Trusting External Identity Providers for Global Research Collaborations

MIND THE GAP

Jim Basney

jbasney@ncsa.illinois.edu

IGTF at CERN (Sep 19 2016)

slideshare.net/jbasney

National Center for Supercomputing Applications
University of Illinois at Urbana–Champaign

About Jim

- Doing IAM at NCSA since 2001
- Operating IGTF CAs since 2007
- Operating CILogon since 2010
- Operating NCSA SAML IdP since 2016
- Contributing to IAM in LIGO, LSST, XSEDE
- Working to improve the security of e-infrastructure
  - Cybersecurity Center of Excellence (trustedci.org)
  - Software Assurance Marketplace (continuousassurance.org)

Topics

- Who do we **trust** to provide identity and access management services for our research collaborations?
- When do we decide to implement it ourselves versus **relying on others**?
- How do we create **incentives** for establishing trust?
- How do we **bridge the gaps** in trust, functionality, and reliability?

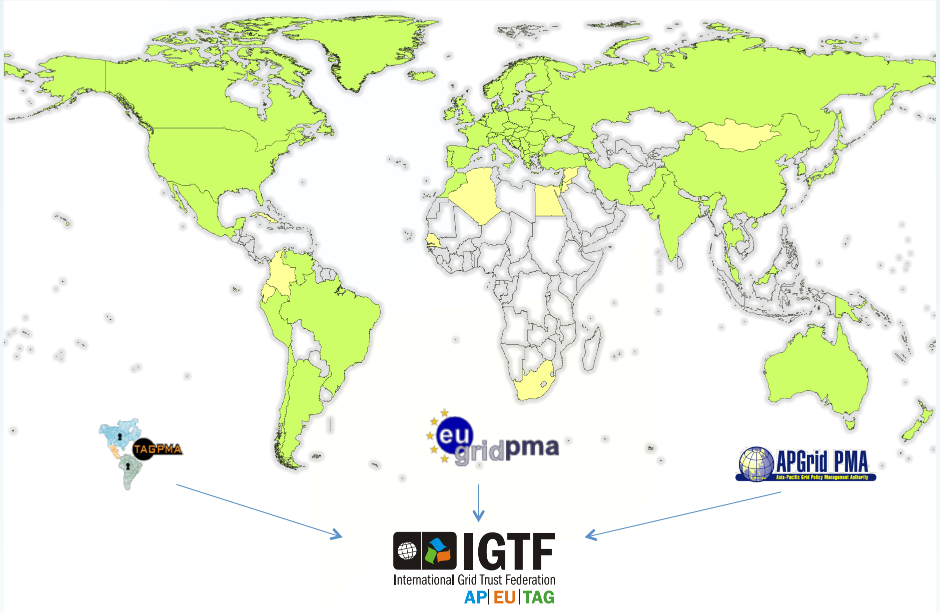
Topics

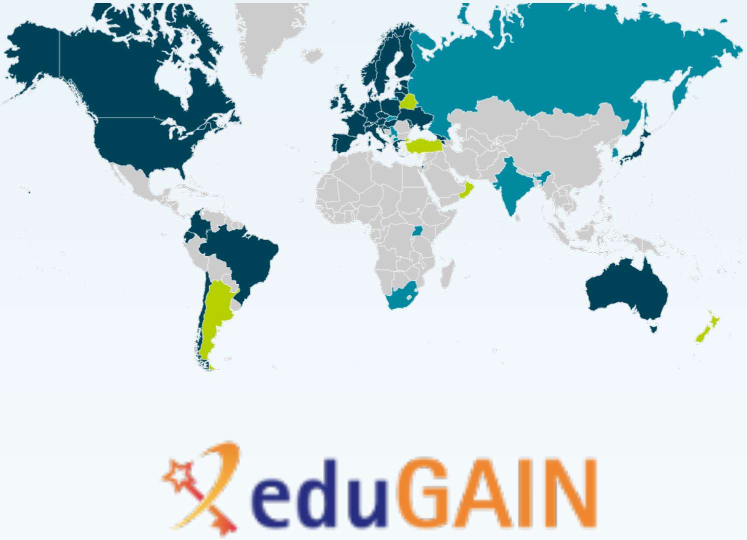
- What new challenges appear when moving from 1 to 10 to 100 to **1000+ identity providers**?
- Why does **identity information flow** more easily in some federations and not others?
- How do we determine what **identity assurance** we need and find providers who can meet those needs?
- How do we **mitigate the risks** of using external identities?
- How do we **effectively federate** services operated by the research community, higher education institutions, NRENs, and commercial providers?

Gaps

- Incentives for trust/interoperability
- Needs/Priorities/Approaches:
  - e-Research
  - Higher Education
  - Commercial
- Protocol standards and implementations
  - Scaling to 2000+ IdPs
- Operational reliability

