Sustainability Models for Guest IdPs

Actual Date:	2016-04-26
Grant Agreement No.:	653965
Work Package:	NA3
Authors:	Wolfgang Pempe, DFN

The research leading to these results has received funding from the European Community’s Horizon2020 Programme under Grant Agreement No. 653965 (AARC).

Abstract

This document gives short overview of sustainability models for Guest IdPs. Depending on the respective target group, the organizational and technical context, different operational and cost models are applicable.

Table of Contents

1 Introduction

This document elaborates on sustainability aspects of so-called Guest or Homeless IdPs – Identity Providers for end users who are not able to access (inter)federated services otherwise, like

1. nomadic users (those without a “home” organization, such as “long-tail” researchers),
2. so-called citizen scientists, and
3. users belonging to an institution that is not able to operate an Identity Provider (IdP), or one which operates a stand-alone IdP which is not part of an established federation.

While target groups 1 and 2 can be categorized as “homeless users”, group 3 will be referred to as “IdP-less users”.

For an in-depth coverage of this topic including definition, motivation and existing technical solutions please refer to the AARC deliverable [MJRA1.2] “Guest Identities”.

2 Overview: Components and Roles

Strictly speaking, an IdP is only some software on top of a data source containing user data which is used to shape digital identities. An IdP implementation acts as an intermediary between federated services (Service Providers) and a (in our case: virtual) home organization by authenticating users and translating user data into (usually) standardized attributes[1]. The IdP transfers that information via specific protocols, usually [SAML2]. The technical and organizational context of the interaction between IdPs and Service Providers (SPs) is provided by Identity Federations, most often operated by NRENs. Thus, considerations about the sustainability of IdPs also have to take into account the maintenance and operational aspects of the underlying systems, their components and the human actors (respectively the abstract roles) involved, cf. figure 1.1.

[1] For a more detailed definition, cf. https://shibboleth.net/products/identity-provider.html.

