the globus alliance
www.globus.org

Authentication for Virtual Organizations: From Passwords to X509, Identity Federation and GridShib

BRIITE Meeting
Salk Institute, La Jolla CA.
November 3th, 2005

Von Welch
vwelch@ncsa.uiuc.edu

NCSA

Outline

- What are Virtual Organizations (VOs)?
- Authentication in VOs
  - Global names
    - X509 and PKIs
  - Identity Federation
    - Shibboleth
    - On-line CAs
- GridShib
- Challenges Ahead

What is a Virtual Organization?

- A dynamic set of users and resources, from different institutions, who operated in a coordinated, controlled manner to achieve a common goal.
- Key attributes:
  - Dynamic
  - All users and resources still belong to original institution
  - Coordinated and controlled
    - Shared policy

What's an institution?

- AKA a "real organization"

- Some relevant attributes:
- Have users
- Professional IT staff and services
  - Define namespace
  - Authentication mechanism
  - Reluctant to change
- Legal standing
- Persistent
- Cares about reputation, legal standing

Some VO examples

- From simpler to more complex

Web-based Collaboration

- Users decide to collaborate
- One user creates a (e.g.) wiki
  - Single resource of interest
- Wiki creator hands out user names and password to all the users
  - This user is now the authority
- Users put usernames and passwords into their web browsers
  - The browser is their wallet

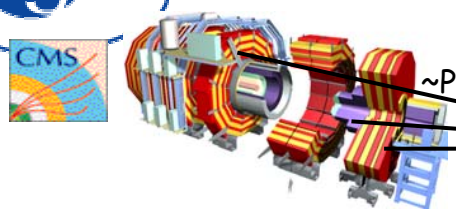
Web-based Community

- One organization brings together users from multiple organizations
  - E.g. IEEE, ACM, AMA
- Organization instantiates a web resource
- Organization creates and hands user names and passwords…
- And users add to their browser wallets

Previous examples

- Both had users from multiple institutions
- Both had only a single resource provider
- Ultimately all the policy was created and enforced in one place
- Moving on to a more complex example…

LHC Data Distribution



~PBytes/sec

Online System

~100 MBytes/sec

There is a "bunch crossing" every 25 nsecs.
There are 100 "triggers" per second
Each triggered event is ~1 MByte in size

1 TIPS is approximately 25,000 SpecInt95 equivalents

Offline Processor Farm
~20 TIPS

~100 MBytes/sec

Tier 0

CERN Computer Centre

~622 Mbits/sec
or Air Freight (deprecated)

Tier 1

France Regional Centre

Germany Regional Centre

Italy Regional Centre

FermiLab ~4 TIPS

~622 Mbits/sec

Tier 2

Caltech
~1 TIPS

Tier2 Centre
~1 TIPS

~622 Mbits/sec

Institute
~0.25TIPS

Physics data cache

~1 MBytes/sec

Tier 4

Physicist workstations

Physicists work on analysis "channels".
Each institute will have ~10 physicists working on one or more channels; data for these channels should be cached by the institute server

NCSA

LHC VO Example

- Users and resources from multiple organizations
  - Resources are computers, scientific instruments, storage, datasets, etc.
  - Often non-web based
- With multiple resource providers there is no longer a single obvious authority

LHC (cont)

- VO picks (and/or establishes) authorities for identity and attributes
- Lots and lots of policy discussions and (hopefully) agreements
- All resources in VO trust authorities
- An attribute authority is established
  - Distributes attribute assertions or a list of member identities
  - All resources trust this attribute authority

VO Challenges

- #1: Protocol and credentials
- Resource has to be able to recognize and authorize users
- Institutions have different credential formats and protocols
  - Passwords vs Kerberos vs LDAP vs Windows Domain
- Unlikely to have a ubiquitous solution any time soon

VO Challenges

- #2: Naming
- Users don't have global, unique names
- Each institution and service provider has their own name for each user
  - But it's hard to leverage these things across institutions (lack of protocols, common credentials)
- Same name at different institutions may be different users
- Names vs Identities
  - Often it's what you are, not who that is important

VO Challenges

- #3: Policy
- Expectation management
- How much effort must go into different operations?
  - How “secure” is “secure”?
- Who is responsible for what when things go wrong?
- How will the VO respond when things go wrong?

VO Challenges

- #4: Scalability
- From the VO perspective:
  - With enough members, it will take professional staff to manage membership, credentials, security services
  - Not a big problem for large VOs
    - E.g. IEEE can afford to set up services, hire staff, etc. to establish and maintain the VO
  - However for smaller VOs, this sort of overhead is an issue
    - E.g. scientific project often do not have the skills and expertise to operate a VO.