Federations

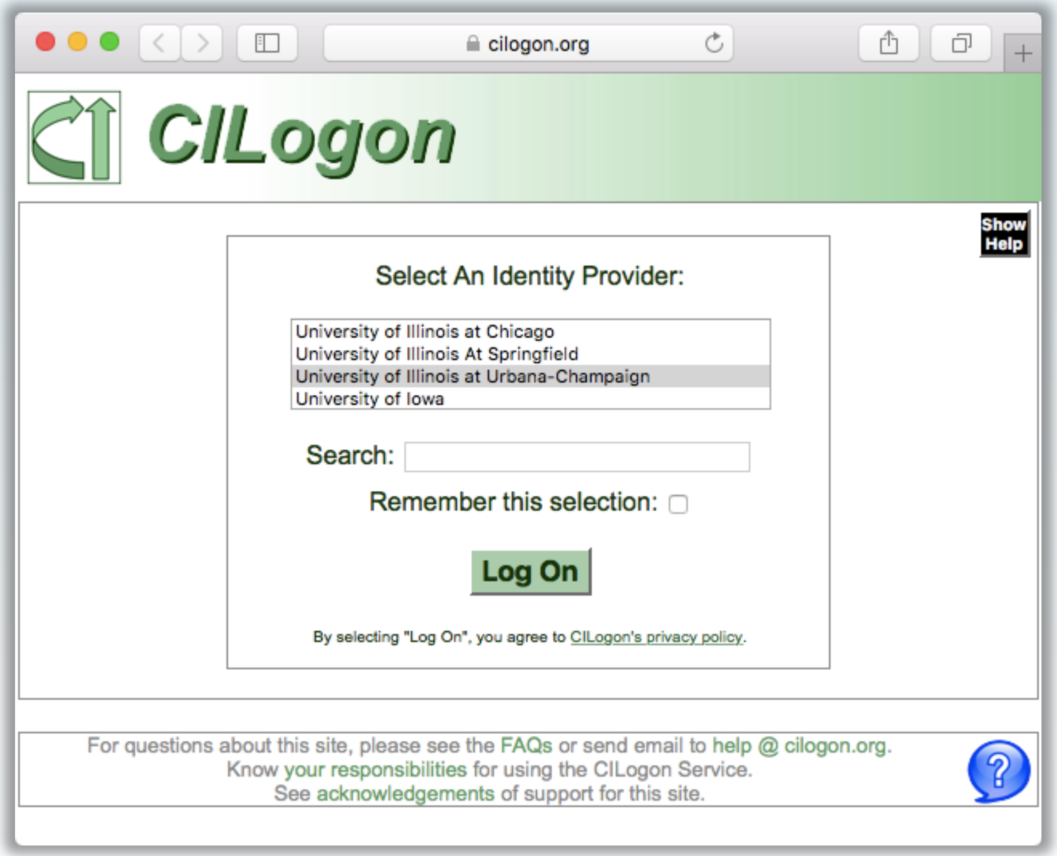

TAGPMA
eu gridpma
APGrid PMA
Asia-Pacific Grid Policy Management Authority
IGTF
International Grid Trust Federation
AP|EU|TAG



2003: IGTF established
2004: InCommon established
2005: SAML 2.0 adopted
2011: eduGAIN operational
2016: InCommon joins eduGAIN

CILogon https://cilogon.org/

- Enables use of federated identities for access to e-infrastructure
- Translates across federations and protocols



Select An Identity Provider:

University of Illinois at Chicago
University of Illinois At Springfield
University of Illinois at Urbana-Champaign
University of Iowa

Search:

Remember this selection: □

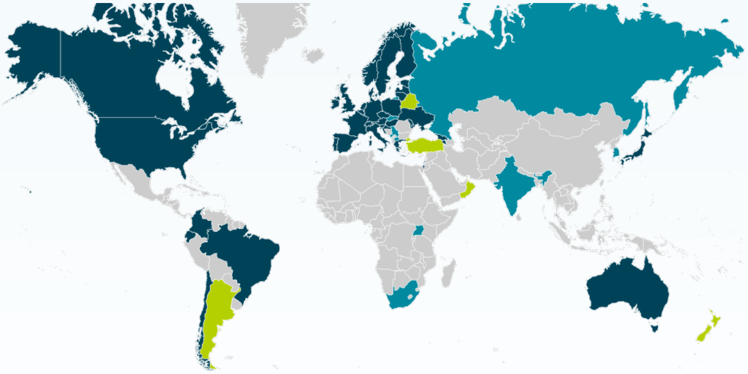
By selecting "Log On", you agree to CILogon's privacy policy.

For questions about this site, please see the FAQs or send email to help @ cilogon.org.
Know your responsibilities for using the CILogon Service.
See acknowledgements of support for this site.

What new challenges appear when moving from 1 to 10 to 100 to **1000 identity providers**?

Scaling to 100+ CAs / 2000+ IdPs

- Risk management
- Interoperability
- Operational challenges



Expanding the identity trust boundary

Local	IGTF	eduGAIN	Social IDs
1 IdP	98 CAs	2187 IdPs	1-20 IdPs
Trusted implicitly	Peer review by CA operators and research project representatives	Registration by NREN federation operators	Contractual relationships
Local policy/procedures (undocumented?)	3 regional Policy Management Authorities	38 federations adopted eduGAIN Policy Framework	Internal procedures (not public?)
Direct relationship with subscribers	Registration Authorities (e.g., VOs, e-infrastructures) vet subscribers	Direct relationship with subscribers (home IdPs)	Consumer relationship with subscribers
Users/Services	Users/Services	Users	Users/Devices

Technology

IGTF	eduGAIN	Social IDs
X.509	SAML	OpenID Connect (OIDC)
Certificates	Assertions	ID Tokens
Distinguished Names	SAML2 NameID (Transient, Persistent), ePPN, ePTID	OIDC (public, pairwise) sub claim
VOMS attributes (group, role, …)	eduPerson attributes (displayName, mail, …)	Claims (name, email, …)

Characteristics of User Identifiers

- https://wiki.shibboleth.net/confluence/x/eYBC

- **Persistent**: doesn’t change across multiple sessions (in contrast to **transient** identifiers)
- **Revocable**: user-identifier link can be severed
- **Reassignable**: a revoked identifier can be linked to a different user (typically after some hiatus period)
- **Opaque**: identifier does not reveal the user’s identity (in contrast to **transparent** identifiers)

