Beyond the Wiki

Federating Applications in Support of Research

Jim Basney (NCSA)
Pål Axelsson (SUNET)
Matthew Economou (NIH)
Scott Koranda (SCG)*

*organizer in absentia
Beyond the Wiki

- Academic research is collaborative, multi-institutional, multi-national
- Federated identity is a perfect fit!
- What do you federate first? The WIKI!
- But the wiki can’t do it all…
- Federate other research applications to provide a consistent experience
  - Single sign-on, authorization, etc.
- This is the topic of our session

By Andrew Laing (Wiki-Wiki!!!)
CC BY-SA 2.0, via Wikimedia Commons
Federating Research Apps: Challenges

- Single identity provider (SAML or OIDC)
- Usernames
- Authorization/provisioning APIs (LDAP, SAML, SCIM, etc.)
- Account linking
Much Progress!

- eduGAIN growth
- REFEDS Research and Scholarship
- REFEDS Multi-Factor Authentication
- REFEDS Assurance Framework
- REFEDS OpenID Connect for Research
- REFEDS SIRTFI
- InCommon Baseline Expectations
federated identity management & collaborative organization management
- Multi-user Jupyter server
- Jupyter Notebooks support authoring/sharing of code, math, text, and multimedia
- Also supports JupyterLab
JupyterHub & CILogon

Use OAuthenticator to support OAuth with popular service providers

JupyterHub's OAuthenticator currently supports the following popular services:

- Autho
- Bitbucket
- CILogon
- GitHub
- GitLab
- Globus
- Google
- MediaWiki
- Okpy
- OpenShift

A generic implementation, which you can use for OAuth authentication with any provider, is also available.

https://github.com/jupyterhub/oauthenticator
JupyterHub & CILogon

- Migration from OAuth1 to OAuth2
- Custom client configurations ("skin")
- Setting JupyterHub username
  - ePPN
  - email
  - custom CILogon sub
- Group-based authorization
Custom CILogon sub

- Default CILogon sub claim is a URI (e.g., http://cilogon.org/serverA/users/534)
- Can be customized per client application
- Approach: use voPersonApplicationUID with COmanage
Thanks!

Interested in using CILogon?
Contact:

jbasney@ncsa.illinois.edu
help@cilogon.org

This material is based upon work supported by the National Science Foundation under grant numbers 0850557, 0943633, 1053575, 1440609, and 1547268 and by the Department of Energy under award number DE-SC0008597. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the United States Government or any agency thereof.
Federating virtual home organizations and applications in Sweden

SUNET

SWAMID
Swedish Academic Identity Federation
Virtual IdP for VOs

- Enable VOs to present an "IdP interface" to federated services
- Leverages existing federated credentials many users already have
- SPs "see" the virtual IdP
  - Users may select their VO from discovery service
- Authentication handled by "traditional" IdPs, e.g., campus IdPs and social gateways
  - IdPs "see" a proxy SP
- The Virtual IdP only accept pre-registered users
- Virtual IdP consumes authentication assertion from authenticating IdP, retrieves VO attributes about user, and asserts VO identity to SP
Virtual IdP for real organizations

- It’s is not cost effective for a really small organization to run their own identity management infrastructure.
- The organization uses Un-Affiliated IdP* with good enough identity assurance as user identity and credentials backend for their users.
- The Virtual IdP for VOs adds organizational attributes to the users, i.e. organizational mail addresses and eduPersonScopedAffiliation.
- The Swedish NREN SUNET will use this for their employees and the coworkers that are employed by other Swedish academic home organizations.

*A more including name on the IdPs of Last Resort (IoLR)*
Virtual IdP for VOs Implementation

- COmanage Registry for collecting and managing VO identities
  - Link organizational identity (authenticating IdP) to VO identity
  - Collect, manage, mint VO attributes and identifiers
  - Each VO is a COmanage CO in a multi-tenant deployment
  - Provision CoPerson objects to LDAP using voPerson schema

- SATOSA Proxy for implementing the virtual IdP
  - Set of SingleSignOnService/binding endpoints for each VO, e.g.,
  - https://idp.comanage.se/saml/{URL safe VO name}/sso/redirect
  - New SAMLVirtualCoFrontEnd class implements the functionality
  - Microservice (plugin) architecture enables lookup of VO identity using authenticated identity
    - Add VO identifiers and attributes to assertion sent to SP
    - Translate, replace, or otherwise modify attributes sent by IdP
COmanage Registry CO Email Lists

- Email Lists are data structures that associate CO Groups and Group Memberships with listservs

- Registry does not provide message delivery capabilities
  - maintains metadata about lists
  - provision to list management software using Provisioner Plugins
  - currently Mailman Provisioning Plugin, more possible in future

- Groups attached to an Email List to define list membership and roles
  - Members
  - Administrators
  - Moderators
COmanage Registry CO Email Lists

Example: if you create a email list called "Researchers" and attach the CO Group "Biologists" as its members group, then any valid CO Person who is a member of Biologists will be subscribed to the Researchers list in the mailing list management software.