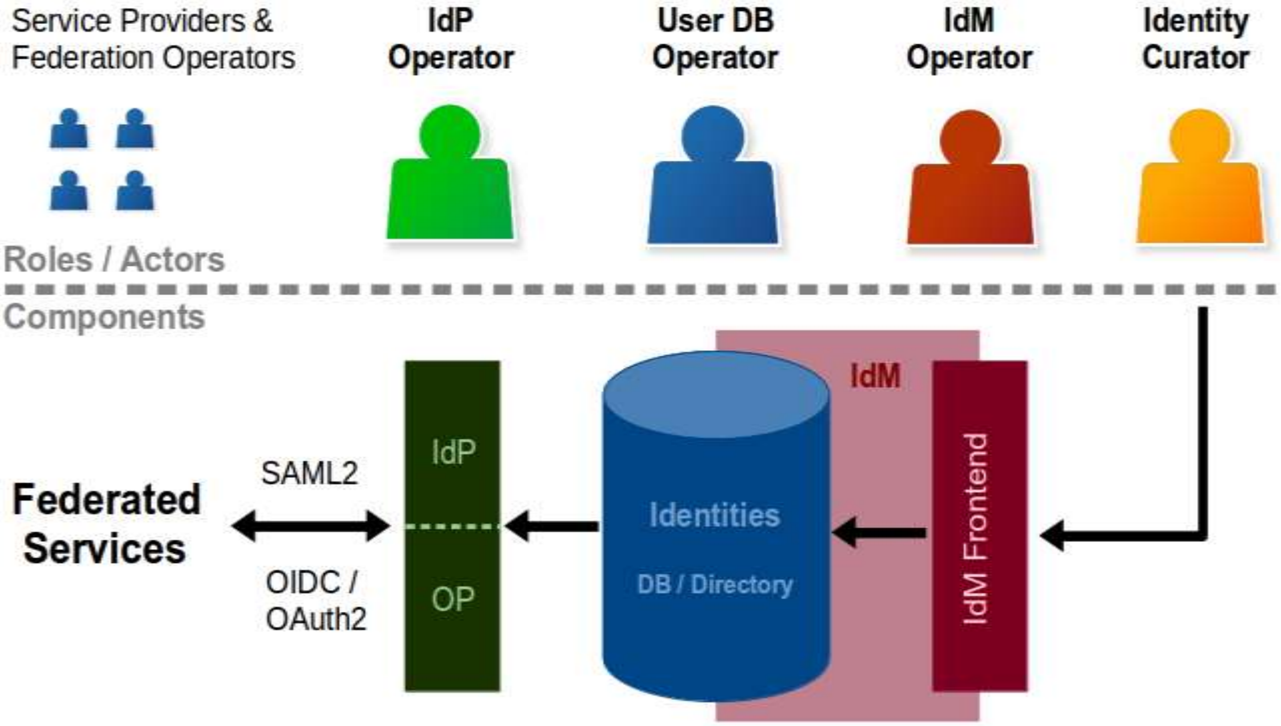

Figure 2.1: Components and Roles

Apart from the IdP software itself, at least the following additional components are required for operating an IdP system:

- A database or directory where user data are being kept
- Some kind of identity management system (IdM) which allows the management, or – more precisely – the “curation” of identities including a user-friendly frontend for human interaction.

All these components could – in theory – be operated by different human and legal entities. Those actors are usually bound to each other or to a trusted third party by contracts and policies. The following section will discuss some operational concepts, i.e. possible constellations of actors/roles and components and the resulting implications in terms of sustainability, both technically and legally.

3 Operational Models for IdP Systems

Considering a classical identity federation, which is typically operated by an NREN, the legal entity acting as federation operator, is the central contractual partner of all participants, both service providers and IdP operators, the latter usually being home organizations. Acting as policy-making authority and legal hub, the federation operator ensures the mutual trust between the participants. Based on this model, the following sections describe in more detail some variants in the relationship between federation operators and the actors involved in operating an IdP, and discuss the applicability of those models to Guest IdPs.

3.1 Home Organization Covering Everything (“all-in-one”)

This is the most common case: All components, IdP, user DB and IdM are operated by one legal entity, for instance a university or research facility which also takes care of the management/curation of identities. Within larger organizations, different departments may fill the different roles, but from a contractual point of view, this is still a point-to-point relationship between federation operator and home organization.

Applying this simple model to the use case of Guest IdP, the only relevant difference is the fact that the end-users are not necessarily members (e.g. students or employees) of the legal entity operating the technical components. In that case, the management/curation of identities takes more effort, as well in terms of identity vetting and keeping account information up-to-date, as in terms of policy-enforcement because the legal relationship between “homeless” end-users and (virtual) home organization is typically not a very close one.

