The Rise of Middleware

Ken Klingenstein, Director
Internet2 Middleware Initiative

Outline

Acknowledgments

Basics –

- *What does middleware do*
- *Middleware and Advanced Applications*
- *What are the technologies in middleware*

What's Been Happening in Higher Ed and NMI

Next steps

http://middleware.internet2.edu/

National Science Foundation

Catalytic grant in Fall 99 started the organized efforts, with Early Adopters and Early Adopters

NSF Middleware Initiative - three year cooperative agreement, begun 9/1/01, with Internet2/EDUCAUSE/SURA and the GRIDs Center, to develop and deploy a national middleware infrastructure for science, research and higher education

Work products are software, community standards, best practices, schema and objectclasses, reference implementations, open source services, corporate relations

Work areas are identifiers, directories, authentication, authorization, GRIDs, PKI, video

Other partnerships

US Government Agencies - NIST, NIH

Educational IT Organizations - EDUCAUSE, CREN, etc.

International Standards Groups - OASIS, IETF, etc.

Corporate Partners - Sun, IBM, Polycom, etc.

International Networking Associations –TERENA, JISC, SURFnet, REDIRIS, etc.

MACE (Middleware Architecture Committee for Education)

Purpose - to provide advice, create experiments, foster standards, etc. on key technical issues for core middleware within higher education

Membership - Bob Morgan (UW) Chair, Scott Cantor (Ohio State), Steven Carmody (Brown), Michael Gettes (Georgetown), Keith Hazelton (Wisconsin), Paul Hill (MIT), Jim Jokl (Virginia), Mark Poepping (CMU), Bruce Vincent (Stanford), David Wasley (California), Von Welch (Grid)

European members - Brian Gilmore (Edinburgh), Ton Verschuren (Netherlands), Diego Lopez (Spain)

Creates working groups in major areas, including directories, interrealm access control, PKI, medical issues, etc.

Works via conference calls, emails, occasional serendipitous in-person meetings...

What is Middleware?

specialized networked services that are shared by applications and users

a set of core software components that permit scaling of applications and networks

a second layer of the IT infrastructure, sitting above the network

a land where technology meets policy

the intersection of what networks designers and applications developers each do not want to do

Scenarios 1 & 2

1. *A member of a university can access digital content regardless of their location anywhere around the world. There is no proxy (or proxy problems) and the access can be controlled by user characteristics (e.g law student, enrolled in Physics 101, etc.) while maintaining privacy.*
2. *A Native American tribe builds an on-line museum. Some artifacts are open to all browsers; some areas should be open only to educators; some (such as tribal songs) should be open only to tribal members, and be available for annotation by individual tribal members.*

Scenarios 3 & 4

3. A campus has licensed on-line content from a museum. The contract permits students to extract up to a 20 sec clip from museum video archives for inclusion in a term paper. Technology to control this access is available to any qualified manager as a simple set of pull down options.

4. A group of scientists at different universities are managing a shared research site. Access to the initial data and instruments is restricted to specific individuals at first, then a few graduate classes are given access to some content. Students in local high school physics classes are permitted to access some of the instruments in a restricted fashion. This is all done transparently to the students, the individuals, etc. All merely login at their home organization as usual.

Scenarios 5&6

5. Particle physics researchers around the world share the data and computational capabilities of CERN and other major international facilities. A variety of machines and services are linked in a seamless mesh; for users, their computing jobs transparently find their data sets and computers to run them on.

6. A user sitting at a remote campus can utilize a shake table sitting at a major university. The user can monitor the table; watch real time as the forces are applied, and control access for other remote participants.

Specifically…

Digital libraries need scalable, interoperable authentication and authorization.

The Grid is a new paradigm for a computational resource; Globus provides middleware, including security, location and allocation of resources, and scheduling. This relies on campus-based services and inter-institutional standards.

Instructional Management Systems need authentication and directories.

Next-generation portals want common authentication and storage.

Academic collaboration requires sharing of restricted materials between institutions.

Core Middleware

Identity - unique markers of who you (person, machine, service, group) are

Authentication - how you prove or establish that you are that identity

Directories - where an identity's basic characteristics are kept

Authorization - what an identity is permitted to do

PKI - emerging tools for security services

Importance to Internet2 Members

Many important applications such as videoconferencing, and interoperable instant messaging could become ubiquitous with identifiers and authentication.

Inter-institutional collaborations require interoperational deployments of institutional directories and authentication.

Core middleware underpins advanced distributed computing environments such as Grids.

Authentication and authorization will be required to implement network based services such as QoS and secure multicast.

Urgent needs in higher education for mobility of users and systems can be addressed with middleware,

Academic requirements for privacy and scholarship need electronic implementation.

Importance to Internet1 Users

Today's Internet lacks a common, standards-based, interoperable middleware layer. Thus there are no

general purpose airport computing kiosks that can be instantly configured with a user's email aliases, bookmarks, etc.

effective, easy to use tools to protect your personal privacy on the web, in email, etc

digital signature technologies in wide usage

general e-commerce collaborations where corporations can exchange internal authorizations and electronic services

Advancing Applications

Internet2 is not just about applications that need advanced networking (typically high bandwidth, perhaps low latency, multicast, etc).

Internet2 is also about enabling new applications through development and deployment of middleware

Sometimes the applications being advanced are "pedestrian" (web browsing, desktop video, etc) and sometimes they may be "high-end" (CAVE's, Grids, etc.)

What I1 did for network connectivity, I2 may do for human collaboration…

Selling infrastructure: "It may be true that you don't know what you have til it's gone, but it's also true that you don't know what you've been missing til it comes…"

Core Middleware Basics

Technical components include: authentication and authorization, directories, community objectclasses, security credentials, identifier management, etc.