The list name will typically become the left hand side of the list's email address.
COmanage Registry Mailman Provisioning Plugin

- Manages GNU Mailman3 mailing lists using COmanage Registry data
  - Provisions to Mailman using a REST API (new in Mailman 3)
- Currently requires patches against Mailman3 codebase
  - Simple extensions to the REST API (less than 50 lines of additional code)
  - As such, it is considered *Experimental*
  - Working to push patches into future Mailman3 release
Edit https://lists-dev.sphericalcowgroup.com/api

The Mailman server must be available and the specified credentials must be valid before this configuration can be saved.

If the specified domain is not already configured on the Mailman server, it will be created.

Mailman3 Admin Server Base URL *
Do not include API version, eg: http://myhost.org:8001

Mailman3 Admin Username *

Mailman3 Admin Password *

Mailman3 List Domain *
Mailing lists managed by this provisioner will be created under this domain (eg: lists.mydomain.org).

Preferred Email Type
If set, email addresses of this type will be used (when available) as the preferred Mailman delivery address.

* denotes required field

SAVE
Test CO 1

**Edit lunch**

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Status</th>
<th>Members Group</th>
<th>Administrators Group</th>
<th>Moderators Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>lunch</td>
<td>People who eat lunch</td>
<td>Active</td>
<td>Lunch Group</td>
<td>CO:admins</td>
<td>CO:admins</td>
</tr>
</tbody>
</table>

* denotes required field

**CO:admins**

**View Lunch Group**

**View CO:admins**

**View CO:admins**
COmanage Registry Atlassian integration

- User and group memberships in Atlassian Crowd
- Atlassian delivers a wide set of applications* and all can use the Crowd software for identity and group management
- Registry does not provide Atlassian application capabilities
  - maintains metadata about users and group memberships
  - provision to Crowd software using Provisioner Plugins
  - authentication in Atlassian applications are done by the Eduix crowd-shibboleth-module

*Examples on Atlassian applications are Confluence and Jira
Add a New Server

**Description**: Crowd API server

**Status**: Active

**Type**: HTTP

* denotes required field

ADD
Edit HTTP Server

Server URL*: https://crowd.sunet.se

Username: restadmin

Password:

* denotes required field

SAVE
Add a New Provisioning Target

Description: SUNET Primary Crowd Service

Plugin: CrowdProvisioner

Status: Automatic Mode

Provisioning Group: [Select Group]

Skip If Associated With Org Identity Source: [Select Source]
Edit Crowd Provisioner Target

Server *
Crowd API server

Username Identifier Type *
ePPN
* denotes required field

SAVE
Thanks!

Questions and thoughts?
Pål Axelsson, pax@sunet.se
Murchison Widefield Array
Murchison Widefield Array

- Low frequency radio telescope (80 to 300 MHz)
- Located at the Murchison Radio-astronomy Observatory (MRO)
- Planned site of the future Square Kilometre Array (SKA)
- One of three telescopes designated as a precursor for the SKA
- International collaboration
  - Australia, Canada, India, New Zealand, and the United States (China soon)
  - ~400 collaborators
Murchison Radio-Astronomy Observatory
Murchison Widefield Array

- 2048 dual-polarization dipole antennas
- 4 by 4 arrays or "tiles"
- 128 tiles
- 112 tiles scattered across 1.5 km core region
- Several hundred square degrees at a resolution of several arcminutes
MWA Identity and Access Management

Access to what?

- Collaboration tools
  - wiki
  - email lists and archives
  - code repositories/version control
    - recent decision to make all code open source
      - Back to initial plan--will deploy federated GitLab
- calendaring, chat, file sharing

- Data
  - web portal
  - command line access (IAM still TBD)
  - access to computing resources
MWA Identity and Access Management

Project started November 2016

- "Cross fertilization" from LIGO
  - Members part of both projects saw benefit to LIGO and wanted it for MWA
- Target single sign-on (SSO)
- Use federated identity as foundation
- Contract with Spherical Cow Group
  - Gap analysis
  - Architecture design
  - Roadmap
  - Deployment
- Services hosted and managed at Curtin University (Perth, Australia)
SAML IdP/SP Proxy-centered Architecture

- COmanage
  - user lifecycle management
  - attribute management
  - identifier management
  - authorization
- SATOSA from IdPy as the proxy
  - Manage spectrum of IdP behavior
  - Future: integration with cloud services
- pyFF from IdPy for IdP discovery/metadata
  - Good "out of box" discovery experience

IdP/SP proxy-centered architecture has become the de facto architecture for VOs (see eg. the AARC Blueprint Architecture).
Welcome to MWA Collaboration. Please select an action from the menus.
## Add a New Email List

**Name**

List name may only consist of alphanumeric characters, dot, dash, and underscore. Although the list name may be changed here, not all Provisioning Targets may support renaming a list.

