View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by KITopen

EPJ Web of Conferences **214**, 03044 (2019) https://doi.org/10.1051/epjconf/201921403044
CHEP 2018

Federated Identity Management for Research

Thomas Barton[1], *Peter* Gietz[2], *David* Kelsey[3], *Scott* Koranda[4], *Hannah* Short[5,*], and *Uros* Stevanovic[6]

[1]University of Chicago and Internet2
[2]DAASI International, Germany
[3]STFC UK Research and Innovation, Rutherford Appleton Laboratory, Didcot, United Kingdom
[4]LIGO
[5]European Organisation for Nuclear Research (CERN), Geneva, Switzerland
[6]Karlsruhe Institute of Technology, Germany

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. In many research communities there is an increasing interest in a common approach to FIM as there is obviously a large potential for synergies. FIM4R [1] provides a forum for communities to share challenges and ideas, and to shape the future of FIM for our researchers. Current participation covers high energy physics, life sciences and humanities, to mention but a few. In 2012 FIM4R converged on a common vision for FIM, enumerated a set of requirements and proposed a number of recommendations for ensuring a roadmap for the uptake of FIM [2]. In summer 2018, FIM4R published an updated version of this paper [3]. The High Energy Physics (HEP) Community has been heavily involved in creating both the original white paper and the new version, which documented the progress made in FIM for Research, in addition to the current challenges. This paper presents the conclusions of this second FIM4R white paper and a summary of the identified requirements and recommendations. We focus particularly on the direction being taken by the Worldwide LHC Computing Grid (WLCG), through the WLCG Authorisation Working Group, and the requirements gathered from the HEP Community.

1 Introduction

Federated Identity Management (FIM) is an evolving set of technologies, policies, and services that national Research & Education (R&E) Federations implement to produce a global trust infrastructure for the R&E sector that enables login to (federated access to) protected resources with users' home organisation credentials. FIM4R [1] (FIM for Research) is a collection of individuals from research communities, research cyber infrastructures that support them, and R&E Federations with a shared interest in enhancing how R&E Federations and research cyber infrastructures integrate to support the work of research communities. A key measure of success for an R&E Federation is its uptake among research and academic communities. To promote this, FIM4R collects and rationalises requirements bearing on

[*] e-mail: hannah.short@cern.ch

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).

technical architecture, services, standards, and operational policies needed to produce harmonious integration between research cyber infrastructures and R&E Federations. These requirements may apply to R&E Federations, the research cyber infrastructures, or proxies, portals, gateways, and other components that link them together. FIM4R members collaborate with organisations in both domains, R&E Federation and research cyber infrastructures, to help implement these requirements.

A version 1.0 white paper [2] was produced in 2012 to document common requirements, a common vision, and recommendations. During 2017 and 2018, FIM4R members collaborated on a second white paper to highlight the progress since 2012 and update requirements and recommendations to reflect current and anticipated trends and challenges.

2 Federated Identity Management for Research White Papers

2.1 FIM4R version 1.0, 2012

The version 1.0 white paper was the result of gathering input from a variety of research communities and cyber infrastructures and distilling common high level requirements and recommendations from the input. These include:

- User friendliness
- Federated access enabled with and without need for browsers
- Support for the multiple credential types in use by research cyber infrastructures
- Support for multiple Levels of Assurance to match different levels of risk associated with protected resources
- Control over authorisation by research communities or research cyber infrastructure operators
- A well-defined set of user attributes that are appropriately available across the entire architecture
- Risk assessment, traceability, and a security incident response capability suited to this environment
- Transparency of the various policies by which organisations manage their users' federated credentials
- Reliable and resilient operations
- Scalable means by which to address legal, policy, and trust concerns with ensuring suitable security and privacy across this international and heterogeneous infrastructure

Substantial developments on all of these have occurred since the version 1.0 white paper's publication. Highlights include:

- Development and implementation of the Research & Scholarship Category [4], a globally adopted program that defines a set of user attributes and helps to manage user privacy by disclosing them only to federated services independently vetted to be purposed for research or scholarly use.
- Development and implementation of Sirtfi [5], a security incident response framework adopted by R&E Federations, Snctfi [6], a suite of policies that facilitate successful integration of cyber infrastructures with R&E Federations, and the GÉANT Data Protection Code of Conduct [7] to address data privacy compliance needs of federated organisations in the EU.

