Federated identity system for scientific collaborations

Summary of user requirements session

Bob Jones

10 June 2011

User Communities Present

- European Photon/Neutron facilities

- Social Science and Humanities

- WLCG High Energy Physics

- Earth Sciences

- Life Sciences

European Photon/Neutron facilities

- Users
  - Approx 10,000 users and need to keep a closer track of their involvement
  - Typically short-stays at facilities
  - 'Part-time' users – not gurus
  - Most are nomadic (use many facilities) and some are "homeless" (no home institution that could issue identity)
- Key concerns
  - Confidentiality of data
  - Remote access to data
  - Potential remote control of experiments
  - PI to take role in authz
- Status: Umbrella project in prototype stage
  - Adopted Shibboleth/SAML
  - Considering a range of organisational schemes (central authority, local offices etc.)

Social Sciences and Humanities

- Users
  - Several hundred
  - Some "homeless" users

- Key concerns
  - Single domain with identity managed by home institute
  - Many diverse and distributed data sets (multi-media, multi-format etc.) with complex relationships
  - SSO with explicit "acceptable usage" signature

- Status
  - eduGAIN (Shibboleth/SAML) prototype in use with CLARIN
  - Signatory process underway with the national IDFs and CLARIN service providers
  - Have created a home for the "homeless"

WLCG

- Users
  - 5900 X509 certs

- Key concerns
  - Security tracking across multiple sites and accounts
  - Users have multiple accounts (not all within IGTF realm)
  - Scalability
  - Easy use (would like to hide certs from users)
  - Need conversion of security credentials

- Status
  - Production usage
  - Relies on IGTF
  - Tests underway for federated approach with BNL

Earth Sciences

- ## Users
  - 5000+ for CMIP5 alone, much greater for other earth science deployments.
  - Currently 25 centres around Europe but likely to grow quickly

- ## Key concerns
  - Integration of many diverse data sets
  - Many technologies and tools already in use

- ## Status
  - Earth System Grid Federation in place using OpenID, X.509 and SAML
  - A standard solution for securing OPeNDAP and other HTTP-based services

Life Sciences

- ## Users
  - Millions access data via the EBI website

- ## Key concerns
  - 20+ datasets that are continuously growing
  - At least 1 data set (European Genotype Archive) currently requires authentication

- ## Status
  - BiomedBridges ESFRI cluster project has security on work plan and will work to bring together up to 10 life science ESFRI projects

Comparing

user community	other projects	# users	chosen technology	status	IGTF
photon/neutron	EUROFEL, PanData, CRISP	10,000	Shibboleth/SAML	Umbrella prototype	no
Social Sciences and Humanities	DARIAH, CLARIN, CESSDAH, (DASISH)	hundreds now, potential for 10000+ across SSH	Shibboleth/SAML	CLARIN SP federation - will see if they can use eduGAIN	yes
WLCG	WLCG	5900 globally	X509	production	yes
earth sciences	Earth System Grid Federation, GENESI-DEC, CMIP5, Metafor, IS-ENES	5000+ for CIMP5	OpenID, X.509 and SAML	production - earth system grid	not yet but foresee for EGI integration
life sciences	ELIXIR & potentially 10 BMI ESFRI projects	millions access data via EBI website	no chosen yet	security included in BioMedBridges project workplan	no

Conclusions

- There are many common needs and hence scope for agreement
  - SSO
  - Easy of use for part-time users
  - Focus is on data access
  - Support homeless users
  - Many tools and technologies deployed
  - Smooth transition from existing systems