8th FIM4R meeting, 3-4th February 2015
Meeting Notes

Agenda page: https://indico.cern.ch/event/358127/
Meeting notes: Romain Wartel

Summary of key points

Objectives of the workshop

Specific focus of the 8th FIM4R meeting:
- Aim at production services
- Highlight the importance of the Pilots
- Express requirements for eduGAIN + AARC (EU)
- Aim at solving a global problem
- Harmonise interoperation, policies and practices
- Discuss attributes definitions & release

Presentations

Introduction:
Ian Bird gave an overview of the LHC computing model, its needs, requirements and future evolution. He highlighted that large science projects have become global collaborations. As a result, federated identities for global communities must be a global effort. We must therefore agree common levels of service and trust, between all major parties (EU, US, etc.). These agreements are essential and urgent for progress and uptake.

Feedback or comments:
Ian clarified that CERN expects to meet the future growth needs through optimising the use of IT resources, getting more capacity through improvements in price/performance of equipment and supplementing with commercial cloud resources (example given was Rackspace/openstack but CERN is also starting to procure limited quantities of cloud service from many commercial providers and these will also need to be integrated into the federated AAI).

AARC - General presentation:
David Groep gave an overview of the AARC project and work packages, highlighting the project would build on and integrate with the very many existing and evolving components ESFRI clusters, eduGAIN, national AAI federations, NGIs, IGTF, SCI, SirTFi, etc.
REFEDS/FIM4R/RDA/EUDAT will not have operational role but can provide requirements & best practises.
ESFRI clusters/GEANT/EGI will handle operations and should have a role in the sustainability beyond the end of the AARC project.
FIM4R and RDA are also a means of working with the global community (beyond Europe).

David explained the main challenges are training, outreach & engagement (IdPs, SPs, users) together with technical and policy aspects.
The AARC project will start on 1st May 2015 for 2 years.

Feedback or comments:
- *The communities welcomed the message that the AARC management is committed to providing production services.*
- *Each AARC pilot should result in a service which can become sustainable.* ***Pilots are a way of measuring the quality of the output of the project****.*
- *There are already a lot of existing AAI technologies available and it will be necessary to filter those technologies that can be integrated and deployed in a production mode.*
- *AARC requested some guidance on where to target the training.*
- *It was noted that 3M€ is a limited amount compared to the work to be performed. (GEANT4 commented they will also have some resources to put in place services. Peter Sologna (EGI) added that EGI-Engage too has resources to focus on AAI services.)*

AARC - Pilots:
Paul van Dijk presented the AARC work packages on pilots. It will be driven by user requirements and focus on integration of existing building blocks. Key aspects will include the ability to support guest users, attribute management, aggregation & consumption, as well and non-Web and commercial (cloud) resources.

Feedback or comments:
- *The ability to identify the same individual as a user even when login in through different means is essential for a number of pilots.*
- *For the photon/neutron community and the umbrella pilot, it is important to track individual over their lifetime even if they change host organisations/employees since the data belongs to the individual not the hosting organisation.*
- *It was suggested that each community could manage its own guest-IdP. This was for example done for DARIAH and was accepted by DFN as part of eduGAIN. However the costs of running a guest IdP should not be under estimated. There may be a business case for a last-resort guest IdP but it cannot rely on the users paying for it directly.*
- *For Elixir, the training/outreach to SPs is essential since they need to be convinced of the added-value of FIM.*
- *REMS is a good example of showing added value for life science community, it should be promoted/reused.*
- *US engagement: it is very difficult to get SP in the US to show-up at REFEDS. Maybe Jim Basney could be a good contact point representing many SPs?*
- *It would be useful to create a list for interested parties to learn about AARC.*

Virtual Campus Hub project:
Niels van Dijk presented the Virtual Campus Hub project, connecting universities via eduGAIN. This enables them to share documents, applications and other online activities. The project produced useful recommendations for the uptake of FIM. It also highlighted the goals of the future OEUVRE VRE proposal.

US/EU collaboration on FIM:
Scott Koranda gave an overview of the US FIM landscape and of LIGO in particular. Key challenges include federating in scalable way with a global network of IdP and attribute release by higher education.
There are still a lot of bi-lateral agreements and 1-by-1 negotiations in InCommon. Wider adoption of the Research & Scholarship (R&S) entity category would clearly facilitate the deployment of FIM.
IdPs and SPs still have no obligation or agreement, moral or legal, to notify other of a security incident, but several key US stakeholders now also participate in Sirtfi. As far as LoAs are concerned, they are mostly ignored now so as to not complicate attribute release issues.

Feedback or comments:
The NSF is now suggesting that they want to know the name/email address of each researcher using InCommon. This could be an interesting policy way of encouraging uptake of FIM. Bob Jones suggested that in Europe this could become a condition of H2020 funding (just as now projects must include a data management plan – they could be forced to include a FIM plan).
- *ECP is seen as necessary for non-web apps on both sides of the Atlantic but it is not foreseen to be deployed by most IdPs anytime soon.*
- *In summary US & EU have very similar problems and are taking a similar approach.*

EGI - Long tail of science:
Peter Sologna presented the EGI services for the long tail of science, aiming at preventing the fragmentation of resources. Science gateways are already in production and support for per-user sub-proxies should be available in a couple of months.

Trust and security issues - Sirtifi/SCI:
Dave Kelsey presented the outcome of the work on Trust and Security for FIM (Sirtfi/SCI), which was requested by the FIM4R communities during the previous meeting.
The key outcome is the Sirtfi document available at http://goo.gl/2xnf2G.
Sirtfi aims at seeking wider discussion and feedback from FIM4R and REFEDS.
The document proposes a Security Incident Response Trust framework for Federated Identity and covers operation requirements on traceability and incident response.
Sirtfi will actively work on producing a public "version 1" document, which will be published through the REFEDS process. The group has excellent participation from the US (InCommon) and now needs more input from Europe. Further feedback from the research communities would be very welcome.

