Networking and Information Technology
Research and Development

The government seeks individual input; attendees/participants may provide individual advice only.

November 14, 2018 2:30-4:30PM ET
Kay Bailey Hutchison Convention Center, Room D175,
650 S Griffin St, Dallas, TX 75202

Middleware and Grid Interagency Coordination (MAGIC) Meeting Minutes[1]

Participants (*In-Person Participants)

Lisa Arafune*	CASC	Brian Lin	UW-Madison
Tom Barton*	UChicago/I2	Lixin Liu*	SFU
Jim Basney*	NCSA	Deep Medhi*	NSF
Joe Breen	Utah	Ben Meekhof*	UMich
Vincenzo Capone*	GÉANT	Laura Paglione	Spherical Cow
Richard Carlson*	DOE/SC	Matyas Selmici*	UW-Madison
Dhruva Chakravorty*	TAMU	Arjun Shankar*	ORNL
Vipin Chaudhary*	NSF	Alan Sill	TTU
Mark Day*	NERSC	Derek Simmel*	PSC
Padma Krishnaswamy	FCC	Harold Teunisse*	SURFnet
Eric Lancon*	BNL	Von Welch*	IU
Joyce Lee*	NCO		

Proceedings

This meeting was chaired by Richard Carlson (DOE/SC) and Vipin Chaudhary (NSF).

Speaker Series:

- *Update on IdM for Research: Some Predictions on the Future of IdM* - Von Welch, Director, Indiana University
- *Identity Management for Research Collaborations* - Jim Basney, Senior Research Scientist, Cybersecurity Group, National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign
- *ORCID Researcher Assurance Model* - Laura Paglione, Spherical Cow Group on behalf of ORCID
- *Update on Internet2/InCommon Trust & Identity Program* - Tom Barton, Senior Consultant for Cybersecurity and Data Privacy, University of Chicago
- *Trust & Identity Services Update* - Vincenzo Capone, Head of Research Engagement and Support, GÉANT
- *Identity Access and Management with Globus* - Rachana Ananthakrishnan, Head of Products, Globus, University of Chicago

[1] Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Networking and Information Technology Research and Development Program.

Speaker Presentations

Update on IdM for Research: Some Predictions on the Future of IdM **- Von Welch**

Observations in Identity (ID) management (IdM) (Slides 7-10)

Passwords

- Password reuse is the users' biggest enemy. Power of computing to crack passwords outpaces human passwords.
- Old heuristics are no longer working (NIST's original guidance for passwords/password schemes have been revised to reduce demands on password changing etc.).
- Entered a different world of password landscape for ID management. (Using passwords in more ways than were originally envisioned).
- Conclusion: Password reliance is breaking down- Not good at authenticating humans over a network.

Temporary rescue (Slides 11-17)

Trends:

- Rise of password safes: standalone and built into browser.
- Rise of single sign-ons in higher education and commercial sector.
- Passwords persist because baked into infrastructure, but note commonality of 2-factor authentication that augment and continue the smart phone's usefulness.
- Biometrics: useful authentication for smart phone; includes voice biometrics (e.g., IoT).

Rise of Cloud: 2 IdM implications (Slide 18)

- Now agents operate on our behalf on the cloud.
- Need to authorize B2B collaboration on our behalf.

Predictions (Slides 20-24)

- Big picture: Death of end-to-end authentication paradigm, which will break down.
- Biometrics will ultimately win despite flaws due to the convenience factor.
- Bifurcated authentication (will lead to devices that will be physically near to us).
    - Authenticating with biometrics will be part of authentication chain.
    - Will split between end-to-end paradigm; device authentication of human via biometrics, and device-to-device authentication using strong cryptography brings the strength of both paradigms together.

Ramifications (Slides 25-28)

- IdM systems will need to think about these devices (smartphone, laptop, etc) as first-class entities.
- Password safes will persist, particularly in web browsers, and become more invisible (akin to software updates to applications). Other safes will be on the edge with command-line tools.
- Re-enrollment issue is an increasing problem. Forgotten password or lost phone/problem; how to re-enroll into new password prone device. Note cost, not technology, is an economic challenge.