VO Challenges

- Scalability from the user perspective:
  - Each VO they are a part of means another name and set of credentials (e.g. username & password)
  - Browsers can solve a lot of this for the Web
    - Unless your disk crashes, you change computers, etc.
  - This is what the identity federation folks are targeting
    - E.g. Shibboleth, Liberty Alliance

Authentication in VOs

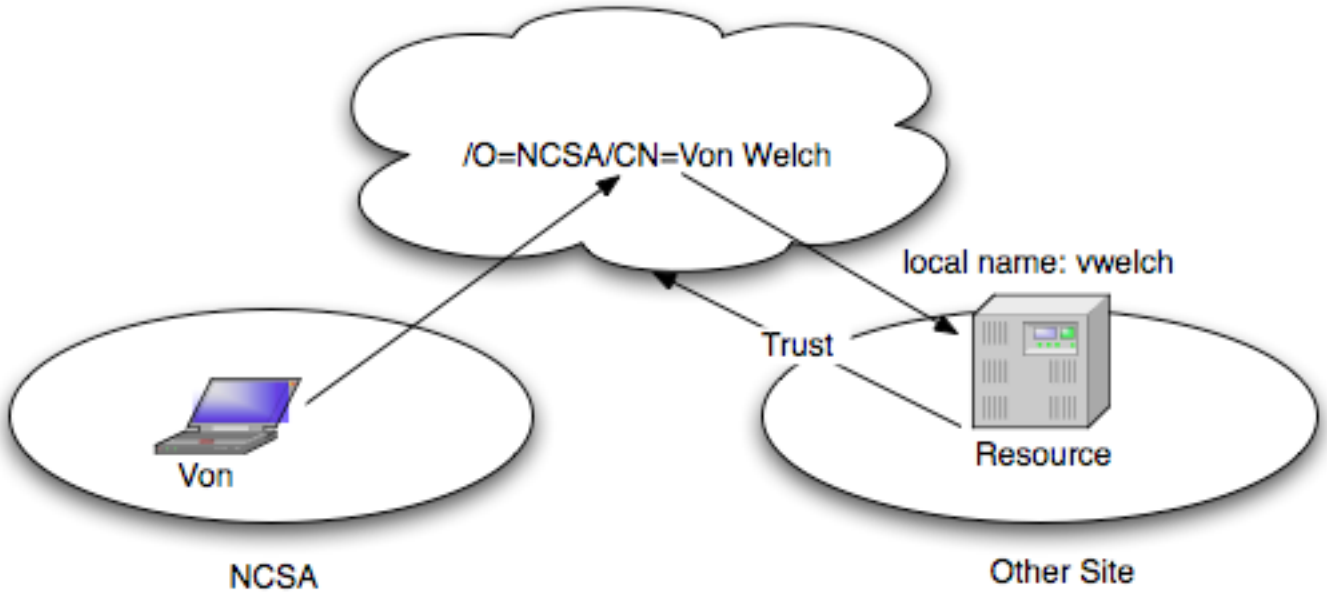
- Some history
  - Grids
  - Shibboleth
- GridShib Work To-Date
- Challenges ahead

Grids

- The Grid uses X509 for authentication and has a lot of experience
- Each user obtains an X509 certificate and identity
- Can be made to scale with enough effort. We have a world-wide trust federation.
  - http://www.gridpma.org
- This identity is that mapped to a local identity at each resource by the resource

X509 Global Namespace

Grid X509 Global Namespace



Advantages to Grid X509 approach

- Lightweight in that it doesn't require sites-to-site agreements
  - Allows a few users from a number of sites to collaborate in VOs without complicated peering
  - Each resource can accept the X509 certificates it wants

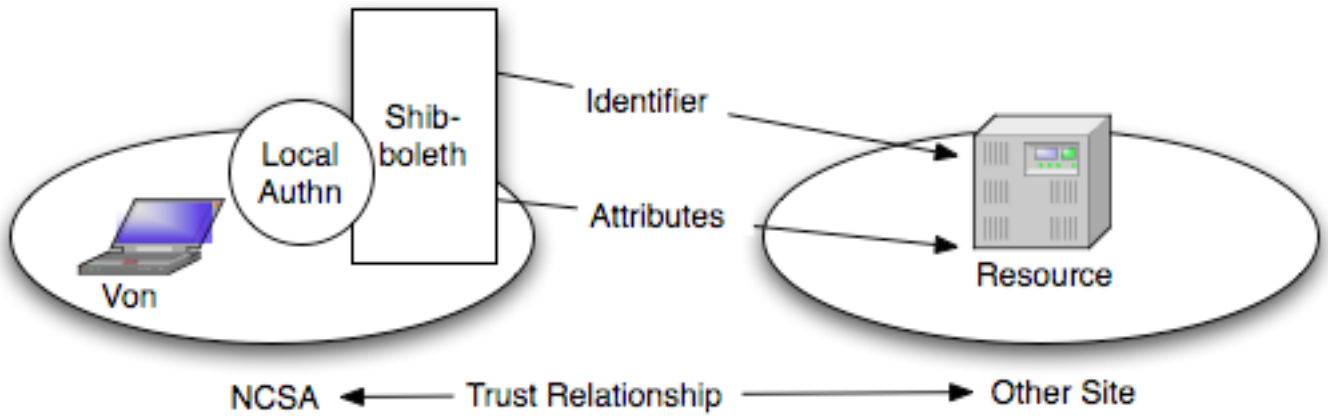
Disadvantages to Grid approach

- Heavyweight in that it buts credential management burden on users
- Users are poor managers of X509 private keys
  - Too long to memorize or write down
- No good place to store keys
  - No ubiquitous support for hardware tokens across multiple organizations
- Lost keys are painful to replace
- Can be hard to tell if a key was compromised
  - Hacker broke in, what keys were on the system?