Characteristics of User Identifiers

- https://wiki.shibboleth.net/confluence/x/eYBC

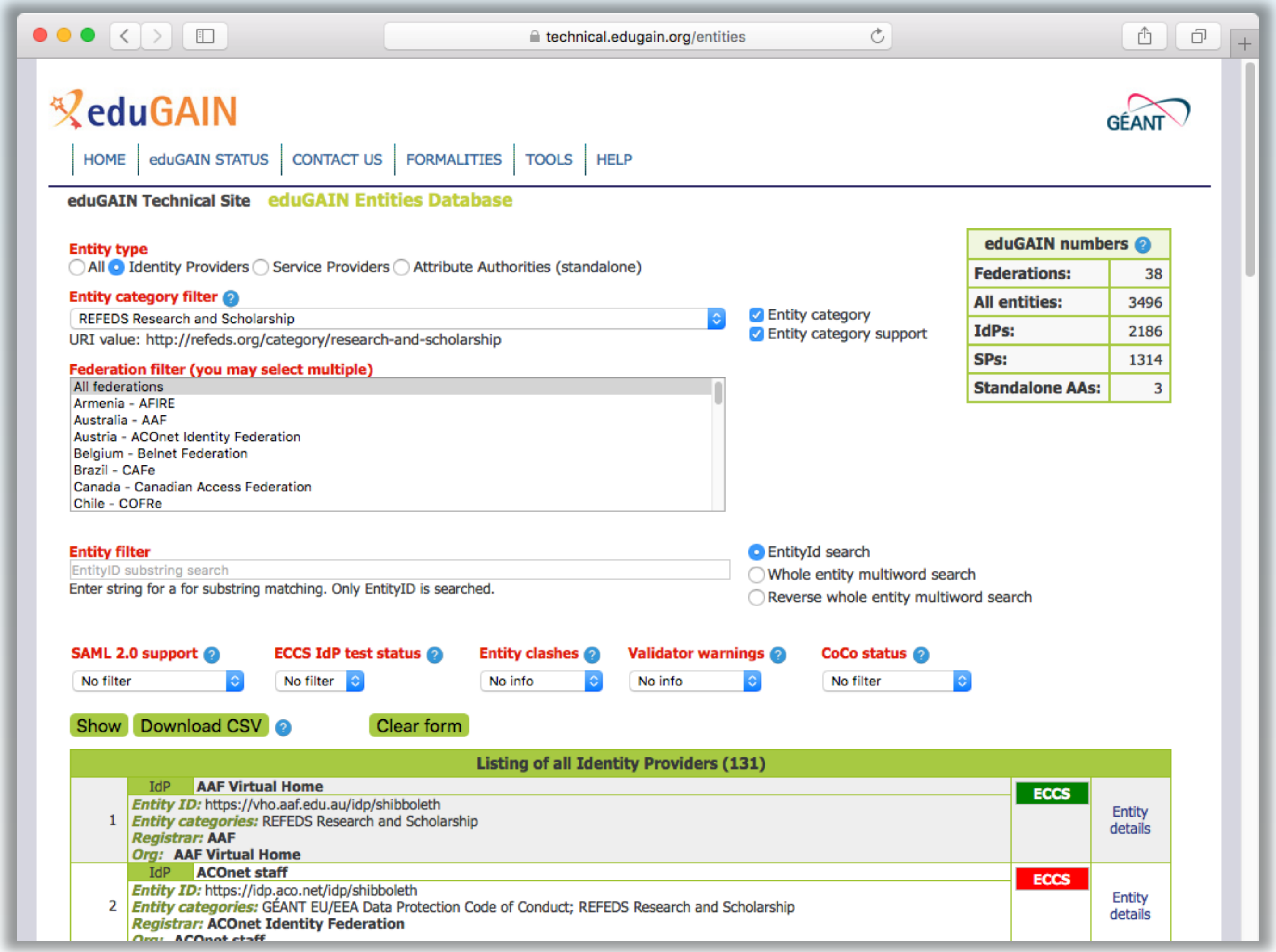
- **Targeted**: specific to a client or relying party, so the user's activities can not be correlated across applications (in contrast to **shared** identifiers)
- **Portable**: usable across security domains
- **Global**: globally unique (e.g., qualified using a DNS domain owned by the issuer)

Attributes and Identifiers

- X.509, SAML, and OIDC all support both user attributes and identifiers
- In IGTF, CAs provide **persistent** identifiers (DNs) and VOs provide attributes (e.g., VOMS)
- In eduGAIN, universities provide attributes (eduPersonAffiliation) with **transient** identifiers
  - Persistent identifiers are the exception!
  - eduPersonPrincipalName may be re-assigned
- For social identity providers, an email address is often used as the identifier
  - May be re-assignable (e.g., Yahoo!)

Why does identity information flow more easily in some federations and not others?

IGTF / X.509	eduGAIN / SAML	Social IDs / OIDC
Grid Acceptable Use, Accounting, and Incident Response policies provide a framework for exchange of user identity info	Attribute release policies vary widely across federations: 130 of 2187 IdPs release the R&S attribute bundle	Information flow is driven by user consent, a required step in the OIDC protocol
CA issues certificate to user who uses it with e-infrastructure	IdP issues assertions through browser redirects for every authentication	IdP is involved in every authentication, using data for commercial purposes
CA is operated by research organization	IdP is operated by academic institution	IdP is operated by commercial entity



technical.edugain.org/entities
eduGAIN
GÉANT
HOME | eduGAIN STATUS | CONTACT US | FORMALITIES | TOOLS | HELP

eduGAIN Technical Site **eduGAIN Entities Database**

Entity type
○ All ● Identity Providers ○ Service Providers ○ Attribute Authorities (standalone)

Entity category filter ?
REFEDS Research and Scholarship
URI value: http://refeds.org/category/research-and-scholarship

☑ Entity category
☑ Entity category support

Federation filter (you may select multiple)
All federations
Armenia - AFIRE
Australia - AAF
Austria - ACOnet Identity Federation
Belgium - Belnet Federation
Brazil - CAFe
Canada - Canadian Access Federation
Chile - COFRe

eduGAIN numbers ?	
Federations:	38
All entities:	3496
IdPs:	2186
SPs:	1314
Standalone AAs:	3

Entity filter
EntityID substring search
Enter string for a for substring matching. Only EntityID is searched.

