Agenda and Notes - 2016-08-03

Per-Entity Metadata Working Group - 2016-08-03
Agenda and Notes

[EtherPad used to create these notes: Agenda_and_Notes_-_2016-08-03.etherpad]

===>> Note the new PIN and meeting URL <<===
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

- David Walker, Internet2
- Ian Young
- Phil Pishioneri, Penn State
- Michael Domingues, University of Iowa
- Paul Engle, Rice U
- Tom Scavo, InCommon/Internet2
- Tommy Doan, Southern Methodist University
- Scott Cantor, tOSU
- Tom Mitchell, GENI
- John Kazmerzak, University of Iowa
- Rhys Smith, Jisc
- Paul Caskey, Internet2
- Walter Hoehn, Memphis
- Chris Phillips, CANARIE

https://public.etherpad-mozilla.org/p/Agenda_and_Notes_-_2016-08-03
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
   a. Should we talk about (functional) requirements for the service before risks?
      i. Qualities of the service -- expected and how close actual existing meets it (why the 'requirement' or expectation is suggested) - CP
      ii. Some people have been assuming a "DNS" model, that the service is very reliable, not usually requiring special client-side mechanisms to accommodate to failures.
4. What are the risks for a per-entity metadata service and the possible mitigations
   a. I suggest we list risks along with their likelihood, impact, and potential mitigation (DHW)
   b. Risks from last week's call (https://spaces.at.internet2.edu/x/pYIABg) and subsequent electronic mail discussion
      i. Availability
         1. Expectations: ability to query for a given piece of metadata at anytime
         2. Failure of the distribution service for IdPs and SPs for longer than ??
         3. Failure of the aggregation/signing service for longer than ??
      ii. Security
         1. Q: will MDQ have any material difference in security than the existing aggregate?
            a. Scott/Michael -- no difference at this time.
         2. Disclosure of the signing key
         3. IdPs and SPs that do not verify signatures
         4. Clients not checking metadata signatures
      iii. Service Delivery

> **For reference: Terms and their meaning around availability and uptime implications**
>
> 3 9's allowed downtime: 8.76hrs/yr, 43.8 min/month, 10.1 min/week
> 4 9's allowed downtime: 52.6 min/yr, 4.32 min/month, 1.01min/week
> 5 9's allowed downtime: 5.26 min/yr, 25.9 sec/month, 6.05sec/week

         1. Expectations:
            a. Q: should perfect reliability assumed? (Scott C)
            b. Observations:
               i. Rhys -- as reliable as the current delivery model, as reliable as possible. Since serving static content, could throw it on a commercial CDN if necessary
               ii. Chris -- similar to Rhys, but in order to deliver 5 9's like experience, caching at various levels to contribute to the whole. +1 to CDN comment

iii. Different clients will present diversity on how to solve availability.
iv. There are mitigations that don't involve mods to the IdP/SP code (e.g., http caching proxies)
v. There are no 100% solutions.
vi. What is an acceptable level?
  1. High 90s (for the aggregation/signing portion of the infrastructure)
  2. At least Akamai (for the distribution portion of the infrastructure) (Walter H)
    a. At least 2 9's, probably 3 or 4.
vii. Consensus (in this call) is that we need at least 3-4 nines of reliability in the distribution service, even better.
viii. Note that retrieving (reading) an MDQ artifact/response is DIFFERENT than being able to UPDATE the content of the MDQ response.
  1. These should be considered separate qualities.
    a. e.g. you may need 5 9's on read/publishing the content, but can tolerate changing the data less reliably (due to cost of offering said reliability)
  2. Clients start up with nothing cached.
    a. Should we recommend something for that?
    b. Is it something that's nice to have our something we *should* have before rolling this out?
      i. CONCLUSION:Consensus existing client-side caching is sufficient. ~~We can, however, tell them what they can do to increase reliability~~
        1. If you point out that people can add additional caching, this might invite questions of reliability of the service. Consensus around it not being worth mentioning that at all.
ix. Does MDQ change the calculus about using federation infrastructure for storing /local/ service metadata?
  1. ANS: YES.
    a. Risk if (single path) internet connection goes down, lose access to metadata for local services.
    b. This could be an argument for enterprise-provided distribution infrastructure.

a. Further discussion of risks
  i. Responsiveness / Capacity
    1. Operations
    2. Expectations: Ability to sign metadata
      a. Q: is it 'real time'?
      b. Q Is it 'online signing?'
    3. The service is up, but unusably slow
    4. Capacity is not sufficiently elastic
    5. Rate of update
    6. Rate of query
      - Malfunctioning entity...
    7. Cost
      a. Cost of elastic capacity not budgeted
        i. Rhys: You can use the Azure CDN with current UK federation level of traffic (50 TB/year) --> 200 GBP per month
      b. Staff time and attention not sufficient
    8. *Your favorite risk here...*
b. Requirements for availability and scalability
c. Next call is August 10, 2016 @ 10:00 AM (America/New York)