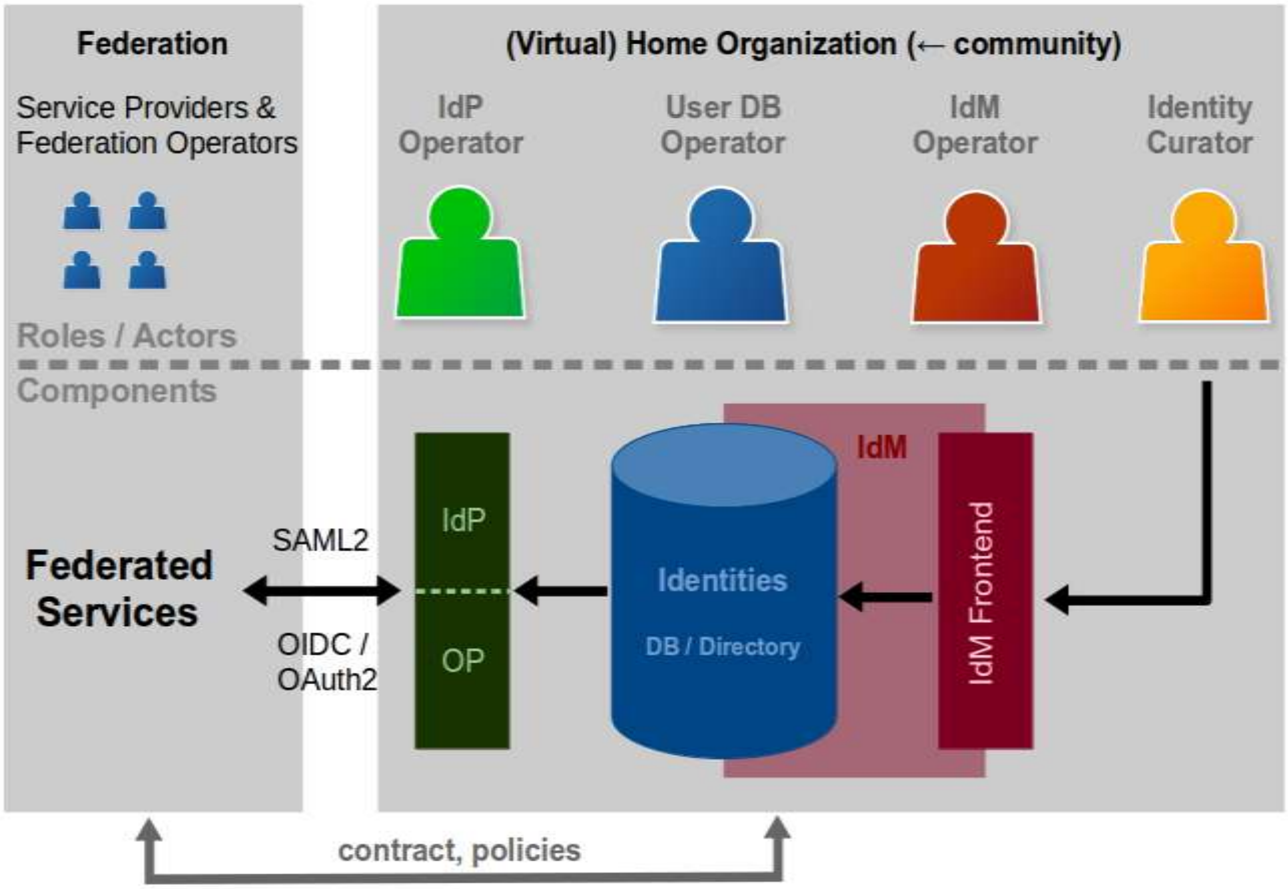

Figure 3.1: HO covering everything

Research communities and projects being in need of a Guest IdP often have the problem that they have to find a legal entity which is able and willing to take over the legal obligations going with joining a federation. Furthermore, it has to maintain the required technical components and take care of the curation of identities (especially with respect to the required Levels of Assurance) – maybe for several years. While long-term (e.g. ESFRI) projects like DARIAH are more likely able to deal with these issues, it is not easy for (non- or loosely organized) research communities and smaller/short-term projects to find some partner within the community or project who is willing to take over all technical, legal and administrative obligations connected with operating a Guest IdP in this way (“all-in-one”), even if funding is granted. Within larger projects, such tasks can also be delegated. In those cases, a trusted third party, like an NREN, GÉANT or a commercial provider may take over at least some of the operational tasks (cf. the following section), while the management of identities is something that should be taken care of by the communities themselves.

An example of this model is the DARIAH Homeless IdP, which is officially registered with the DFN-AAI by the [DARIAH] project partner GWDG Göttingen but technically operated by other DARIAH partners.

3.2 Trusted Third Party as IdP Operator (VHO Service)

While the effort of maintaining the necessary technical infrastructure for operating a Guest IdP is easily manageable, the user management including identity vetting, de-provisioning etc. is a task that should reasonably be performed by the respective user community.

In this model, the technical components are operated by a trusted third party, either the federation operator (VHO as a service) or by another organization providing a Virtual Home Organization (VHO) service and taking over the role of the legal counterpart of the federation operator. This implies that the VHO operator guarantees the respective federation and the outside world that the VHO complies with the federations policies including the “quality” of identities and – possibly – with specific Levels of Assurance. Assuming those liabilities, the VHO operator has to make sure that the Identity Curators for their part take all efforts to ensure the compliance with the federation policies. To this purpose, the VHO usually enters into some kind of written agreement with the legal entities representing the projects and/or communities using the VHO. This kind of VHO is often implemented as a multitenant system so that it may be used by more than one project or community. Such a setup only scales as long as no specific attribute filter rules are needed. Thus, the VHO IdP is expected to release a set of standard attributes including a persistent identifier, which may be used by SPs to retrieve community- or project-specific attributes from Attribute Authorities as e.g. community-specific entitlements.