- Various solutions to non-browser federated access needs.
- Experiments with defining and fielding solutions to Level of Assurance needs [8] [9].
- Emergence of a proxy architecture [10] as the approach taken by multiple research cyber infrastructures to simplify their integration with R&E Federations.
- European Commission funding for the AARC and AARC2 projects [11], which convened and focused FIM4R members and others on identifying means to address the various requirements of the version 1.0 white paper, whose efforts helped to produce several of the above items and more.

2.2 FIM4R version 2.0, 2018

Although the developments listed above are quite substantial, they have not fully addressed the problems at which they are aimed. Most R&E organisations do not yet participate in the Research & Scholarship Category [4] or Sirtfi [5] programs, and the Data Protection Code of Conduct [7] has had to be revised in light of the General Data Protection Regulation [12], a process that is incomplete and whose outcome is as yet uncertain. The InCommon Federation in the United States developed its Bronze and Silver Levels of Assurance but discovered that most of its member organisations found them too onerous to implement absent evidence of uptake by resource providers, and resource providers similarly did not rely on them because of, among other reasons, insufficient uptake by users' home organisations.

Whereas in 2012 many practitioners envisioned that every service in a research cyber infrastructure would directly join an R&E Federation, experience has shown that it is more practical and scalable to implement a proxy for them that is joined to R&E Federation. This model is articulated in the AARC Blueprint Architecture [10] and centralises credential translation, authorisation management, and other functions in one place, avoiding the need to do so in each service within a cyber infrastructure and join it to R&E Federation. A proxy helps to mitigate the shortcomings of the Research & Scholarship Category program by providing an alternate locus for managing needed user attributes. More generally, standards, technologies, architectures, and services have evolved over six years, as has what practitioners can envision. There are now good open source proxy platforms that address these needs, and services such as ORCID [13], for example, that provide new approaches to meeting some of the needs of research communities.

In early 2017, FIM4R members determined that it was time to look anew at how integration of FIM and research cyber infrastructures should continue to evolve and began a new cycle of gathering input from research communities and cyber infrastructures. Representatives of 14 research fields across physics, astronomy, climate and planetary science, life sciences, infectious diseases, and humanities, to name but a few, and their supporting research cyber infrastructures, provided input. Five face-to-face meetings focused on this effort took place in Europe and North America. At three of them presentations by research communities and cyber infrastructures were heard, followed by discussion to appropriately integrate their specific requirements within a single catalog. At another meeting sets of specific requirements were assigned to break-out groups to reconsider whether they were the right requirements, which led to some requirements being removed, others merged, and sharpening of the language used to express those remaining. At the final face-to-face meeting, research communities were asked to endorse requirements, ensuring that the published list reflected genuine needs.

An editorial team was established to complete the final paper, which was published in June 2018 on Zenodo[14], in line with the group's affiliation with Open Research. Members of the editorial team prepared a set of presentations at meetings in Europe (TNC18, RDA, CHEP), North America (Internet2 Technology Exchange), and Asia Pacific (ISGC) to inform the wider community and seek further input.

2.2.1 Summary of Recommendations

The FIM4R white paper version 2.0 [3] has identified sets of recommendations that the authors strongly believe is beneficial to the academic community. As in version 1 [2], the paper starts with a comprehensive sets of requirements that were identified together with all the stakeholders. These requirements address, among others, attribute release, web and non-web access and technologies, security, authentication, and authorisation. An expansive analysis of requirements, both by their distribution (i.e. who has expressed particular requirements) and their importance (i.e. how important are the requirements), has produced a set of recommendations aimed to address the identified problems and to increase the uptake of FIM. The paper further seeks to facilitate the adoption of best practices by mapping the identified concerns to relevant groups and stakeholders best suited to address them.