● EntityId search
○ Whole entity multiword search
○ Reverse whole entity multiword search

SAML 2.0 support ?	ECCS IdP test status ?	Entity clashes ?	Validator warnings ?	CoCo status ?
No filter	No filter	No info	No info	No filter

Show | Download CSV ? | Clear form

Listing of all Identity Providers (131)

1	IdP **AAF Virtual Home** *Entity ID:* https://vho.aaf.edu.au/idp/shibboleth *Entity categories:* REFEDS Research and Scholarship *Registrar:* **AAF** *Org:* **AAF Virtual Home**	🟩 ECCS	Entity details
2	IdP **ACOnet staff** *Entity ID:* https://idp.aco.net/idp/shibboleth *Entity categories:* GÉANT EU/EEA Data Protection Code of Conduct; REFEDS Research and Scholarship *Registrar:* **ACOnet Identity Federation** *Org:* **ACOnet staff**	🟥 ECCS	Entity details

https://technical.edugain.org/entities

When do we decide to implement IAM ourselves versus **relying on others**?

Motivations and Driving Use Cases

- IGTF
  - Enable access to e-infrastructure
  - Per-user accounting, access control, incident response
- eduGAIN
  - Access to online academic journals by library patrons
  - Strong privacy protections
  - Access to contracted cloud services
- Social IdPs
  - Facilitate social connections
  - Advertising and user activity tracking for commercial purposes

Unique IAM needs for e-Science

- Commercial CAs
    - Why are IGTF requirements different from CA/Browser Forum?
        - Acceptance of 3 year server certificates?
        - Relax namespace requirements for server certificates?
        - Rely on standard DV/OV/EV verification for server certs?
    - Comodo, DigiCert, QuoVadis CAs now in IGTF distribution
    - Pros/Cons of keeping separate
        - Not impacted by attacks on wider commercial CAs
        - Can’t use commodity CA services – letsencrypt.org
- Commercial IdPs
    - Lack of support for scalable, multi-lateral federation
    - Privacy concerns around commercial cloud/social providers

Unique IAM needs for e-Science

- Virtual Organizations
  - Many small VOs well served by Google Apps?
  - Large VOs are like other large multi-national organizations?
  - Scaling to 100+ CAs / 2000+ IdPs
- Non-browser applications (SAML ECP)
  - Increasing use of web apps, clouds, and mobile apps in e-science
- Delegation, long-running workflows (proxy certificates)
- Interoperability across e-infrastructures
  - Related to cloud interoperability?

How do we **mitigate the risks** of using external identities?

Scoped Identifiers - Namespace constraints

- Enforcing a unique namespace for each issuer (CA, IdP) reduces the impact of compromise of any one issuer
- Namespace constraints not a standard feature of X.509
  - S/MIME verification based on email address in subjectAltName
  - HTTPS verification based on domain name in subjectAltName
  - Not required by CA/Browser Forum
  - Constraints on DNs not relevant for S/MIME & HTTPS use cases
  - Instead: HTTP Public Key Pinning (HPKP, RFC 7469)
- Namespace constraints not a standard feature of SAML
  - Not needed for primary use cases: bi-lateral federation, attribute-based authorization, transient identifiers
  - Shibboleth metadata extension for multi-lateral federation

Office365 Auth Bypass Vulnerability

- Jointly discovered in Dec 2015 by Klemen Bratec from Šola prihodnosti Maribor and Ioannis Kakavas from Greek Research and Technology Network
- Microsoft Office 365 SAML Service Provider implementation failed to check scope of IDPEmail attribute used for authorization
- Allowed Office365 tenant to impersonate users from another Office365 tenant
- Microsoft fixed the vulnerability within **7 hours** of report

http://www.economyofmechanism.com/office365-authbypass

Namespace constraints

```
TO Issuer "/DC=org/DC=cilogon/C=US/O=CILogon/CN=CILogon OSG CA 1" \
  PERMIT Subject "/DC=org/DC=opensciencegrid/.*"
```

```
access_id_CA   X509    '/DC=org/DC=cilogon/C=US/O=CILogon/CN=CILogon OSG CA 1'
pos_rights     globus  CA:sign
cond_subjects  globus  '"/DC=org/DC=cilogon/C=US/O=CILogon/CN=CILogon OSG CA 1"
                        "/DC=org/DC=opensciencegrid/*"'
```

```
<EntitiesDescriptor xmlns="urn:oasis:names:tc:SAML:2.0:metadata"
                    xmlns:shibmd="urn:mace:shibboleth:metadata:1.0" ...>
  <EntityDescriptor entityID="https://idp.ncsa.illinois.edu/idp/shibboleth">
    <IDPSSODescriptor ...>
      <Extensions>
        <shibmd:Scope regexp="false">ncsa.illinois.edu</shibmd:Scope> ...
```

Namespace constraints

- OpenID Connect (OIDC) relies on unique issuer identifier for globally unique identifier