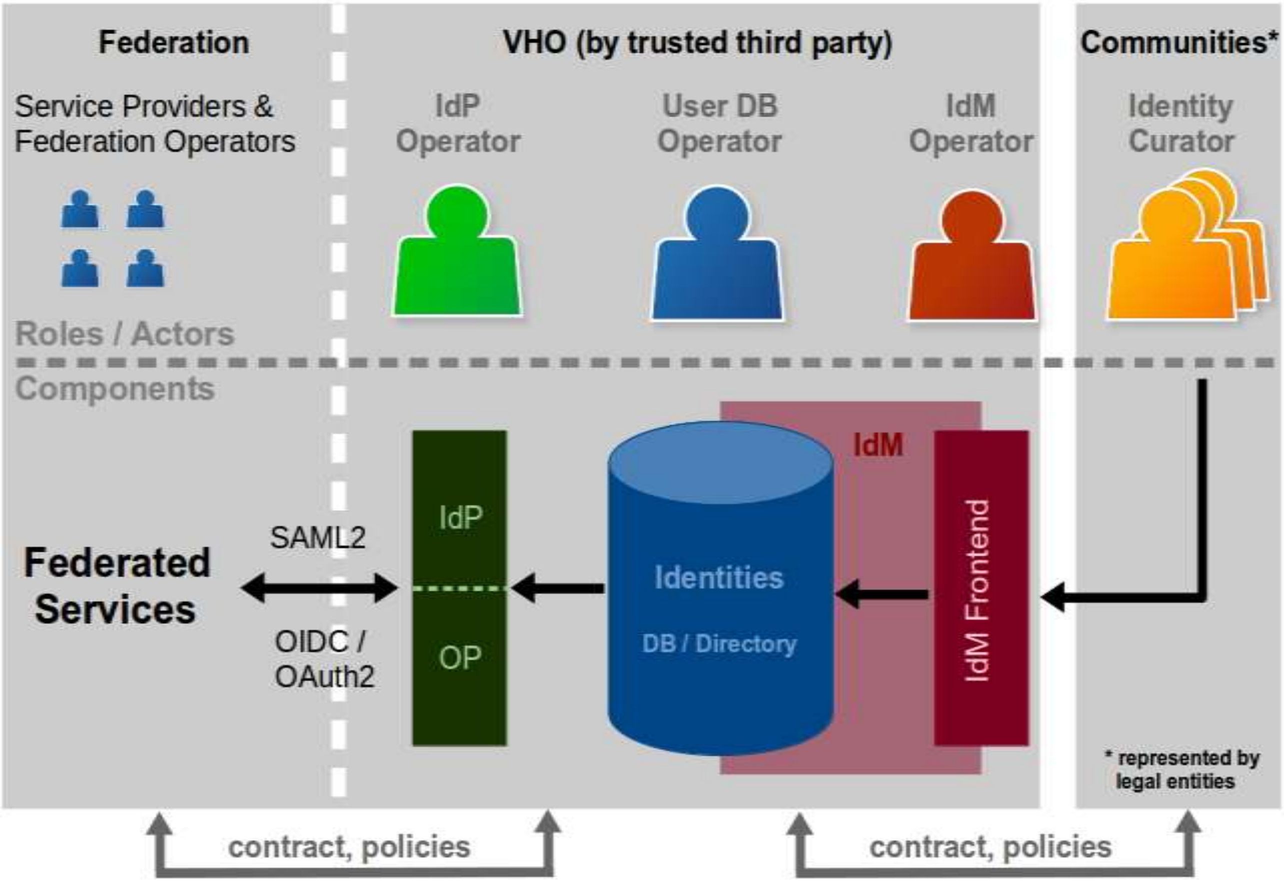

Figure 3.2: VHO by trusted third party

An example of this model is the VHO operated by the “Verbundzentrale des GBV (VZG)”, Göttingen , which represents the so-called Common Library Network of the German States Bremen, Hamburg, Mecklenburg-Vorpommern, Niedersachsen, Sachsen-Anhalt, Schleswig-Holstein, Thüringen and the Foundation of Prussian Cultural Heritage. The latest use case is the user registration and management for the DFG-funded, so-called “Fachinformationsdienste”, [FID] (documentation only in German). VZG and DFN are currently discussing the possibility that DFN takes over the operations of this service (or offer an adequate solution) once funding ceases.

