Ongoing Challenges in the VO Space

Table of Contents

Introduction

Virtual organisation (VOs), in particular research collaborations, are one of the strongest use cases for identity federation. Membership is driven by the research more than by geopolitical boundaries. Academic institutions, National Research and Education Networks (NRENs), and funding agencies all encourage the collaboration between researchers; they share the goal of extending the boundaries of human understanding. With science and research as the ultimate driver, the campuses and institutions providing the ICT infrastructure to research group members must do what they can to make their services support rather than impede the science, and a big part of that is to make authentication and access something that happens without the need for replicated ICT infrastructure within each and every research group. In other words, identity federation.

The research collaborations have been fairly clear in what they require and in offering strong recommendations regarding how support services can be improved, and federated identity management is at the heart of those recommendations. Groups like the Federated Identity Management for Researchers (FIM4R) pulled together the requirements and core use cases in a white paper, now hosted at CERN. From the abstract:

> *Federated identity management (FIM) is an arrangement that can be made among multiple organisations that lets subscribers use the same identification data to obtain access to the secured resources of all organisations in the group. Identity federation offers economic advantages, as well as convenience, to organisations and their users. For example, multiple institutions can share a single application, with resultant cost savings and consolidation of resources. In order for FIM to be effective, the partners must have a sense of mutual trust. A number of laboratories including national and regional research organisation are facing the challenge of a deluge of scientific data that needs to be accessed by expanding user bases in dynamic collaborations that cross organisational and national boundaries. Driven by these needs, representatives from a variety of research communities, including photon/neutron facilities, social science & humanities, high-energy physics, atmospheric science, bioinformatics and fusion energy, have come together to discuss how to address these issues with the objective to define a common policy and trust framework for Identity Management based on existing structures, federations and technologies. This paper will describe the needs of the research communities, the status of the activities in the FIM domain and highlight specific use cases. The common vision for FIM across these communities will be presented as well the key stages of the roadmap and a set of recommendations intended to ensure its implementation.*

In response to the FIM4R paper, and coming at the support issues from a more NREN and funding source model, the "Advancing Technologies and Federated Communities" paper, published by TERENA (now GÉANT) in 2012, described a slightly different set of recommendations around technology, policy, funding, and legal issues. These recommendations supported the message out of the FIM4R paper: Federated technologies are key, and the infrastructure needs to be improved to take advantage of those technologies. The existing models of authentication and authorisation no longer scale. This work is not easy, and campuses, NRENs, funding bodies, and legal bodies need to work together to make identity federation easier to do.

Those papers were published three years ago. Unfortunately, the problem space has not progressed very far; the recommendations made are still waiting to be fulfilled. The one key item that would make identity federation a truly compelling story for support institutions and research groups is the release of attributes. With data protection and privacy laws unclear or in flux, campus auditors, security officers, and registrars are uncertain of how to protect the safety and integrity of their institution's data. This short paper provides a view in to where progress has been made, what the limitations have been, and what the support community can do to realize the goal of identity federation and support for virtual organisation.

Progress Made Since 2012

Entity Categories

Technology providers, such as REFEDS, have approached the biggest problem of attribute release by creating Entity Categories. Entity Categories, such as the Research & Scholarship (R&S) category defined by REFEDS, provides a clear structure that allows Identity Providers (IdPs) to make attribute release decisions based on purpose or some other common set of criteria. In the case of R&S, for instance, the community felt it would be easier to encourage IdPs to release attributes when there was some assurance that they would be used for purely academic purposes.

From the entity categories' purpose statement on the REFEDS wiki:

> *Entity Categories group federation entities that share common criteria. The intent is that all entities in a given entity category are obliged to conform to the characteristics set out in the definition of that category.*
>
> *While Entity Categories have multiple potential uses, they were initially conceived as a way to facilitate IdP decisions to release a defined set of attributes to SPs without the need for detailed local review for each SP. The decision by the IdP would instead be based on the criteria detailed in each SP entity category specification. Categories were also conceived for IdPs to indicate support for the SP categories; SPs would use this information to tailor discovery and other aspects of the user experience.*

The entity category model is sound, but there is still a long way to go to see this fully embraced by the community. In the eleven months since the publication of the R&S specification in 2014, for example, 43 out of 1440 IdPs in eduGAIN, the interfederation service, have indicated that they support use of R&S.

Data Protection Code of Conduct

To further support a sense of trust and to address certain liability concerns, a team came together to draft the Data Protection Code of Conduct. From their wiki page:

> *The Data protection Code of Conduct describes an approach to meet the requirements of the EU Data Protection Directive in federated identity management. The Data protection Code of Conduct defines behavioral rules for Service Providers which want to receive user attributes from the Identity Providers managed by the Home Organisations. It is expected that Home Organisations are more willing to release attributes to Service Providers who manifest conformance to the Data protection Code of Conduct. For more information, see Introduction to Code of Conduct.*

While this self-asserted set of information is targeted towards organisation concerned with the EU Data Protection directive, it is a model that other regions may choose to follow as a way to help build understanding and trust in how information can and should be handled by the Service Providers (SPs) based in their region. As with the entity categories, however, there is a long road to adoption. As of October 2015, 69 of the 991 SPs registered in eduGAIN assert that they follow the Data Protection Code of Conduct.