Identity Management for Research Collaborations - **Jim Basney**

CILogon: Working on expanding from federated identity management system to a collaborative organization management system. (Slides 2-3)

- Brought an identity management platform capability to science applications, which are being plugged into the platform. Mix of applications using interfaces in different ways are connecting to platform (e.g., JupyterHub)

CILogon: one of supported Open Authorization (OAuth) service providers in JupyterHub authenticator framework:

- Can support controlling types of identifiers and different attributes.
- JupyterHub has strict requirements regarding guests' usernames because need to spawn containers. Can map through COmanage, which recognizes JupyterHub.

voPerson: Attributes for virtual organizations. Complement of eduPerson, which is being used by CILogon and COmanage. Preparing to hand over to R&E federations international body for standards (REFEDS) group for more participation beyond CILogon project. (Slide 4). Includes:

- *voPerson*: Some new attributes defined in voPerson aim to capture voPerson affiliations. Affiliations in the research collaboration differ from affiliations of the home institution.
- *voPerson ApplicationUID*: application may need a different type of user ID than that on campus.
- *Author name*: for publication may be different from display name.
- *voPersonExternalID*: ID enables linking together identifiers, better support researchers' lifetime in collaboration and track collaboration number. (e.g., status)
- *voPersonPolicyAgreement* and *voPersonStatus*

Continued growth of Active CILogon Users.
Continuing increase of users; more than 6k active user threshold. (Slide 5)

CILogon Top 20 ID providers:List of frequently used identity providers that many campuses are spinning out (Slide 6)

X.509: Not dead yet. Demand still high for short-lived certificates. (Slide 7)

- GridFTP protocol and GSISSH (secure remote login command) are still widely used although they are not supported as part of the Globus toolkit.
- Grid community forum and Grid community toolkit were formed to support core grid software.
- Continue to support certificate issuance and improve ability to use it.
    - *Example*: introduction of REFEDS level of assurance (see Tom Barton's talk). Can update CILogon policies to match the REFEDS level of assurance for X595. Desire identity providers to adopt this standard to enable certificate issuance.

Higher Assurance for XSEDE's InCommon Identity Providers (IdP) (Slide 8)
Adopted multi-factor authentication standard for REFEDS using DUO MFA (Multi-Factor Authentication). Supports low and medium levels of assurance for REFEDS. IdPs can also query status in XSEDE database.

SciTokens: Aiming to use capabilities for distributed scientific computing for authorization. (Slide 9)
OAuth 2.0 Standards: for distributed authorization and delegation.

- Wish to use (e.g., HTCondor platform).
- Update: Software exists and is available at GitHub.

Custos: Identity and Access Management (IAM) for Science Gateways (Slide 10)

Implement new generation of access and management for science gateways – effective web applications that enable successful science applications and cyber infrastructure (e.g., Apache Airavata).

- Science gateways are conducting identity management, group and credential management in different ways.
- Brought together a new team to find best approaches across these platforms and how we can use the same software components and processes.

***ORCID Researcher Assurance Model*- Laura Paglione**

Background (Slides 2-3)

ORCID is an open nonprofit run by and for the research community that: funds and supports open- source software and provides unique, individually-owned researcher identifiers accompanying the individual throughout their career. As identifiers are insufficient for adoption and use, robust integration points for ORCID IDs are incorporated into existing workflows that are part of research infrastructure (e.g., paper submission).

ORCID by the Numbers (Slide 4)

- Over 1000 members from 44 countries; many members are organized into regional consortia. Member organizations create connections and integrate ORCID IDs into their workflows.
- Over 5.5M active researchers use ORCID identifiers. About 65% of identifiers are created during a process (e.g., publishing a paper).
- 66 publishers and over 7k journals support requiring ORCID IDs as part of the application process.