Another example is the VHO Service operated by SWITCH, cf. [SWITCH VHO]. So-called VHO administrators fill the role of the Identity Curator.

3.3 Hosted IdP

Theoretically, every single component listed above could be operated by a separate organization – arbitrary combinations are conceivable. In practise, not every possible constellation makes sense; especially models where too many actors are involved wouldn't scale, neither technically nor from an organizational or juridical point of view. Therefore, we will discuss only one further model, which may not be as common as deployment model for Guest IdPs, but which proved to be quite attractive for smaller and medium-sized institutions in some federations like SWITCHaai or DFN-AAI, the model of IdP hosting. In this scenario, a federation operator or another trusted third party operates an IdP instance on behalf of a home organization which is not able or willing to deal with yet another piece of software which has to be understood, configured, monitored and otherwise maintained. All other components are operated by the home organization itself.

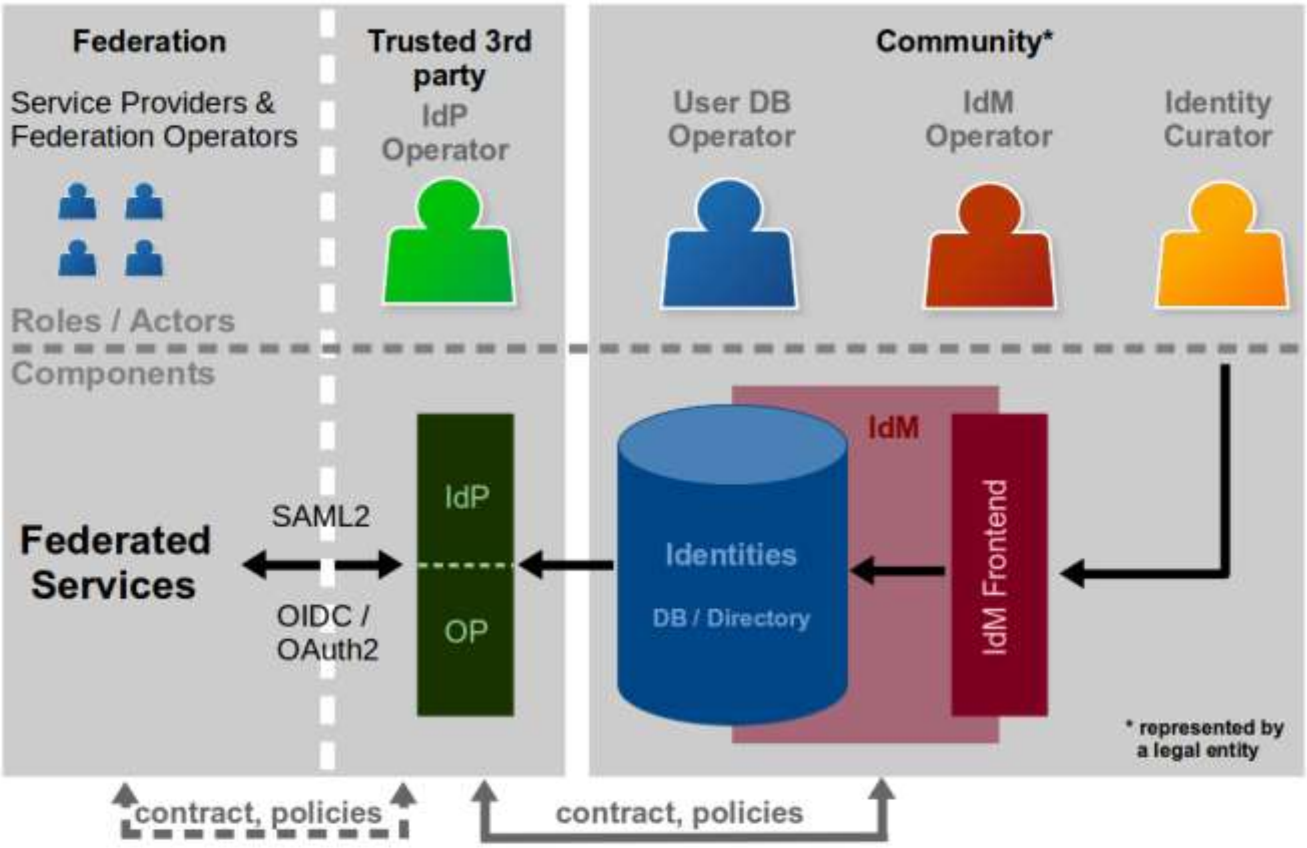

Figure 3.3: Hosted IdP

If we try to apply this model to the target groups listed in the introduction, it only fits to the third one, the “IdP-less” users (“users belonging to an institution that is not able to operate an Identity Provider”). Some federations already provide this kind of service, namely SWITCHaai[2] and DFN-AAI[3]. GARR even goes a step further by providing virtual machines including all required technical components including a simple IdM[4].

4 Cost Models and Maintenance Aspects

Considering the three operational models as outlined above, different possible cost models are applicable, depending on several factors:

- Target group,

[2] Cf. https://www.switch.ch/aai/join/idp-hosting/
[3] Cf. https://www.aai.dfn.de/der-dienst/ausgelagerter-idp/
[4] FIXME – some reference to this service.

- Technical setup,
- “Organizational” setup, i.e. which legal entity fills in what role.

Furthermore, the level of the costs to be covered depends on

- the required degree of reliability of identities and user data, and
- the intended period of maintenance.

Home Organization Covering Everything (“all-in-one”)

This concept fits best for “homeless users” who are associated with a research community, represented either by a project or by some other entity like a consortium. Operational costs would be taken over by the funding body or as in-kind contribution by one or more project partners. In that case, the period of maintenance of such a Guest IdP would most probably be coupled with the project’s lifespan. If operated by a consortium, this kind of setup would be more sustainable. Charging annual fees to end users may be an option, but only in theory. It is most unlikely that researchers are willing to pay for this kind of service.