Governance and coordination Representation of researchers and research e-Infrastructure operators within large Infrastructures should be increased. This would help to ensure the continued alignment of interests of the researchers with the intended mission of large Infrastructures, i.e. supporting research and scholarly aspects. It is also essential that the FIM services are operated sustainably, reliably, and with the appropriate user support. Due to the ever changing technological and operational environment, providing a forum, owned and attended by research communities, where common issues can be discussed is beneficial.

Baseline of Research User Experience This section identifies several well-established practices, of which increased adoption would significantly boost the usability of federated access mechanisms. Releasing a sufficient set of attributes (as identified in R&S entity category [4]) would increase the value of federations, and reduce the impediments the researchers face in accessing remote services. Better harmonisation of import/export practices by R&E operators of their metadata is recommended, and also providing a process through which certain research organisations (that are not legal entities) can be admitted, either at the national or international level. Usability can further be increased by better presenting errors to the users and by providing means to support user mobility (e.g. ORCID [13]).

Security Incident Response Readiness All participants in FIM and federations should support best practices for operational security (such as Sirtfi [5]). Having a security incident response plan is strongly recommended, as well as periodic testing of such abilities. The issue of cooperation for security purposes is also recognised and encouraged by legislation, e.g. Recital 49 of the GDPR [12].

Harmonisation of Research Community Proxy Operations and Practices As previously mentioned, proxies have emerged as an answer to needs unmet by R&E federations. They have by now matured as a solution, hence their stability, support, and sustainability is becoming significantly important. This includes following the identified best practices, such as the AARC Blueprint Architecture [10] and related guidelines, and reuse of certain AAI services, when applicable.

Sensitive Research User Experience Employing strong controls of authentication and access management is paramount. This is necessary, for example, to ensure confidentiality of restricted research data, integrity of basic research data, or fine grained access to expensive instruments and computing resources. REFEDS Assurance Framework [8] in its first version has issued guidance responding to these needs, and it is encouraged that these instructions are applied by relevant parties.

3 Federated Identity Management for High Energy Physics

3.1 HEP input to FIM4R

The High Energy Physics Community, was one of the founding research communities of FIM4R, represented by the Worldwide Large Hadron Collider Computing Grid (WLCG) [15]. This activity has been sustained by individuals from CERN and STFC-RAL who ensured that requirements specific to the field were taken into account in FIM4R version 2.0. An extended paragraph on WLCG requirements is included in the white paper and highlights the following points as priorities for FIM integration for critical services:

- Adoption of security frameworks such as Sirtfi, R&S and Assurance profiles
- Mature, tested Security Incident Response tools and procedures and their widespread adoption
- Strengthened operational support for Identity Federations and Interfederation including help desks and guidance
- Sustainable operation of shared components
- Identity Federation support for tokens beyond SAML, notably OpenID Connect (OIDC)

Insights gathered thanks to WLCG's participation as a key use case in the Helix Nebula Science Cloud Pre-Procurement Project [16] were also included in the FIM4R version 2.0 white paper. These are particularly relevant as HEP, the nuclear physics communities and other research communities begin to seriously explore the potential of commercial cloud providers and their integration in Authentication and Authorisation models.

3.2 Ongoing Projects

Although WLCG has been successfully using X.509 certificates for many years, there is increasing interest to take advantage of non-X.509 identities, both for user Authentication (through SAML Identity Federations and standalone OIDC providers) and for Authorisation within the grid (through OAuth2, OIDC and Macaroons). Multiple proof-of-concept implementations and projects (e.g. ALICE Tokens, SciTokens) have begun to explore the ways in which the community can benefit from token based access control.

A WLCG Authorisation Working Group has been established with two primary aims:

1. Identify an appropriate solution in line with the AARC Blueprint Architecture to control access to WLCG services

2. Define a common schema for tokens issued and exchanged by such a solution

This group is working towards the aims for Authentication and Authorisation identified in the HEP Software Foundation Community White paper [17]. Critically, the group includes participants from many of the ongoing projects, with the aim to facilitate interoperability in the long term. A principle of the WLCG Authorisation Working Group is to follow standard industry and R&E practices, FIM4R version 2.0 and the AARC Guidelines [18] being important assets to consider.