Shibboleth

- Uses identity federation approach
  - Very much aligned with Liberty Alliance
  - Identity == what you are, not necessarily who
- Site-to-site trust arrangements allow for expressing identifiers and attributes across sites
- Features for privacy
  - Resource knows only what you are, not who

Shibboleth Id Federation



Advantages of Shibboleth

- Uses existing authentication system
  - No new credentials for the user to learn and manage
- Privacy
- XML-Buzzword-compliant
  - Might be an advantage, certainly hipper
- Flatter, simpler hierarchies than PKI
  - At least for now

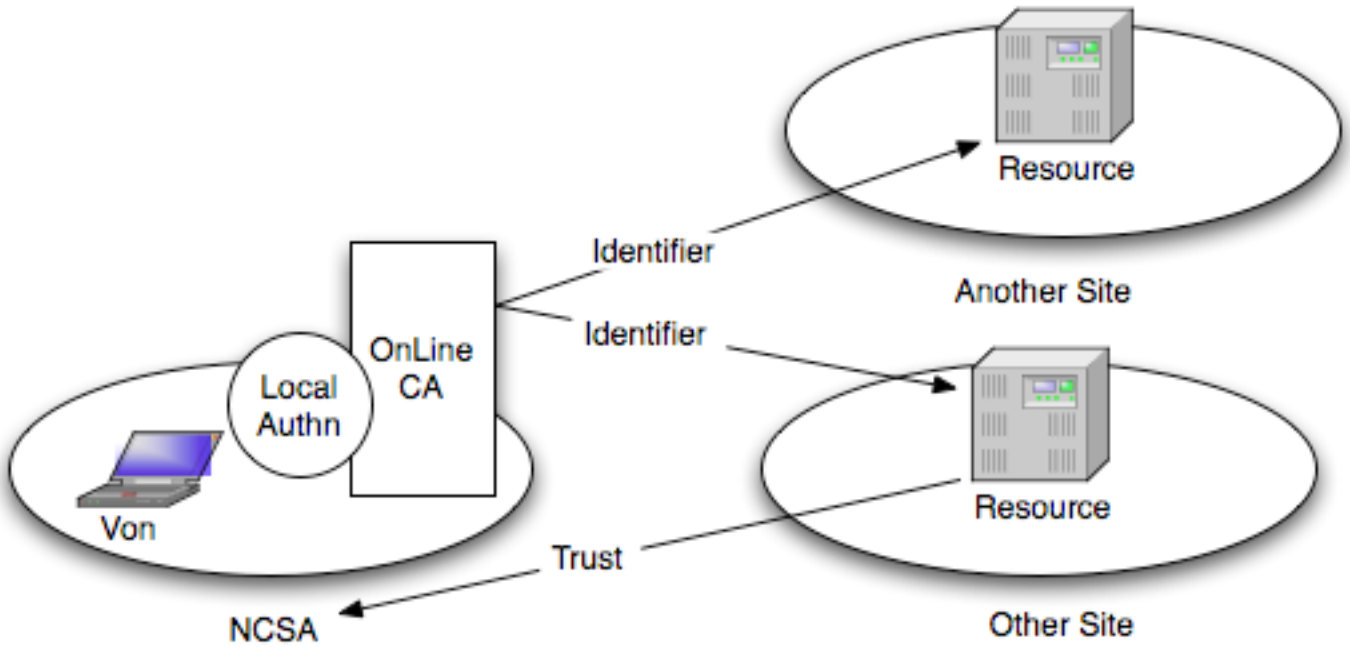
Disadvantages of Shibboleth

- Identity federation requires institutions to agree
  - Slower than user-to-user trust
  - Requires high-level of motivation to ensure that it will happen
  - Lawyers
- Technology is currently focused on web browser applications
  - Lack of delegation
  - Protocol assumes lots of browser features
    - Redirection, auto-refresh of credentials, etc.

The online CA Approach

- An alternative to traditional PKIs
- Online CAs leveraging existing institutional authentication
  - E.g. KCA, MyProxy
  - Deployments at FNAL, NERSC
- User uses local authentication to obtain short-lived X509 credential (with persistent name)

Online CA


Resource
Another Site
Identifier
Identifier
OnLine
CA
Local
Authn
Von
Resource
Trust
NCSA
Other Site

Online CA

- Advantages
  - No new passwords for the users
  - Works with existing Grid infrastructure
- Disadvantages
  - Still have short-lived credential. Is it short-lived enough we can ignore revocation?

