CDN versus hosted servers
    a. What are the discriminators other than cost?
    b. (We ran out of time for this.)
10. Charter review: what have we missed?
    a. https://spaces.at.internet2.edu/display/perentity/Per-Entity+Metadata+Working+Group+Charter
    b. Everyone please review this before next call to make sure we aren't missing anything.
11. Timeline going forward (2 calls)
    a. Propose that we focus on the report deliverable
    b. Use calls to efficiently work through issues on text/diagrams
    c. Any remaining time is spent brainstorming on discovery