ORCID Trust: Policy and Controls (Slide 5)

- Establish policy and controls around trustworthiness, which is essential to providing core infrastructure necessary for the researchers' daily work.
- Individual control, at core of ORCID IDs, supports individuals making choices regarding connections.
- Ensure reliable service over time and accountability to community and integrity to ORCID identifiers and records.

Building Reputation over time through actions (Slides 6-9, see diagrams)

Core issue: verifying person's identity and trustworthiness of attached ORCID record

- Connections build trust over time.
    - Information attached to ORCID records are from organizations interacting with researchers.
    - Claims are created by people in the scholarly communication infrastructure.
- Diagram of Researcher surrounded by:
    - Talent Stewards - work directly with research institutions
    - Resource Stewards - provides resources such as funding, access to equipment, etc.
    - Contribution Stewards - outputs (publications, data)
    - Each contributes information to the ORCID record and facilitates information exchange across communities; researchers can control access and the information included in those data groups.
- Diagram of types of information exchanged by the Talent, Resource and Contributions Stewards and types of organizations and their core contributions.
- Individual creates an ORCID record that includes ID and other information (e.g., biography).

- Researcher can control visibility of information in their ORCID record. Can add permission screen to workflow for the researcher to provide permission; enable distributed flow of information (e.g., university can add researcher's affiliation).

Update on Internet2/InCommon Trust & Identity Program **- Tom Barton**

Trust & ID program areas (Slide 2)

- InCommon, as well as software, brought under new brand (InCommon Trusted Access Platform).
- InCommon Certificate service is widely used.
- eduroam
- Engagement: Enable community: new InCommon Ecosystem meeting in Summer 2019.

Global R&E federated access ecosystem (Slide 3)

- eduGAIN: 58 countries have national research and education federations;
  - local policy: IdPs (users login from) and Service Providers (SPs) (login to), some are published in eduGAIN and available internationally.
  - More than 5k entities published across 15 national federations; more than 10M U.S. users.
- Additional SPs usually have a bridge/proxy to enable federated access for users.

Federation's value to research collaborations (potential) (Slide 4)

- Ubiquity: More users increase value. Minimal attributes when engaged in R&S.
- Interop: correctly regardless of location; multi-protocol.
- Trustworthy: Increase with mutual reliance on credentials; adequate security; voice in governance.

Increasing the value. Much activity in InCommon and REFEDS () (Slide 5)

- International Standards:
  - SIRTIFI (Security Incident Response Trust Framework for Federated Identity).
  - REFEDS Assurance Framework (roughly in NIST 863 space, but narrow response to research communities' input to manage risk), includes single/multi-factor authentication.
- Response emerging to the need to address Cloud First strategies.
- Deep engagements with different kinds of researchers and research communities.
  - Science Gateways Community Institute Partner (participating along with Internet2).
  - Collaboration Success Partners Program: modeled on Campus Success Program.
- Baseline Expectations (Slide 6-11)
  - Need to accommodate biggest obstacle to achieving full value of federation: IdP and SP operators who aren't paying attention and ineffective Federation operator management
    - Identity Providers– "Eat our own dog food with our credentials". Good to include security and federation metadata. (see Slide 7)
    - Service providers: privacy issue- will respect data given to them.
    - Federation Operators' primary goal: trustworthiness; promote similar programs.
  - How the Baseline Expectations program helps:
    - Metadata health check alerts operator about issues.
    - Community Trust and Assurance Board runs processes with InCommon's support.
    - Community Consensus process: increasing expectations.
    - Contractual Participation Agreement: to maintain expectations.
    - Dispute resolution process: detect, correct and minimize deviations.

Attribute Release: Research and Scholarship Entity Category (Slide 12)

- Program with international standards from REFEDS. Each federation operator can tag an SP in their federation that substantially supports research and scholarship (R&S). Thus, can automatically release attributes to R&S SPs (opt-in option, but not much participation).
- Baseline expectations work better than opt-in.
- Community working group recommended adding R&S for IdPs added to baseline. Possibility.