On to GridShib…

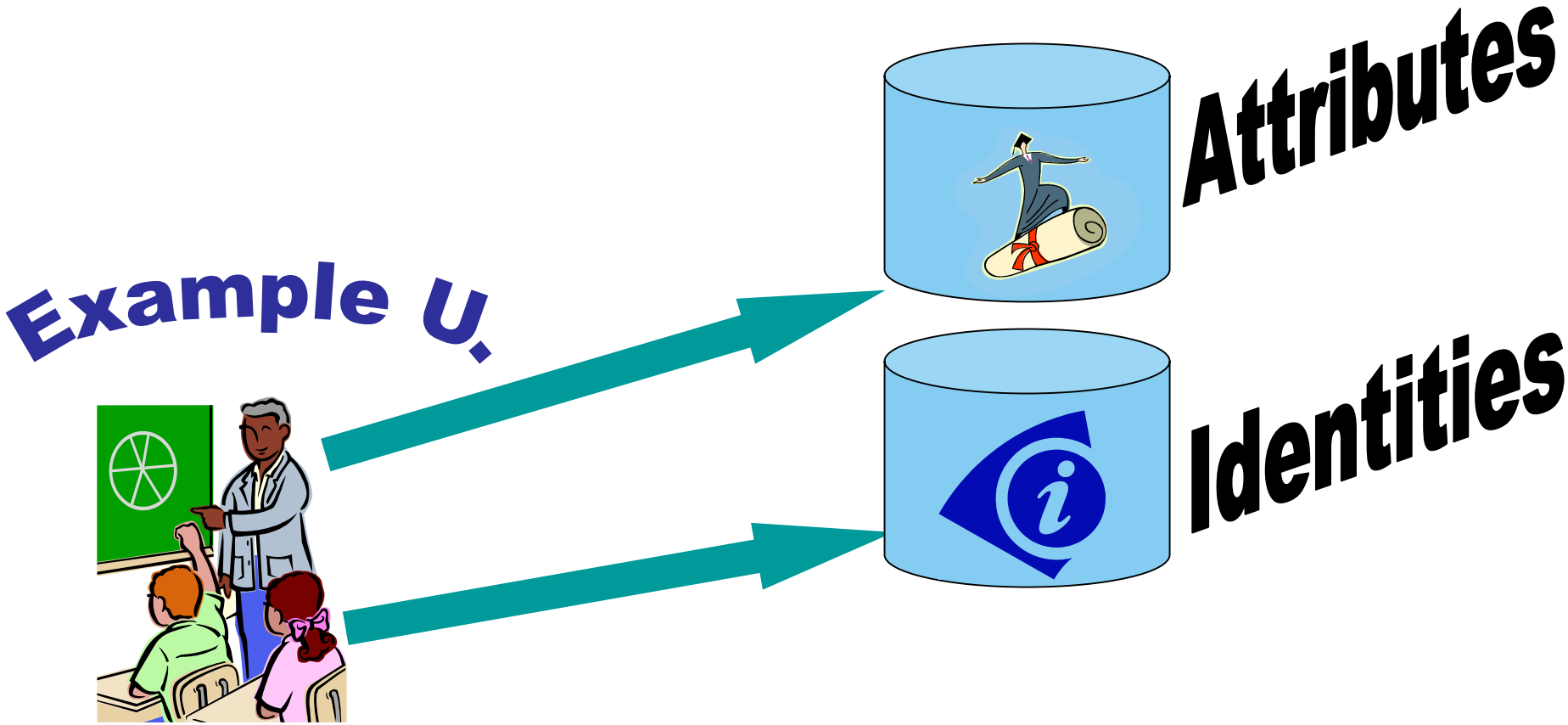
What is GridShib

- NSF NMI project to allow the use of Shibboleth-issued attributes for authorization in NMI Grids built on the Globus Toolkit
  - Funded under NSF NMI program
- GridShib team: NCSA, U. Chicago, ANL
  - Tom Barton, David Champion, Tim Freemon, Kate Keahey, Tom Scavo, Frank Siebenlist, Von Welch
- Working in collaboration with Steven Carmody, Scott Cantor, Bob Morgan and the rest of the Internet2 Shibboleth Design team

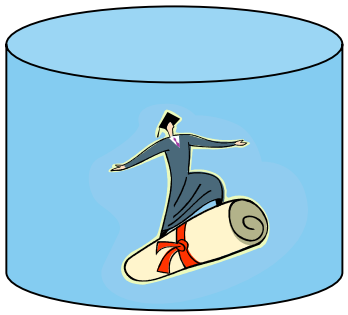
Motivation

- Many Grid VOs are focused on science or business other than IT support
  - Don’t have expertise or resources to run security services
- We have a strong infrastructure in place for authentication in the form of Grid PKIs
- Attribute authorities are emerging as the next important service

Campus Infrastructure



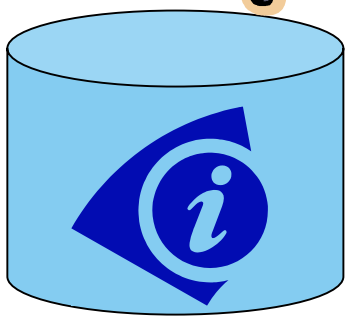
Example U.