- mwa-members

**Description**

All MWA Members

**Status**

Active

**Members Group**

Members of this group will be provisioned as recipients of messages to the list

- CO:members:all

**Administrators Group**

Members of this group will be provisioned as administrators of the list, if supported by the mailing list service

- CO:admins

**Moderators Group**

Members of this group will be provisioned as moderators of the list, if supported by the mailing list service

- CO:admins

* denotes required field

**ADD**
Status & Roadmap

- COmanage Registry, SATOSA proxy, pyFF discovery service in production
- SATOSA SP published into Australian Access Federation (AAF)
  - Registered as REFEDs Research and Scholarship (R&S) SP
  - Pushed into eduGAIN by AAF
- Application integration (in parallel)
  - Wiki (Confluence) completed
  - GNU Mailman 3 in testing
  - GitLab planned
  - Data portal (Django/Python) in design
- Leverage infrastructure for ASVO
  - All-Sky Virtual Observatory
  - Access data across federated network of datasets, from all types of astronomical facilities in Australia
  - Python based
  - Use pySAML2 to add SAML functionality
    - Made tractable because only needs to manage integration with SATOSA proxy
Thanks!

Questions and thoughts?

Scott Koranda/skoranda@sphericalcowgroup.com

Greg Sleap/greg.sleap@curtin.edu.au
The NIAID Science Forum
Federated CMS, simplified

Then:

Not shown: AD FS custom attribute store, Shib EDS hacked into an AD FS web theme, MDQ kludges for metadata consumption, etc.

Now:

SATOSA (and soon, pyFF) do all the work. AD FS is just another SAML 2.0 SP. Blink and you’ll miss it.
Federated CMS, opportunities

How can we integrate better with our meta-directory?
Federated CMS, opportunities

How can we integrate better with our meta-directory?

How can we make permissions management easier?
Federated CMS, opportunities

How can we integrate better with our meta-directory?

How can we make permissions management easier?

How can we kick off enrollment flows from CMS authentication or authorization failures?
Federated EDC/ePRO

That was easy!

<table>
<thead>
<tr>
<th>Security &amp; Authentication Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authentication Settings (System-level)</td>
</tr>
<tr>
<td>Authentication Method</td>
</tr>
<tr>
<td>Used for all global pages and as default authentication for newly created projects</td>
</tr>
<tr>
<td>Shibboleth (see custom settings below)</td>
</tr>
<tr>
<td>For details on each authentication method, see the wiki authentication page, which describes them and lists how to install, configure, and enable each in REDCap.</td>
</tr>
<tr>
<td>Need help troubleshooting your LDAP configuration?</td>
</tr>
</tbody>
</table>

| Additional Shibboleth Authentication Settings: |
| Shibboleth Username Login Field |
| REMOTE_USER |
| Name of the server variable that contains the user's username that Shibboleth defines in PHP (e.g. $_SERVER['REMOTE_USERNAME']) |

| URL for Shibboleth Logout Page |
| https://mcap.scienceforum.sc/Shibboleth.sso/Logout |
| (Full URL with qualifiers) |
Federated EDC/ePRO, interrupted

That was easy....

...except JIT provisioning isn’t supported
Federated EDC/ePRO/SMOP

We can build JIT provisioning.
We have the technology:

- Reverse-engineer the user and access control tables.
- Map COmanage groups to REDCap projects/groups/ACLs.
- Develop a session hook callable by Shibboleth.
Federated EDC/ePRO, enrollment redux

My Projects

Project Title

You do not have access to any projects
Federated EDC/ePRO, mobile edition

The mobile app does not support SSO. Instead, the user must request an API token and configure that in the app.

(API access bypasses SSO.)
Federated Cloud/SaaS

The Good.
Federated Cloud/SaaS

The Good.

The Bad.
Federated Cloud/SaaS

The Good.

The Bad.

The Ugly.
Federated HPC
Federated Discovery

Today, Shibboleth EDS:

Soon, pyFF:
Can we make it better through analytics? Machine learning? Email address mappings?

How do we train our users?
Federated Future

Works in progress:

- BioBank Pro
- GoToMeeting
- GitHub
- Huddle
- Kubernetes
- Meraki Dashboard
- Moodle
- OpenStack
- Rancher
- SharePoint 2019
- Tableau

Unmet needs:

- Bioinformatics tooling
- Business process/workflow management
- Lab data management
- Specimen tracking
Thank you!