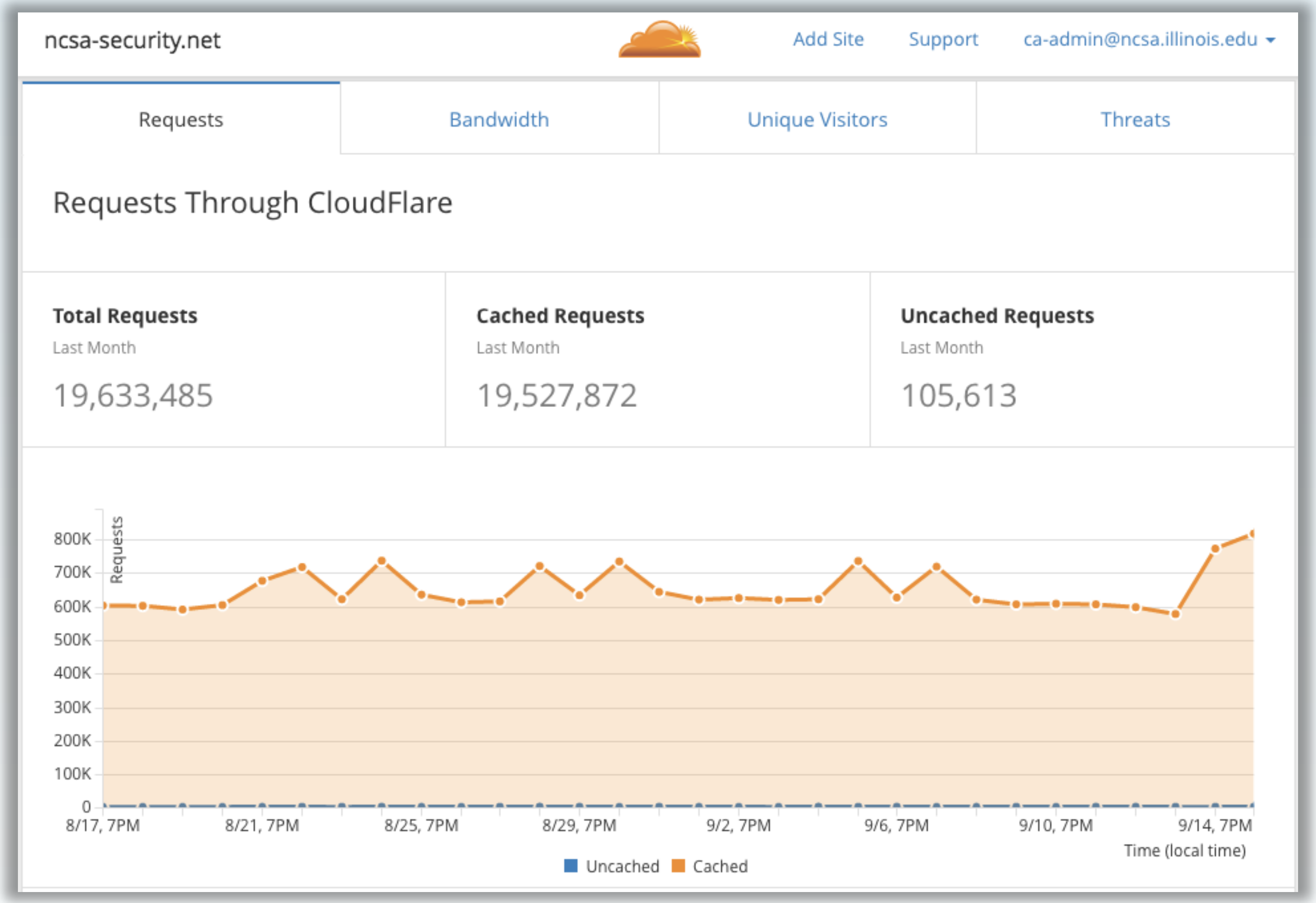
> The sub (subject) and iss (issuer) Claims, used together, are the only Claims that an RP can rely upon as a stable identifier for the End-User, since the sub Claim MUST be locally unique and never reassigned within the Issuer for a particular End-User, as described in **Section 2**. Therefore, the only guaranteed unique identifier for a given End-User is the combination of the iss Claim and the sub Claim.

How do we **effectively federate** services operated by the research community, higher education institutions, NRENs, and commercial providers?

Operational Gaps – Network Challenges

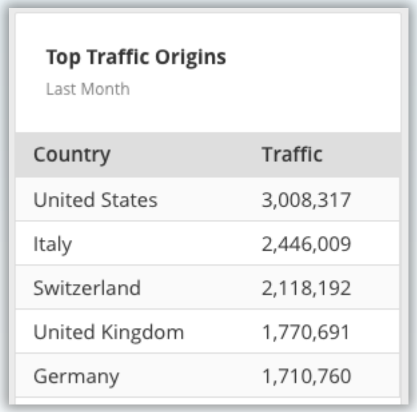
- IPv6
  - 29 of 69 IGTF CRL distribution points have AAAA records
  - 16 of 38 eduGAIN federation metadata distribution points have AAAA records
- X.509 CRLs, SAML metadata, and CDNs
  - CILogon and Comodo using CloudFlare CDN for CRLs
  - samlbits.org CDN for SAML metadata
  - Per-entity metadata https://spaces.internet2.edu/display/perentity

We love CloudFlare!



Requests Through CloudFlare

Total Requests	Cached Requests	Uncached Requests
Last Month	Last Month	Last Month
19,633,485	19,527,872	105,613