Trusted Third Party as IdP Operator (VHO Service)

This operational model also rather applies to the target group of “homeless users”. The cost model for this concept depends on the trusted third party that operates the VHO. A commercial provider would need to charge the operational costs to the target groups. In that case, the considerations concerning the “all-in-one” concept would also apply here. If the VHO is operated by a larger non-commercial entity like an NREN or an organization like GÉANT, this kind of service could be considered as part of the overall service portfolio and free of charge at least for end users. Assuming the latter, long-term maintenance would be more likely, too.

Hosted IdP

This solution is most convenient for “IdP-less users” who are members of a home organization that doesn’t operate an IdP. IdP hosting is a well-established service in many federations, in some cases as part of the overall service portfolio (DFN) or as an additional, unbundled service (SWITCHaai) which entails some extra costs. In any case, no fees are charged to end users. However, this kind of service is usually only available for home organizations that are already members of a federation offering this kind of service. Insofar the challenge here is to find a solution for “federation-less” home organizations – perhaps a service to be provided by an entity like GÉANT.

5 Summary

In this document, we have surveyed and discussed several operational and costs models for Guest IdPs. It could be shown that there is no standard solution for the sustainable operation of a Guest IdP. Beginning with the target groups listed in the AARC proposal, “homeless” users and institutions without an (inter-)federated IdP, various operational and cost models are applicable, depending on the individual conditions. In any case, well-established institutional partners, contractual frameworks and long-term funding are important factors for every sustainability model.

References

[DARIAH] https://wiki.edugain.org/DARIAH_UseCase

[FID] http://www.fid-lizenzen.de/ and http://www.fid-lizenzen.de/ueber-fid-lizenzen/dateien/FID-Info_Registrierung_und_Authentifizierung_20141201.pdf

[MJRA1.2] AARC JRA1 Task 1.3: Guest Identities, Deliverable MJRA1.2 (not yet published)

[SAML2] http://www.oasis-open.org/committees/security and http://www.oasis-open.org/committees/download.php/27819/sstc-saml-tech-overview-2.0-cd-02.pdf

[SWITCH VHO] https://www.switch.ch/aai/join/vho/