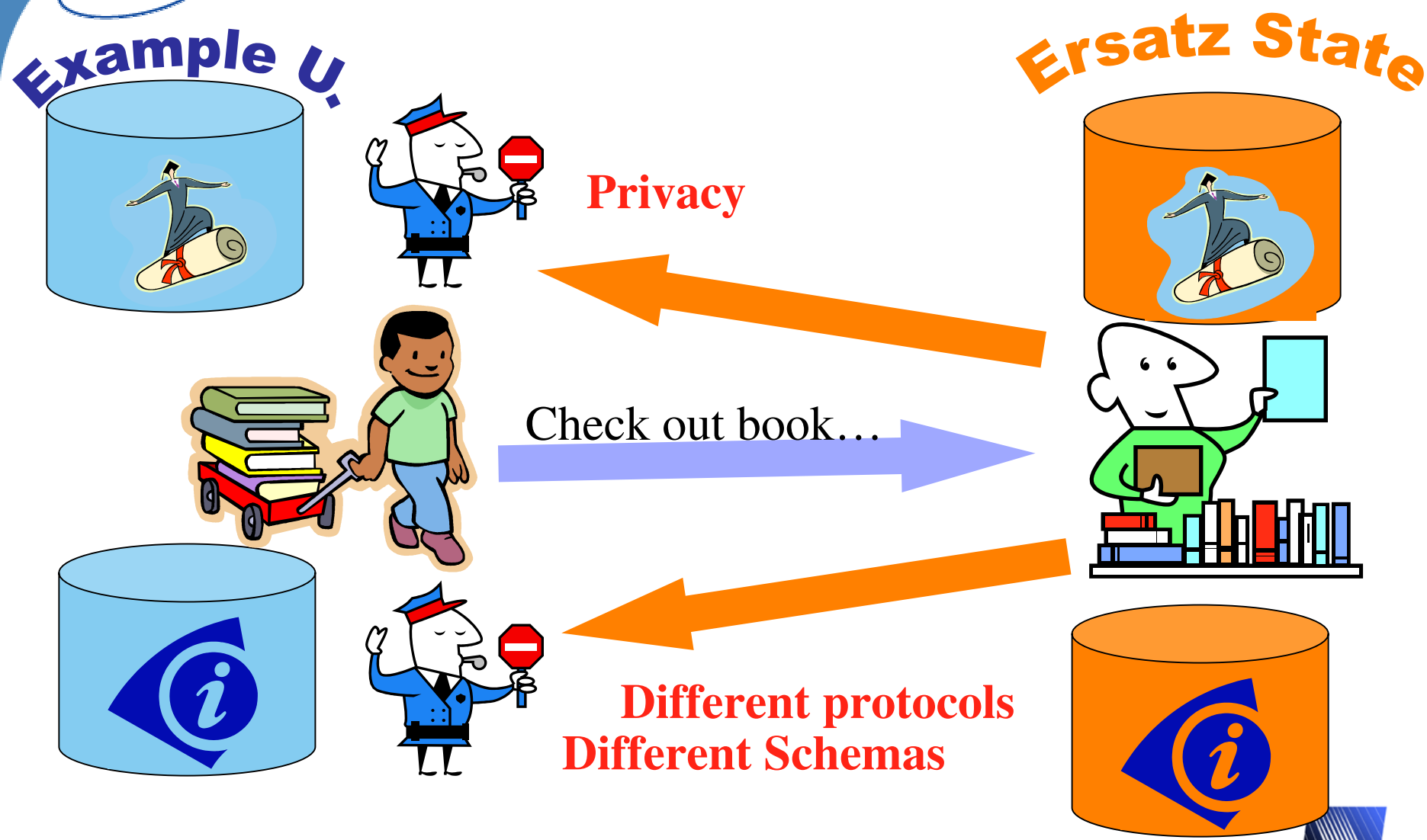
Student?

Check out book…

Access student records…



Is student John Smith?


Example U.
Ersatz State
Privacy
Check out book…
Different protocols
Different Schemas