Top Traffic Origins

Last Month

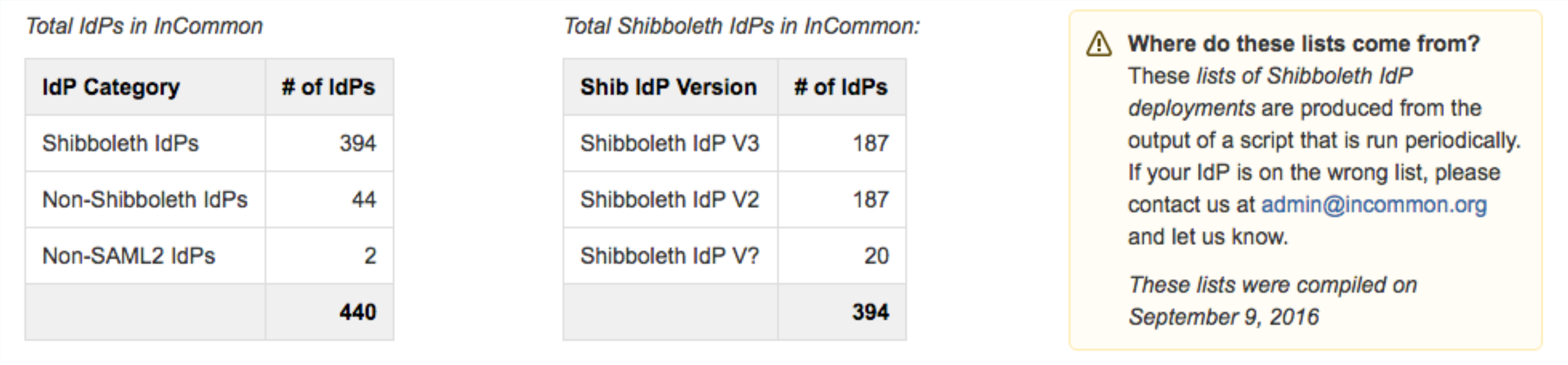
Country	Traffic
United States	3,008,317
Italy	2,446,009
Switzerland	2,118,192
United Kingdom	1,770,691
Germany	1,710,760

Operational Gaps – Algorithm Agility

- SSLv3 → TLSv1 → TLSv1.2
  - CILogon currently sees 85% TLSv1.2
  - Browsers good at updating; CLIs & web apps not so much…
- MD5 → SHA-1 → ~~SHA-512~~ → SHA-256
  - 2013-2015 IGTF transition https://www.eugridpma.org/documentation/hashrat/sha2-timeline
  - 2013-2014 InCommon transition https://spaces.internet2.edu/x/AYbYAg
- RSA key bits: 512 → 1024 → 2048 → 4096

Operational Gaps – Software Maintenance

- Shibboleth IdPv3 EOL Jul 31, 2016



Total IdPs in InCommon

IdP Category	# of IdPs
Shibboleth IdPs	394
Non-Shibboleth IdPs	44
Non-SAML2 IdPs	2
	440

Total Shibboleth IdPs in InCommon:

Shib IdP Version	# of IdPs
Shibboleth IdP V3	187
Shibboleth IdP V2	187
Shibboleth IdP V?	20
	394

⚠ **Where do these lists come from?**
These *lists of Shibboleth IdP deployments* are produced from the output of a script that is run periodically. If your IdP is on the wrong list, please contact us at admin@incommon.org and let us know.

These lists were compiled on September 9, 2016

https://spaces.internet2.edu/x/0IIQBg

Operational Gaps – Software Maintenance

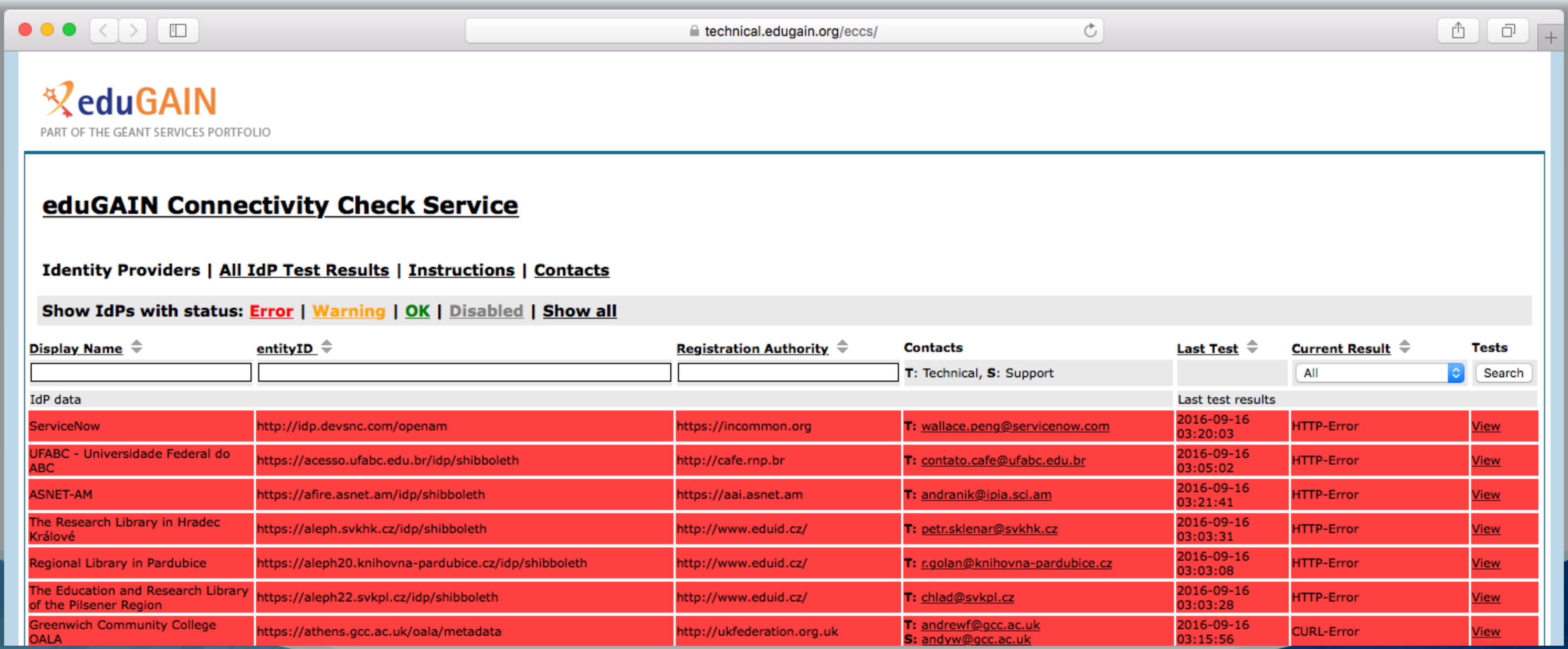
- OpenSSL 1.0 changes CA filename hashing algorithm

> Date: Mon, 15 Feb 2010
> It has come to the attention of the IGTF that the developers of the OpenSSL software (www.openssl.org) are about to release a new version of their software (version 1.0) which is fundamentally incompatible with both any pre-existing versions of their own software, as well as bring incompatibility with many other software products that use a directory-based trust anchor store (such as Apache's mod_ssl, the gLite Trust Manager, gridSite or VOMS)...