Maturing Federation (Slide 13, link to paper)

- Baseline must happen everywhere - other federations are paying attention.
- Not just Security Assertion Markup Language (SAML) anymore.
- As ecosystem matures, will need more authority sources in the trust chain beyond National R&E Federation operator.
- New working group: Strategic planning at Federation 2.0 (January 2019).
- Federated Identity Management for Research v2: review and recommendations of global federation; what's not/working and desired changes.

Trust & Identity Services Update **- Vincenzo Capone**

GÉANT supports and represents over 40 National Research and Education Network (NRENS) across Europe. (Slide 2)

Challenge:

More complicated for individual users, regardless of field, to obtain access to expanding services from R&E world, primarily commercial world.

GÉANT heavily relies on eduGAIN ecosystem (Slides 4-6)

EduTEAMS

- Deployed new system based on eduGAIN.
- Service that joins different institutions, IdPs, SP under a common umbrella; maintain individual identities; facilitates access wider set of services under a common function (e.g., large experiments/collaborations- comprised of researchers from different countries and institutions) Create experiment-wide IDP, user profiles in a friendly and scalable way.
- Relies on AARC Blueprint architecture: harmonize tools, software, technologies; maintain interoperability.
- Relies on Proxy and Identity Hub. (Slides 7-10)
    - Proxy: allows different service providers common access to EduTEAM implementation. Each will have a common interface to the specific EduTEAM implementation.
    - Proxy Identity Hub: allows integration of different IdPs, including commercial IdPs, and base providers.
    - Membership Management Service: (Slide 9) most relevant part of eduGAIN. Common system with a common set of attributes that are centrally managed, regardless of location. Allows access very different set of services.
    - Discovery Service and Metadata Service: Integrate directly with each other and with proxy service to facilitate the exchange of metadata and coherence between IdP and SPs.

eduTEAMS Service Offerings (Slides 11-14)

- eduTEAMS Service: Shared platform people can join; maintained internally by GÉANT, which also defines possible policies and attributes.
- eduTEAMS Dedicated Service: White box, can be provided to standalone service offering. Managed by the community and operated by GÉANT . Day-to-day operation and hosting. But community maintains authentication, roles.
- eduTEAMS Bespoke Service: Tailored service that can be used to serve specific needs; community can host internally.

eduTEAMS Communities Roadmap (Slide 15)
Just released a centralized service. In Q2 2019, we will deploy a dedicated service offering.

eduroam managed IdP: Currently working on. (Slides 16 -17)

- Would meet needs of smaller organizations.
- Cloud-based eduroam infrastructure is managed and maintained by the eduroam operations team. Institution can use service from cloud, but controlled by institution.

eduVPN initiative (Slide 18-22) Based on same concept as eduroam.

- Shared access; can roam globally through VPN access; not need hardware or software on site.
- User-friendly with simple application.
- Connect WIFI and fire up application on your computer.
- Seeking additional partners to increase effectiveness.

Identity Access and Management (IAM) with Globus **- Rachana Ananthakrishnan**
Globus Auth: Foundational IAM service (Slide 1)

- Hosted, highly scalable OAuth authorization service with very rich token management functionality.
- Standards-based, and based on OAuth2 and OpenID Connect.

Identity broker for Research Applications (Slide 5)

- Use Globus Auth to broker authentication and authorization across all entities in research ecosystem.
- Provides security framework for interactions within this ecosystem.

Use cases:
Log in with Globus

- Can add Globus Auth to web applications.
- When people come to use web applications, can use one of hundreds of identities that Globus Auth supports; quick way to allow federated login to your application (e.g., Jetstream, KBase).

Authorization Code Grant (Slide 8)
Application: a client calling a service on behalf of their users.

- E.g., GPCR Data Explorer: Data management at back end; use tokens to authenticate users.

Native Apps calling services (CL applications) (Slide 6)
Used internally and also adopted.

- E.g., Parsl allows high performance panelized throughput computing; use global transfer.