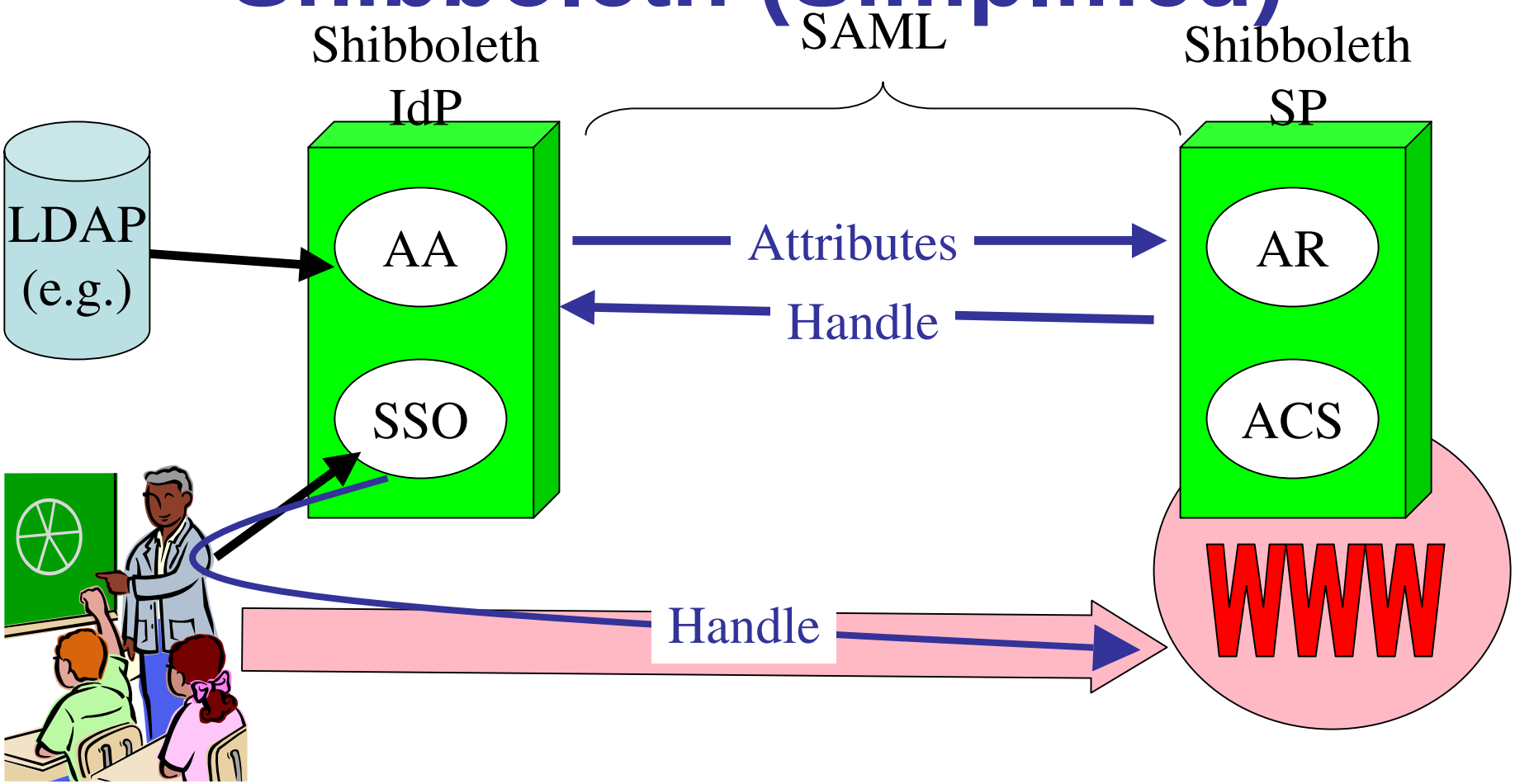
Shibboleth

- http://shibboleth.internet2.edu/
- Internet2 project
- Allows for inter-institutional sharing of web resources (via browsers)
  - Provides attributes for authorization between institutions
- Allows for pseudonymity via temporary, meaningless identifiers called ‘Handles’
- Standards-based (SAML)
- Being extended to non-web resources

Shibboleth

- Identity Provider composed of single sign-on (SSO) and attribute authority (AA) services
- SSO: authenticates user locally and issues authentication assertion with Handle
  - Assertion is short-lived bearer assertion
  - Handle is also short-lived and non-identifying
  - Handle is registered with AA
- Attribute Authority responds to queries regarding handle

Shibboleth (Simplified)



Globus Toolkit

- http://www.globus.org
- Toolkit for Grid computing
  - Job submission, data movement, data management, resource management
- Based on Web Services and WSRF
- Security based on X.509 identity- and proxy-certificates
  - Maybe from conventional or on-line CAs
- Some initial attribute-based authorization