- OpenSSL 1.1 (Aug 2016) changes APIs…

Operational Gaps – Reliability

- eduGAIN CCS - https://technical.edugain.org/eccs/
    - 2047 IdPs: 1300 OK / 396 error / 179 warning / 172 disabled



eduGAIN
PART OF THE GÉANT SERVICES PORTFOLIO

eduGAIN Connectivity Check Service

Identity Providers | All IdP Test Results | Instructions | Contacts

Show IdPs with status: Error | Warning | OK | Disabled | Show all

Display Name	entityID	Registration Authority	Contacts	Last Test	Current Result	Tests
			T: Technical, **S**: Support		All	Search
IdP data				Last test results		
ServiceNow	http://idp.devsnc.com/openam	https://incommon.org	**T:** wallace.peng@servicenow.com	2016-09-16 03:20:03	HTTP-Error	View
UFABC - Universidade Federal do ABC	https://acesso.ufabc.edu.br/idp/shibboleth	http://cafe.rnp.br	**T:** contato.cafe@ufabc.edu.br	2016-09-16 03:05:02	HTTP-Error	View
ASNET-AM	https://afire.asnet.am/idp/shibboleth	https://aai.asnet.am	**T:** andranik@ipia.sci.am	2016-09-16 03:21:41	HTTP-Error	View
The Research Library in Hradec Králové	https://aleph.svkhk.cz/idp/shibboleth	http://www.eduid.cz/	**T:** petr.sklenar@svkhk.cz	2016-09-16 03:03:31	HTTP-Error	View
Regional Library in Pardubice	https://aleph20.knihovna-pardubice.cz/idp/shibboleth	http://www.eduid.cz/	**T:** r.golan@knihovna-pardubice.cz	2016-09-16 03:03:08	HTTP-Error	View
The Education and Research Library of the Pilsener Region	https://aleph22.svkpl.cz/idp/shibboleth	http://www.eduid.cz/	**T:** chlad@svkpl.cz	2016-09-16 03:03:28	HTTP-Error	View
Greenwich Community College OALA	https://athens.gcc.ac.uk/oala/metadata	http://ukfederation.org.uk	**T:** andrewf@gcc.ac.uk **S:** andyw@gcc.ac.uk	2016-09-16 03:15:56	CURL-Error	View

How do we determine what **identity assurance** we need and find providers who can meet those needs?

Identity Assurance

- InCommon Assurance program www.incommon.org/assurance/
  - Silver (LOA 2) and Bronze (LOA 1)
  - IGTF CILogon Silver CA
  - Lacking incentives: apps didn't show up
  - Implementation hurdles: audits and contractual agreements
  - Virginia Tech certified at LOA 2 but didn't renew
  - Too onerous to map from existing certifications (e.g., DOE Labs)
- FICAM - idmanagement.gov
  - Google certified at LOA 1 but didn't renew

Identity Assurance

- IGTF
  - Peer review
  - Active engagement from relying parties
  - Move to technology-agnostic assurance profiles https://www.igtf.net/ap/authn-assurance/
- REFEDS Assurance WG - https://wiki.refeds.org/x/MgDI
  - Apply IGTF-style assurance to SAML IdPs?
  - CILogon-FNAL-LIGO-OSG engaged
  - Still a question of incentives:
    - Enable university IdPs to simply self-assert what they do
    - Enable e-research IdPs to assert a higher assurance

Identity Assurance – Multi-Factor Auth

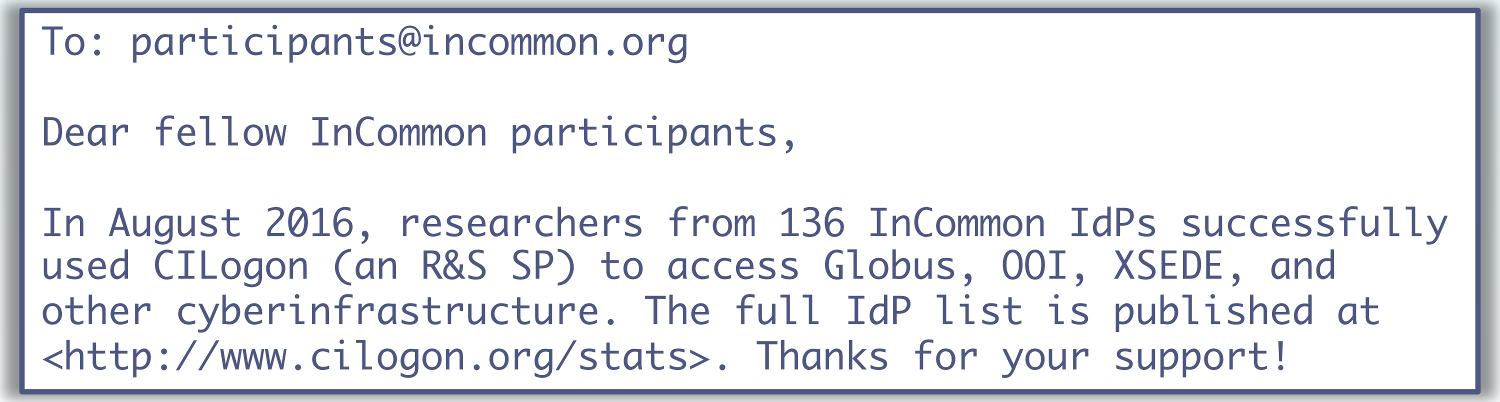
- e-Science use cases requiring multi-factor authentication
  - Blue Waters supercomputer at NCSA
  - TACC resources (XSEDE)
  - and others…
- NCSA Two Factor CA (IGTF)
- InCommon MFA Interoperability Profile Working Group
  - https://spaces.internet2.edu/x/CY5HBQ
  - InCommon MFA Profile – spec almost fits on one page!
  - http://id.incommon.org/assurance/mfa