Applications needing offline access for long running tasks (Slide 7)

- E.g., Parsl: when submit a job, not know when job will start running. Need to page in data.
- So need offline tokens (standard refresh tokens). Globus has support for obtaining, managing, validating and revoking tokens.

Use case: Apps invoking service as its itself (Slide 8). Client Credentials Grant

- E.g. Research Data Archive (RDA): data processing for users (service account). When data is ready, portal begins transfer to put in staging location. Many users in data distribution cases.
- Whole Tale project: makes data available for users.

Securing service's REST API (Slide 9)
Outsource all to IAM to Globus Auth, but not authorization.
Internal use and also external adoption:

- NIH Data Commons: built some supporting services. Service REST API secured using Globus Auth.
- Object resolution service: handles persistent identifiers. Used Globus Auth to secure services.

Invoking Dependent services (Slide 10)
Restricted delegation down the call chain. Often service calls another service on your behalf.

- NIH Concierge Service: Service needed to manage data collection sets.
    - Dependency chain: needs to do transfers to put data together, needs identifier, transfer service sometimes needs to call the group service for authorization.
    - Can have user provide restricted delegation down the chain; allow portal to just do those sets of activities for you and manage whole token infrastructure and authentication.
- Increased used outside of Globus; used as a platform.

High assurance support in Globus Auth (Slide 11)
In the past year, preparing to use Globus with our protected data.

- Built out functionality in product to support use.
    - Fine grained access (provided more metadata about the authentication).
    - Create "session context" specific to OAUTH application. Made available via token introspection.
    - Providing additional information to services, which still make access control decisions.
    - Goal to make it more friendly. Make information available and user decides.
    - High assurance: available in public API. Services using Globus Auth as a platform can leverage this and do additional security.

Services using high assurance features (Slides 12-16)

- Globus transfer and groups.
- First time we have added additional authentication assurance; can enforce user authentication with specific identity.
- Application instance isolation: treats OAUTH client as the entity. Login and access higher assurance endpoint; need to login again (although same browser session).

Additional resources. (Slide 18)
See slide for additional resource links.

Discussion

Individuals are using distributed systems and giving consent to different authorizations. How are we helping to manage the consent?

- Expiration: since not giving consent indefinitely; will need to be prompted to give again (e.g., expire in 90 days).
- Similar to authentication. Note balancing between usability and security.

What can nonexperts do? Is there an Incremental way for federating your own identity mechanisms?

- Existence proof security convergence/alignment. Is there a prescriptive position for the community to adopt?
- ARC (Authentication and Authorization for Research Communities) project.
    - Main outcome: many communities overcame this issue and no longer solved this problem on their own.
- While this issue seems complicated to end user, there has been improvement. As the community is collaborative, we can do this together. Balance between needs of community and effective use.
- Note the InCommon baseline and ARC blueprint architecture, which recognizes this pattern of proxy that takes care of the complexity and interoperability. Can get quite far with the judicious placement of proxy.
- As infrastructure providers, our mission is to abstract for users, but also build on top of the work performed by InCommon and eduGAIN.
- In ORCID, the complexity comes from the ecosystem acting in similar ways which benefit the overall community. Trying to focus each message on each research community/ part of ecosystem to be as simple as possible (collect authentic ORCID IDs as simply as possible). Took 6 years to get to this point. Observing commercial services' approach to simplifying their commercial offerings to make them more understandable to the communities they're trying to serve.
- Eduroam is Holy Grail; successful due to its simple use case. Distilling use cases within a cookbook, of use cases. Slowly maturing to this point (including maturing infrastructure). That scientific communities are now discussing how to do this is a big gain from 20 years ago.

Note the Creation of new community forum where we have open grid forum. We have been working to extend/create tools for hosting software – open to discussing with other proponents of GCF?

- Need to support code. DFTO is a successful Apache incubator project; many communities need to support code. A natural extension for GCF would be to develop a way for code projects to avoid duplicating this process.

Next meeting

TBD