Challenges for Federated Identity for Research Organisations

Warren Anderson, LIGO

Advantages of Federated Identity (FI) for Research Orgs

1. Provides SSO without managing credentials
   - Let home institutions handle securing credentials, password resets, phished credentials, ..
2. Quicker spin-up times, especially for small agile teams
   - Authentication infrastructure already exists
3. Research orgs do not need IAM expertise
   - Frees up FTEs to do research rather than computing, no computing hires required
4. Easier for the researcher
   - Using credentials they already have, no password proliferation
5. Easier for services
   - Plug into proxy and get authentication and authorization information handed to you

But …

“We want to use our favorite SaaS platforms!”

1. Using FI often requires subscriptions for SaaS that is otherwise free:
    a. Subscriptions cost money:
        i. GitHub - requires GitHub Enterprise Cloud subscription- US$250+/user/year
        ii. Google Drive - requires Google Workspace subscription - US$72+/user/year
        iii. Slack? - requires paid subscription - US$87+/user/year
        iv. Zoom? - requires a paid subscription - US$150+/user/year
        v. Total - Up to $550+/user/year
    b. Subscriptions require contracts:
        i. Many research orgs aren’t legal entities, can’t sign contracts
        ii. Get a university to sign for you, but:
            1. Business office taking on fiducial responsibility for people from other universities
            2. Legal office taking on legal responsibility for people from other universities
            3. Difficulties around multiple subscriptions/campus, internet domains, …

“We Need Command-Line (CL) Tools!”

1. SSH (again)
   a. Web-flow interactions are clunky for regular use
   b. Use ssh keys?
2. Cloud-based dev-ops
   a. Users prefer authenticated CL tools like terraform, ansible,etc. to deploy containers
   b. Use long-lived API keys?
3. GItHub CI/CD against RESTful APIs
   a. Automated deployment does not allow users to login with federated IDs
   b. Use long-lived API keys?
4. Automated grid job submissions
   a. Does not allow users to renew credentials after launch
   b. Use JWTs (WLCG or SciTokens)?

Strategies for Managing CL Credentials with FI

- Generate them for researchers upon request using FI enabled “proxy”:
  - Pros
    - Easy for researchers, login and get your credentials
  - Cons
    - Must design, implement and maintain “proxy”
    - Must integrate with downstream services
    - How does it work for automated processes?
- Let researchers generate them with each service via FI login to that service
  - Pros
    - Downstream services probably already generate them for authenticated users
  - Cons
    - No easy way to manage user lifecycle
      - Track creation of credentials, deactivate them during offboarding, …
      - Credential registry? How to enforce?
    - How does it work for automated processes?
    - Requires FI access to services, which often requires expensive subscriptions

Advantages of Federated Identity (FI) for Research Orgs

1. Provides SSO without managing credentials
    - Let home institutions handle securing credentials, password resets, phished credentials, ..
2. Quicker spin-up times, especially for small agile teams
    - Authentication infrastructure already exists
3. Research orgs do not need IAM expertise
    - Frees up FTEs to do research rather than computing, no computing hires required
4. Easier for the researcher
    - Using credentials they already have, no password proliferation
5. Easier for services
    - Plug into proxy and get authentication and authorization information handed to you

But …

Advantage 1: Provides SSO without managing credentials

- Many command-line activities require alternate credentials
- Linking ssh-keys, API tokens, JW tokens, etc to FI requires
  - money (paid subscriptions with SSO) and/or
  - effort (credential linking/management service provided by research org).
- Automated process require credentials that are not generated through webflows (e.g kerberos)

Advantage 2: Quicker spin-up times

- Using FI with SaaS requires subscriptions
  - Setting up a subscription without a campus or other partner may require becoming a legal entity or negotiating a special agreement with SaaS provider - neither is quick
  - Setting up a subscription via a campus usually requires negotiations with at least one campus business office, campus legal, etc
- Setting up FI based management for CL credentials requires either or both
  - Creating credential management infrastructure
    - Architect, implement, integrate
    - Management of user lifecycle
  - Setting up subscriptions
    - See above

Advantage 3: Research orgs do not need IAM expertise

- Using FI to manage CL credentials - two strategies:
  - Generate them for researchers upon request using FI enabled “proxy”
    - Architect and Implement proxy
    - Operate and Maintain proxy
    - Integrate CL generated credentials with downstream services
  - Let researchers generate them with each service via FI login to that service
    - Research org does not need IAM expertise but
      - How to link credentials to researchers?
      - How to manage lifecycle of credentials?
      - How to offboard researchers without losing research?

Advantage 4: Easier for the researcher

- Using credentials they already have, no password proliferation
  - True if:
    - There is a proxy to get all needed CL credentials using federated login
      - Large up front effort requirement from research org
    - All required SaaS tools accept SSO/FI
      - Large up front cash outlay from research org
  - Not true if:
    - Researcher needs to create all needed keys and tokens themselves
      - Registering credentials with research org is an extra burden, won't happen
    - SaaS tools do not have SSO
      - Use personal accounts for GitHub, Google Workspace, …
      - Work product is in control of individual researchers - what happens when they leave?

Advantage 5: Easier for services

- True for web services that can accept SSO/FI from SAML/OIDC proxy
- The list of services for which this is not straightforward without investments of effort and/or money include:
  - Shell access
  - Most web APIs
  - AWS
  - GitHub
  - Google Workspace
  - CI/CD integrations
  - Other automated jobs
  - …

Take-aways

- When federated identity works “out of the box” it is a powerful tool
- Modern research (at least in astrophysics) is adopting powerful SaaS tools, which usually require expensive subscriptions for FI integration
- Researchers often access SaaS via CL tools for which FI is not suited
- For small research orgs, the money and effort to use FI in these contexts might be prohibitive
- For large research orgs, it is perhaps more feasible but:
  - If you have the effort and money, why not just issue your own credentials?
  - Even for large research orgs, signing contracts on behalf of researchers from several institutions can be tricky or impossible.