Account LInking SErvice - ALISE

1. The problem

In intertwin, we are developing services that operate on behalf of the user. These services include managing data ingress/egress and launching HPC jobs that train machine learning models.

We imagine such managing services will operate (for the most part) outside of any particular HPC facility. This is because the users will (typically) use the services of many facilities in concert. Therefore, in order to support these activities, we need to operate HPC facilities “on behalf of” the individual user.

These **users are already known to the HPC facility**. They have already gone through the facility's usual enrolment procedure. Their application for access will have been vetted, according to the facility’s own policies. The users have accepted the facility’s AUP and any other related agreements.

Using the current authentication schemes typically supported by HPC facilities (username + password, SSH public/private key pairs) require an extremely high level of trust between the HPC user and the interTwin central service. This is because they require the HPC user to share highly sensitive material (a password or an SSH private key) currently with a central service. Such sharing of sensitive material could even be in violation of the HPC facility’s policies, which the user agreed to when enrolling.

Moreover, the recent trend toward adopting two-factor authentication (2FA) or multi-factor authentication (MFA), makes such approaches currently impossible. This is based on the assumption that the HPC user cannot share the second factor with the central services.

2. The proposed solution

The proposed solution involves the adoption of an industry-standard approach to sharing access: OpenID-Connect (OIDC). This is widely used at scientific and commercial infrastructures, such as WLCG, ELIXIR, “Erasmus Without Paper”, Google, Facebook, and GitHub.

In more detail, we anticipate the HPC facility recording the user’s federated identity, much in the same way as many facilities already support storing an SSH public key and (for some facilities) a user’s X.509 Distinguished Name. The user’s federated identity would be associated with the user’s identity under the HPC facility (usually the username).

HPC facilities currently don’t support building this association between a user’s username and their federated identity. Therefore, we are proposing a new identity-mapping service that would allow a user to add their federated identity. This new service would be automated, **requiring no HPC admin intervention**.

A typical user experience would involve the HPC user logging into the identity-mapping service using their HPC credential; e.g., username + password, possibly involving a second factor (2FA). Once they are logged in, the user would authenticate via the federated authentication service. This provides the identity-mapping service with both the HPC username and the federated identity. These two pieces of information are then stored.

Other services (those that support OIDC authentication) would use the identity-mapping service to identify the HPC user's identity. Existing HPC services (those that do not support OIDC authentication) would not use the identity-mapping service. Therefore, this approach should introduce a minimal risk on existing services.

We imagine that this identity-mapping service would be run within the HPC facility. Over time, this functionality may be replaced by similar behaviour in the HPC facility's IAM service.

3. Context

The identity-mapping service is called ALISE. ALISE is used to collect information about different accounts a single user may have. We expect ALISE to be used at Computer Centres (sites) that operate their own user database, and wish to understand which federated identities a given user owns. For this, ALISE provides multiple logins into a single web-session, so they can prove their identities from different external and internal identity providers.

This is useful for sites that plan to enhance their existing services with federated authentication mechanisms.

4. Technical Description

The Account Linking Service (ALISE) is a service with a web and a rest interface.

The web interface allows users to log in with various external and internal authentication mechanisms. The external (federated) mechanisms will rely on OpenID-Connect, while the internal mechanisms will be adapted to individual Computer Centres (sites) requirements.

We distinguish between two different kinds of accounts:

- Internal Account: Privileged account, which was created a priori, following the procedures (e.g. ID-verification, Paper-Forms, ...) that are in place at a given site.
  The information associated with the internal accounts are (at least)
    - Unix User ID
    - Primary Unix Group ID

- External Account: Federated Account, which typically comes from an AARC-BPA Community AAI, such as EGI-Checkin, WLCG-IAM, Elixir-AAI, Geant-Core-AAI-Platform, Helmholtz-AAI, Google, ORCID, Meta, github, ...
  The information associated with the external accounts are (at least)
  - sub + iss: OpenID-Connect claims to uniquely identify any user
  - voperson_id: Community identifier to uniquely identify any user
  - Name + Email
  - Home Organisation + Affiliation in Home Org.(e.g. student@kit.edu)
  - Assurance: To understand how well the user is known at issuing IdP, most often used to differentiate between a university and a "social" login. Details here and here.

We expect one instance of ALISE per site.

The external mechanisms will typically rely on OpenID-Connect, while the internal mechanisms will be adapted to individual Computer Centres (sites) as required.

Once the identities are linked, services, that need a mapping between federated and local account

5. User-Flow

- User Logs in to ALISE using the **internal** account of the Computer Centre
- Once logged in, user clicks "add community identity" in ALISE
- User is prompted for which community to choose (e.g. EGI-Checkin, ...)
- Identity is linked. Any number of external identities can be added.

6. Service Interaction

Services that want to provide access to services based on the federated identity, and tailored to the corresponding owner of the internal identity. One example could be webDAV/OIDC.

7. Hosting

During the development phase ALISE will be hosted at https://alise.data.kit.edu during development. Hosting by sites will be possible once a release has been made.