Authentication and Authorisation for Research and Collaboration (AARC)

Funding agencies in the EU are supporting efforts in the identity space by providing money and structure to coordinate identity federation efforts. The AARC project is a multi-year effort, started in 2015 and funded by the European Commission, to help develop and mature the ICT infrastructure required by academia and research groups in Europe.

> ***AARC is an EC funded project that brings together 20 different partners among National Research and Education Networks (NRENs), e-Infrastructures Service Providers and libraries to develop an integrated cross-discipline AAI framework, built on production and existing federated access services.***
>
> *The AARC project vision is to avoid a future in which different e-Infrastructures and (new) research collaborations develop and operated independent (and not inter-operable) AAIs.*

Still in its early stages, the project is focusing on verifying requirements and assessing the landscape to see what tools already exist to support research and academia so as to improve and integrate them rather than rebuilding where it is not necessary. For the latest on the AARC project, see their website.

US National Science Foundation (NSF)

In the US, federal agencies like the NSF offer grants for targeted development in the cyberinfrastructure space. While these grants are not limited to the identity federation space, work like CILogon 2.0 aims to provide a strong identity and access control platform that will meet the needs of the research community. Previous awards, including the recently concluded Software Development for Cyberinfrastructure (SDCI) grant that funded the COmanage project, shows a history of commitment to improving the tools that are necessary for further deployment of the identity federation model.

While funding support such as what's being provided by the EC and by the US is critical for any forward progress in the identity federation space, until laws and policies are clear at the same levels identity federation is expected to function, we will continue to see barriers to deployment that cannot be overcome with technology and money.

The VO Barriers

With technology providers and funding agencies providing the resources needed to make some forward progress in the identity federation space, there are still a few key issues that make VOs a unique support challenge. Issues such as the legal status of a VO: they are often not legal entities at all, and therefore cannot sign contracts, agreements, or MoUs–and not having IT staff that can overcome the high bar for technical knowledge to deploy some of the more challenging tools and applications in this space.

It appears to be a intractable problem: without an immediate and obvious solution that lessens their support burden in the identity management space, VOs will not put the time and resources into making identity federation work for them. Without VOs pushing the problem, support agencies will be unable to argue for the resources they need to push for clearer laws, policies, and technologies. And without clear laws, policies, and technologies, there is no immediate solution for VOs.

Gaining Traction

> "When eating an elephant take one bite at a time" – *Creighton Abrams*

When presented with an apparently intractable problem, one needs to consider what parts of that problem can be moved at all. In the case of identity federation, the fundamental requirement to make this concept work for research collaborations and others is attribute release. Attribute release will not happen until other areas in the problem space shift, such as the policy and legal aspects. But while that very massive problem blocks the road, other requirements can still be matured.

Making the Technology Easier

In the US, Internet2 is coordinating funding for the Trust and Identity in Education and Research (TIER) project. Internet2, with grant and institutional support, has helped sponsor work on Shibboleth, Grouper, and COmanage. Each of those platforms requires strong system administration and possibly programming skills in order to deploy and support; often more skill than can be found in a small research group. The TIER project seeks to make these tools more easily deployed, thus lowering that technical bar of knowledge required to get started in the identity management space.

As mentioned earlier, AARC is also looking to improve the identity and access management space by improving the tools already used by authentication and authorisation infrastructures around Europe. In addition, they have in their mandate a strong training component to help organisation learn more about identity management, the existing tools, and the best practices in the identity federation space.

Improving Security in a Federated World

One of the areas of concern in the identity federation space is security. If one institution, through participation in an identity federation, is trusting the information from another, they need some information when the information they are using has been compromised. While institutions often have security policies for incidents within their own, constrained environment, there needs to be further work on how to share that information beyond the institutions borders.

Work is underway through the Security Incidents Response Trust Framework for Federated Identity (SIRTFI) to help fill that gap in the security model. Draft guidelines are under development, but additional support–particularly in the form of institutions agreeing to follow that framework–is necessary.

Providing a Value Proposition

Establishing identity federation is a global challenge. While some parts of the world have been working in this space for over fifteen years, other regions are just getting started. To help those regions–or any group that is still trying to build the business case for why their institution or group should provide resources–a clear value proposition is required. At the 40th Asia Pacific Advance Network (APAN) meeting in March 2015, the discussion in the new APAN IAM Task Force made this need quite clear.

This became a work item for REFEDS, and a draft was presented at the REFEDS meeting in October 2015. A final document is expected by the end of 2015, though updates will be incorporated as feedback is received.

Conclusion

The research collaboration model is just one facet of the VO space. VOs can be intergovernmental agencies. They can be public-private collaborations. They can be inter-departmental teams on a single campus. If the technology providers, the funding agencies, and the research community can work together to solve the challenges for research collaboration, we have solved most of the issues for a much broader community. While attribute release remains the most critical issue to resolve, we can still make progress on making the technology underlying identity federation and the best practices around policy and security more accessible to a broader audience.