Feedback or comments:
- *There are already multiple tools and systems to securely share incidence response information and the intention is to re-use existing tools but we need to agree on which ones. As scale gets larger then some form of automation will be necessary.*
- *NREN CERT teams have similar problems and hence could potentially use same tools. NREN CERT focuses on hosts and IPs, while federations concentrate on identity aspects. Both are very complementary.*

- *The communities requested concrete examples on how Sirtfi would help during incident response, and how it will actually works. Going through various use cases during the next FIM4R workshop would be very useful.*
- *Q: What will happen when Sirtfi has to interact with commercial IdPs? Most security teams, for example in NRENs and large organisations, have links to commercial SPs/IdPs, though it is difficult to have a formal agreement with them.*

Technical Discussion - Levels of assurance:
Marcus Hardt gave a short overview of the different levels of assurance for identity.

Feedback or comments:
- *KIT has large uptake for ECP because the students all want to use a 'dropbox' like function which depends on ECP. This is good news to help with the adoption of ECP for other non-Web use cases.*

Technical Discussion - Geant Data protection code of conduct:
Mikael Linden presented the GEANT Data Protection Code of Conduct (CoC). The goal is to encourage the EduGAIN Service Providers to adopt this CoC, so that EduGAIN Identity Providers would feel more confident in releasing user attributes to the Service Providers. The CoC has been endorsed by the CLARIN, DARIAH, DASHIS, ELIXIR and WLCG communities. It has also been submitted to the EC Article 29 Working Party.
In addition, an international CoC for non-EU IdPs/SPs is actively being worked on.

Feedback or comments:
- *It would be very useful to get the whole InCommon community aware of this activity.*

Technical Discussion - How to Use Entity Categories to Get More and Better Attributes:
Lukas Hammerle presented the Entity Categories concept and how they can help the research communities. Specific Entity Categories can help IdPs decide to when and how to release attributes to SPs, as well as group or tag SPs to facilitate filtering.
They can also be used to highlight IdPs that follow an agreed-on incident response policy (e.g. Sirtifi). While the technical mechanisms are available, some Entity Categories, like Research & Scholarship, are still not widely deployed.

FIM4R pilots' status update:

- ELIXIR: will have a single proxy IdP (like Umbrella). The Interface with e-infrastructures is seen as important and where the EGA data archive is important. It will use clouds around Europe and hence will need FIM.
- ESA: will make use of a commercial partner (Siemens) and support ECP. Shibboleth V3 is also being investigated. Some limitations have been found with the current Shibboleth profile and LDAP back-end.
- WLCG/HEP: The pilot involves transforming SAML into X.509 and is also using Microsoft ADFS (unusual configuration). Major generic issue: no persistent identifier available to uniquely identity eduGAIN users over time. This possibly a blocking issue for the adoption of eduGAIN in production.

- DARIAH: Pilot uses LDAP and SAML. There is a struggle to get wider adoption of eduGAIN or federations among participating DARIAH organisations. Attributes release is also a major issue. The infrastructure provides a homeless-IdP.
- Umbrella: Investigating both eduGAIN and Moonshot. Good progress on the pilot. Authorization is delegated to other entities, just like in WLCG/HEP.
  (Some in the audience commented that Moonshot requires too many and intrusive modifications of the existing applications, including clients to aim at wide adoption at this stage).

What should eduGAIN/AARC do?

- ELIXIR: Improve the IdP releasing of attributes to SPs, and provide training material focused on SPs and users (not just IdPs)
- ESA: Provide clear FIM contact points for communities and international organisations. Some stakeholders are spread across several countries and are not attached to a NREN.
- WLCG/HEP: eduGAIN absolutely needs to provide unique, non-reusable identifiers for each user. Operational security and incident response activities , like Sirtifi, are also essential.
- Umbrella: A Single Logout would be useful. Improving the attribute release situation is also very important.

Conclusions

- A key aspect of future e-infrastructures is the underlying assumption of federated identity services.
- Agreements on interoperation, attributes (definition, release), policies, between all major parties (US, EU, etc.) are essential and urgent for progress and uptake.
- Large science projects are global collaborations now – agreement on a way to harmonize basic identity management is essential.

- Common issues among the research communities:
  - Too many 1-by-1 negotiation to solve issues (technical, bugs, attributes not shared etc.)
  - IdPs and SPs still have no obligation or agreement, moral or legal, to notify other of a security incident
  - Pilots ignore LOA now so as to not complicate attribute release issues
  - Lack of participations from the SPs

- Feedback and main points to address for eduGAIN and AARC:
  - **International Authentication**
  - **Attribute harmonization**
  - **Unique, non-reusable identifier for each user**
  - **Attribute management for Authorization**
  - **Wider adoption of the CoC**
  - **Outreach, training material for SPs**
  - **Clearer contact points in eduGAIN**
  - **Non-web use case**

  - **LoA (ability to express it)**
  - **Logout functionality**
  - **Take into account lessons learnt from GN3+**

- Actions for next FIM4R
  - Lots of expectations on eduGAIN and AARC progress report by next meeting?
  - Investigate “large VO entity categories?” (Alternative to attribute aggregation?)
  - Present some practical examples using the Sirtfi framework
  - Report on non-EU Data protection CoC

- Next meeting
  - Elixir-NL is considering hosting the 9th FIM4R and co-locate with an AARC meeting.
  - TNC2015 is probably too full to be able to co-locate FIM4R with the event. Same for RDA where the focus is more on high-level aspects (informing others of the work) and working with American and Australian participants.