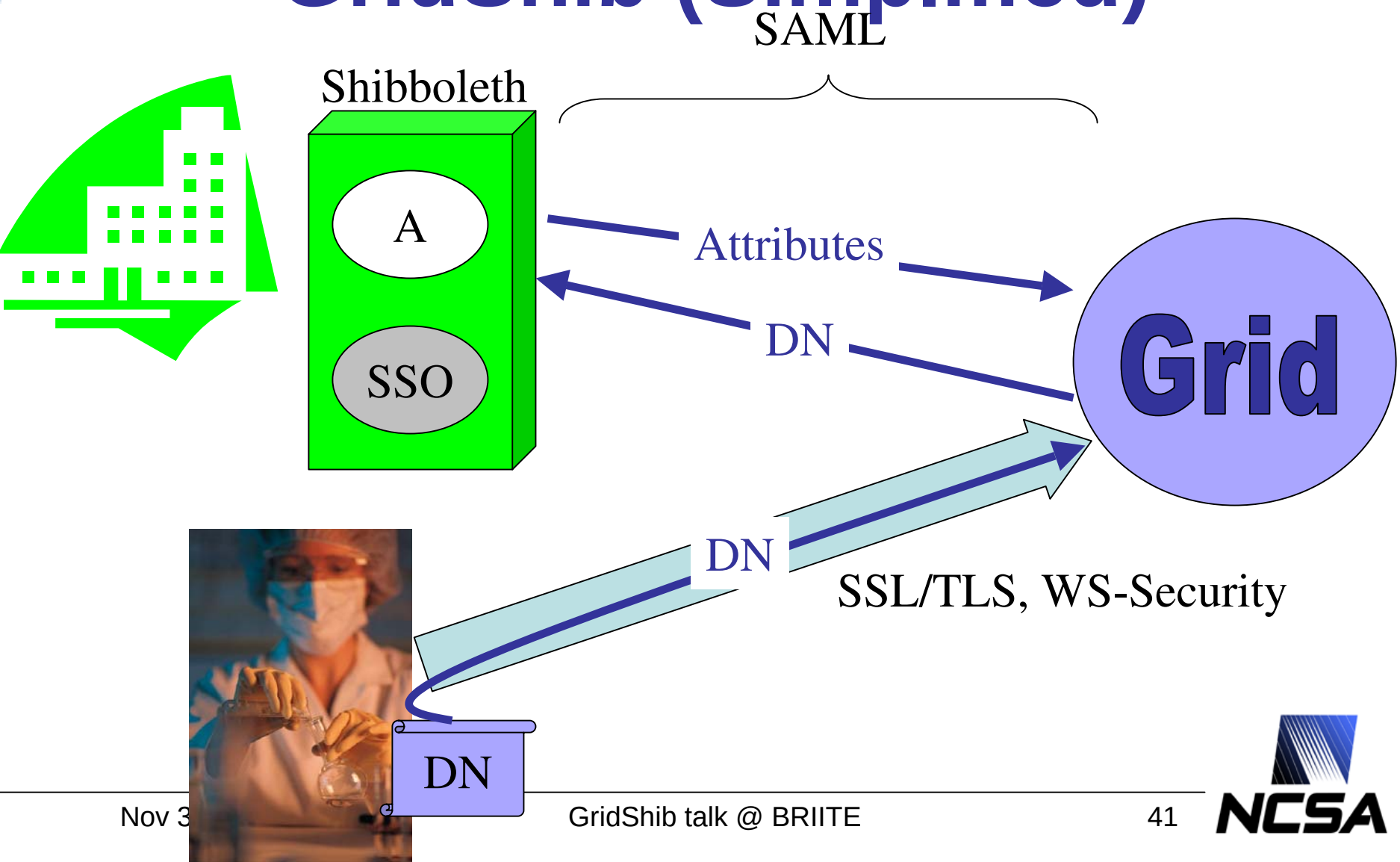
Grid PKI

- Large investment in PKI at the international level for Grids
  - http://www.gridpma.org
  - TAGPMA, GridPMA, APGridPMA
  - Dozens of CAs, thousands of users
- Really painful to establish
- But it's working…
  - And it's not going way easily

Integration Approach

- Conceptually, replace Shibboleth's handle-based authentication with X509
  - Provides stronger security for non-web browser apps
  - Works with existing PKI install base
- To allow leveraging of Shibboleth install base, require as few changes to Shibboleth AA as possible

GridShib (Simplified)



Authorization

- Delivering attributes is half the story…
- Currently have a simple authorization mechanisms
  - List of attributes required to use service or container
  - Mapping of attributes to local identity for job submission

Authorization Plans

- Develop authorization framework in Globus Toolkit
  - Siebenlist et. al. at Argonne
  - Pluggable modules for processing authentication, gathering and processing attributes and rendering decisions
- Work in OGSA-Authz WG to allow for callouts to third-party authorization services
  - E.G. PERMIS
- Convert Attributes (SAML or X509) into common format for policy evaluation
  - XACML-based

GridShib Status

- Beta release publicly available
- Drop-in addition to GT 4.0 and Shibboleth 1.3
- Project website:
  - http://gridshib.globus.org
- Very interested in feedback

Challenges Ahead…

Distributed Attribute Admin

- The Problem…
- NCSA runs the attribute authority
- But lots of people issue attributes about me
  - IEEE, ACM, TeraGrid, GridShib, etc.
  - Every group I'm a member of is an attribute
  - Many of these group are their own authority
- Think of all the credentials in your purse or wallet…

Distributed Attribute Admin

- Many of these groups will simply set up their own attribute service
- Two issues:
  - Users need a way to manage this virtual wallet
    - What attribute authorities should be consulted when - what are my roles at the moment?
  - Some groups are too small to set up their own attribute services