How do we **bridge the gaps** in trust, functionality, and reliability?

Bridging the gaps: eduGAIN reality

- Some % of IdPs will always be broken
  - Automate error handling
  - Provide work-arounds
    - identity linking, account recovery, catch-all IdPs
- Increase awareness
  - Make it easy for participants to declare their practices / requirements (attributes, assurance, error handling, etc.)
  - Monitoring / alerts
  - Sharing requirements (e.g., FIM4R)

Bridging the gaps: showing appreciation



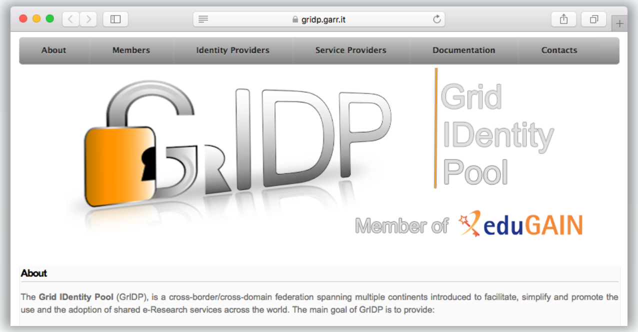
To: participants@incommon.org

Dear fellow InCommon participants,

In August 2016, researchers from 136 InCommon IdPs successfully
used CILogon (an R&S SP) to access Globus, OOI, XSEDE, and
other cyberinfrastructure. The full IdP list is published at
<http://www.cilogon.org/stats>. Thanks for your support!

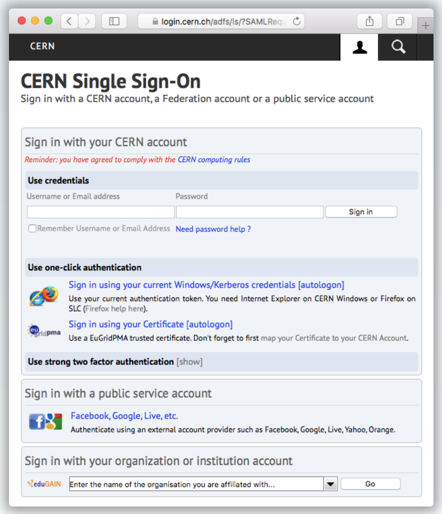
Bridging the gaps: role for e-infrastructure

- Continued need for e-infrastructure IdPs / federations
- Potential new roles for IGTF



About

The **Grid IDentity Pool** (GrIDP), is a cross-border/cross-domain federation spanning multiple continents introduced to facilitate, simplify and promote the use and the adoption of shared e-Research services across the world. The main goal of GrIDP is to provide:



CERN Single Sign-On

Sign in with a CERN account, a Federation account or a public service account

Sign in with your CERN account

Reminder: you have agreed to comply with the CERN computing rules

Use credentials

Username or Email address

Password

□ Remember Username or Email Address Need password help ?

Use one-click authentication

Sign in using your current Windows/Kerberos credentials [autologon]

Use your current authentication token. You need Internet Explorer on CERN Windows or Firefox on SLC (Firefox help here).

Sign in using your Certificate [autologon]

Use a EuGridPMA trusted certificate. Don't forget to first map your Certificate to your CERN Account.

Use strong two factor authentication [show]

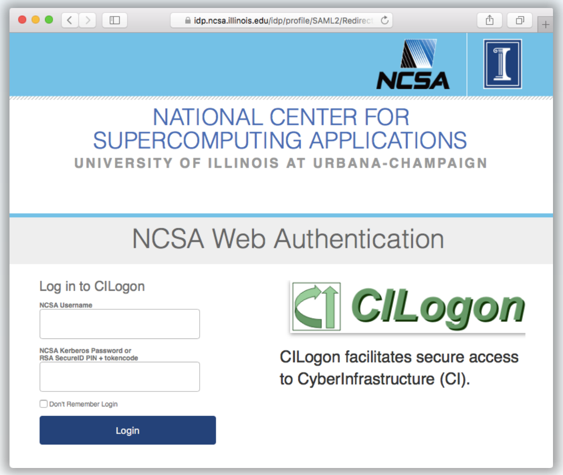
Sign in with a public service account

Facebook, Google, Live, etc.

Authenticate using an external account provider such as Facebook, Google, Live, Yahoo, Orange.

Sign in with your organization or institution account

eduGAIN Enter the name of the organisation you are affiliated with...



NATIONAL CENTER FOR SUPERCOMPUTING APPLICATIONS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

NCSA Web Authentication

Log in to CILogon

NCSA Username

NCSA Kerberos Password or RSA SecureID PIN + tokencode

□ Don't Remember Login

CILogon facilitates secure access to CyberInfrastructure (CI).

Thanks!

jbasney@ncsa.illinois.edu

keybase.io/jbasney

F867 03EA 84ED 456A 7D3E
388E 04DA 0074 775D 6316

@JimBasney

slideshare.net/jbasney