4 Conclusion

The publication of FIM4R version 2.0 [3] in June 2018 drew significant interest from the wider community. As of November 2018, the paper has been viewed approximately 800 times online and has been presented at multiple international conferences. In addition, major organisations that support R&E Federation such as Internet2 [19], GÉANT [20], and REFEDS [21] are already taking the version 2 findings into account as they plan their further activities.

A version 2.1 of the white paper is envisioned in which input received too late for the version 2 paper can be incorporated. The Research Data Alliance [22], in particular, may provide a potential link to additional research communities who could contribute to, or learn from, the experiences of the FIM4R community.

The FIM4R version 2.0 white paper has enabled the High Energy Physics to express its needs to FIM stakeholders. Conversely, FIM4R also provides a forum for following the evolving best practices of the field and the results of FIM4R version 2.0 are a key reference for the WLCG Authorisation Working Group.

References

[1] *Federated Identity Management for Research, FIM4R*, https://fim4r.org/

[2] D. Broeder, B. Jones, D. Kelsey, P. Kershaw, S. Lüders, A. Lyall, T. Nyrönen, R. Wartel, H.J. Weyer, Tech. Rep. CERN-OPEN-2012-006, CERN, Geneva (2012), `https://cds.cern.ch/record/1442597`

[3] C.J. Atherton, T. Barton, J. Basney, D. Broeder, A. Costa, M. van Daalen, S. Dyke, W. Elbers, C.F. Enell, E.M.V. Fasanelli et al., *Federated Identity Management for Research Collaborations* (2018), `https://doi.org/10.5281/zenodo.1307551`

[4] *Research and Scholarship - REFEDS*, `https://refeds.org/research-and-scholarship`

[5] N. Harris, T. Barton, J. Basney, D. Groep, L. Johansson, D. Kelsey, S. Koranda, R. Wartel, A. West, H. Short, *Security Incident Response Trust Framework for Federated Identity (Sirtfi)* (2015), `https://doi.org/10.5281/zenodo.1256531`

[6] *Snctfi - the Scalable Negotiator for a Community Trust Framework in Federated Infrastructures*, `www.gridpma.org/snctfi/`

[7] *Data Protection Code of Conduct Home - REFEDS wiki.*, `https://wiki.refeds.org/display/CODE/Data+Protection+Code+of+Conduct+Home`

[8] *Refeds Assurance Framework*, `https://refeds.org/assurance`

[9] *Incommon Assurance*, `https://www.incommon.org/assurance/`

[10] *AARC Blueprint Architecture – AARC*, `https://aarc-project.eu/architecture/`

[11] *Authentication and Authorisation for Research and Collaboration, AARC*, `https://aarc-project.eu/`

[12] *EU General Data Protection Regulation*, `https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1532348683434&uri=CELEX:02016R0679-20160504`

[13] *ORCID*, `https://orcid.org/`

[14] *Zenodo*, `https://zenodo.org/`

[15] *Worldwide LHC Computing Grid*, `http://wlcg-public.web.cern.ch`

[16] *HNSciCloud*, `https://www.hnscicloud.eu/.`

[17] HEP Software Foundation including, J. Albrecht, J. Alves, Antonio Augusto, G. Amadio, G. Andronico, N. Anh-Ky, L. Aphecetche, J. Apostolakis, M. Asai, L. Atzori et al., ArXiv e-prints arXiv:1712.06982 (2017), `1712.06982`

[18] *Guidelines - AARC project.*, `https://aarc-project.eu/guidelines/`

[19] *Internet2*, `https://www.internet2.edu/`

[20] *GÉANT*, `https://www.geant.org/`

[21] *REFEDS*, `https://refeds.org/`

[22] *Research Data Alliance*, `https://www.rd-alliance.org/`