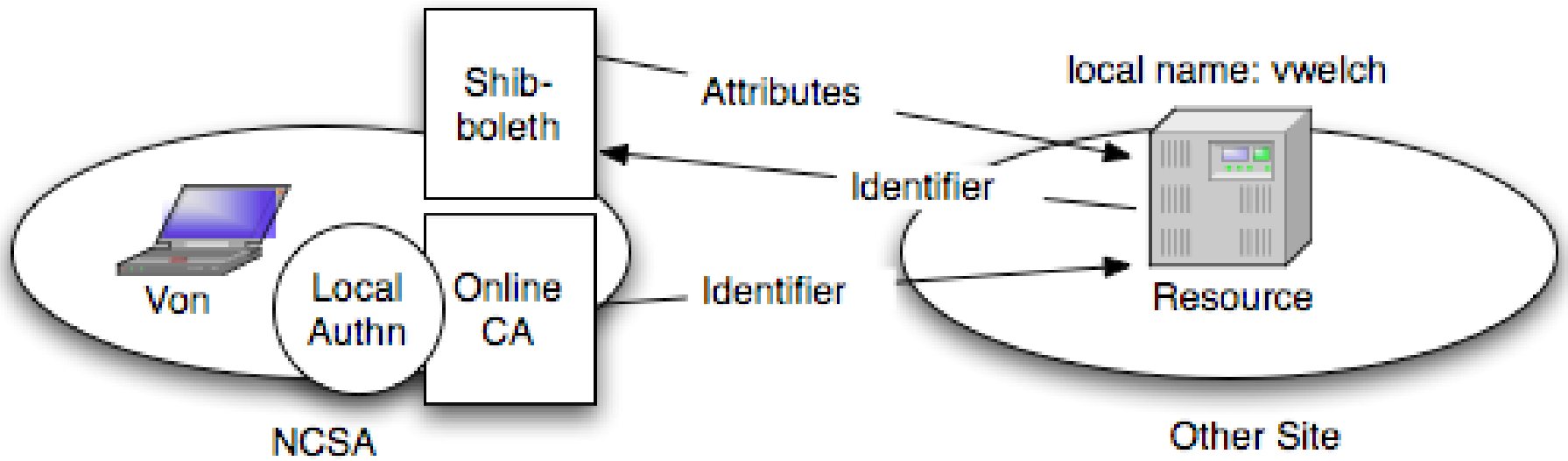
Distributed Attribute Admin

- Need ways for a user to point at the attributes services they want to be consulted
  - Push attributes?
  - Push references to attribute authorities?
  - We exploring both of these paths
- Signet/Grouper integration for distributed attribute administration
  - Tom Barton @ U. of Chicago
  - Allow small groups to set attributes in your attribute server
  - Technical issues, probably bigger policy issues

GridShib/Online CA Integration

- X509 Credentials still have large problem with user-managed credentials
  - See slide 21
- Use of online CA at campus to issue credentials helps with this
- If we integrate an online CA such that the identifiers it issues can then be used to get attributes from a Shibboleth AA we get a full attribute-based authorization system
- Collaboration with Jim Basney

GridShib/MyProxy Integration



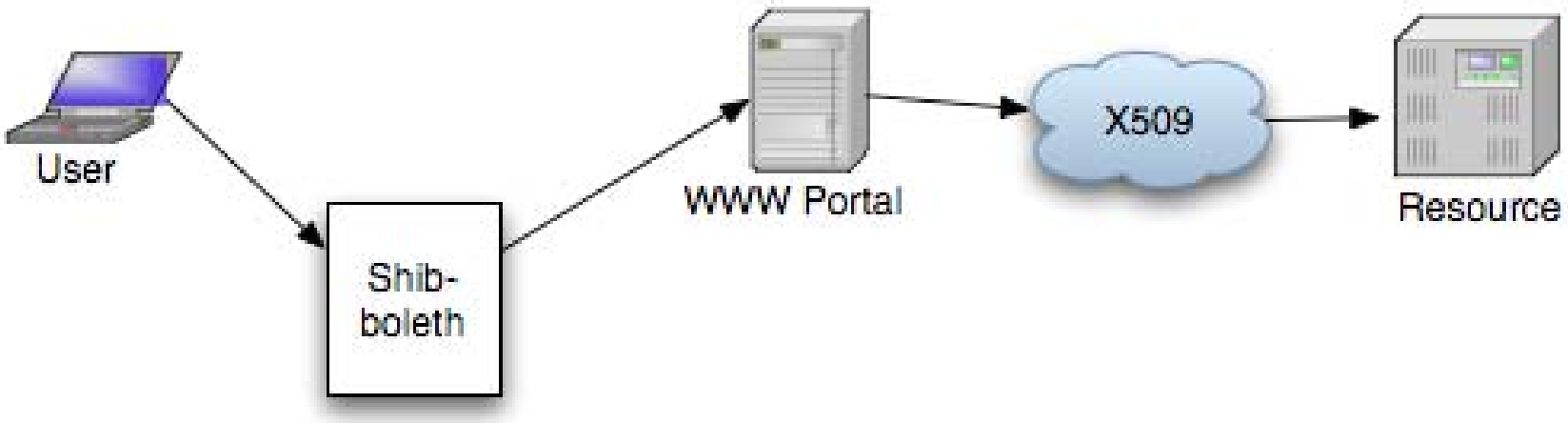
GridShib/MyProxy Integration

- Challenge is one of name management
- User's local name must be mapped to X509 DN and then back to name meaningful to attribute authority
- Is algorithmic approach better or can we assume database of mappings?
- Who should do the mappings?

Grid Portals

- Web portals are important
  - Clients already installed, easily customized, users familiar with them
- But protocols are rather difficult to customize
  - There is a rich set of features, but adding new features (for security) or otherwise is difficult
  - Lots of portal developers to convince

Grid Portals



SAML

X509

Thank You

- My email:
  - vwelch@ncsa.uiuc.edu
- GridShib
  - http://gridshib.globus.org
- Shibboleth
  - http://shibboleth.internet2.edu/
- Globus Toolkit
  - http://www.globus.org/
- MyProxy
  - http://myproxy.ncsa.uiuc.edu/