Policy components include: trust models, privacy legislation and regulations, community standards developments, deployment models, etc.

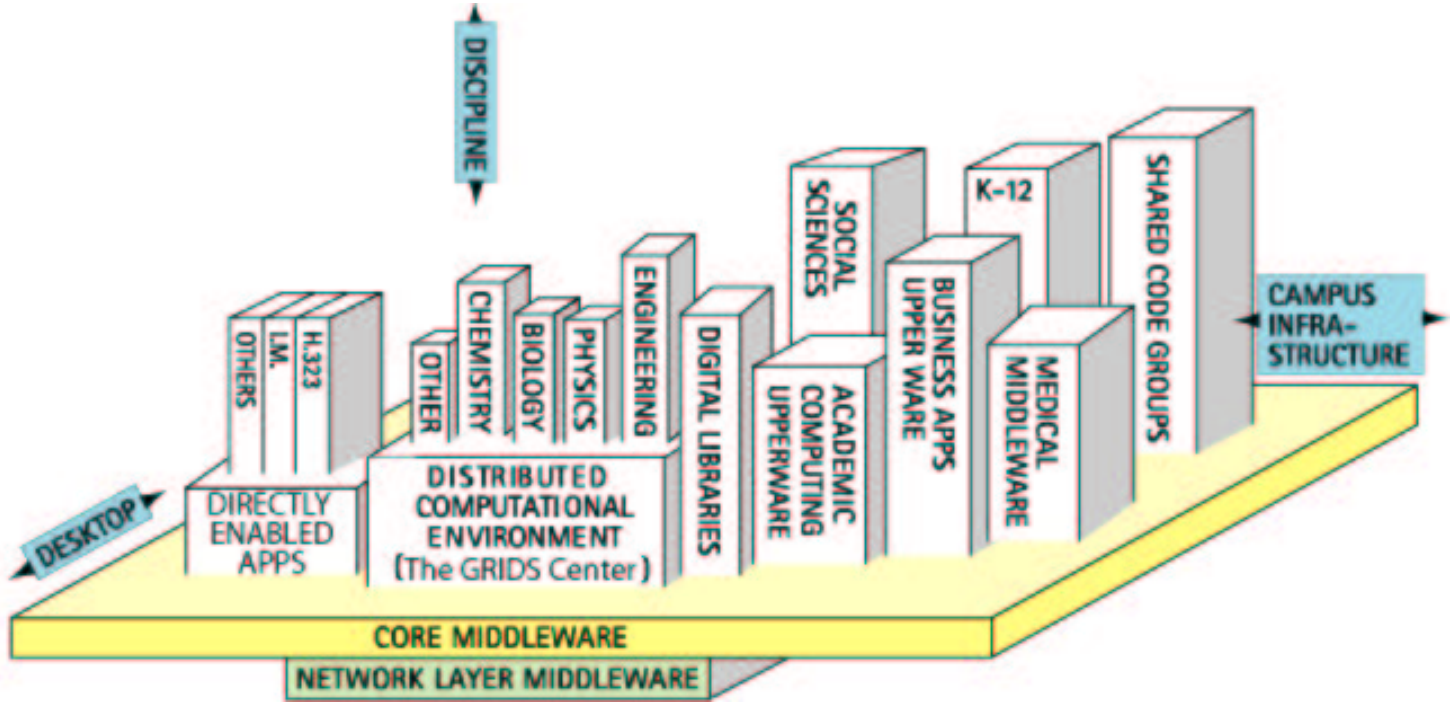
Middleware enables

ease of access/ease of control for digital content

security with privacy and accountability

scalable usability for advanced apps such as Grids, digital rights management, desktop video, etc.

A Map of Middleware Land



What Are Our Activities

Consensus building - overall architectural model for middleware

Standards setting - standards in directory objectclasses, naming, security approaches, interrealm exchange of attributes

Technology transfer - from leading edge institutions to other universities; from higher ed labs to corporate sector

Fostering of basic research - PKI Labs, protocol development in XML, interrealm metadirectory tools, etc.

Tool development – web single sign on, directory anaylzers, interrealm authorization

Activities in Higher Ed

First, a standardization on the information that institutions might exchange for collaborations

eduPerson 1.5, eduOrg 1.0

Then an architecture and an open source implementation to exchange that information in a secure but privacy-preserving standards based technology.

Shibboleth

Then the NSF Middleware Initiative to accelerate and disseminate the work and integrate it to other science

Soon applications in video and DRM that leverage these approaches

With increasing and consistent middleware deployments on campuses

NMI Release 1 (05/07/02)

Globus Toolkit 2.0, Condor-G

Network Weather Service, KX.509

Pubcookie 2.0

eduPerson 1.0, commObject

Architecture docs for videoconferencing, Shibboleth

Best Practices in Campus Directories, Groups in Directories, metadirectories, etc.

Sample Policies for PKI, campus account management, etc.

NMI Release 2.0 (10/25/02)

Updates to Globus Toolkit 2.0

Shibboleth 1.0

LDAP Analyzer 1.0 – a tool for directory testing

Pubcookie 3.0

eduPerson 1.5 and eduOrg 1.0 final

Architecture docs for interrealm metadirectories, digital rights management, etc.

Updates on Best Practices and Policies

Campus infrastructure: local and for interrealm collaboration

Basic:

Enterprise directory services with feeds from core legacy systems and driving most enterprise applications

Campus-wide name spaces and authentication

Appropriate policies for identity, permissions, etc.

Interrealm:

Application-specific enterprise-wide directories

eduPerson and eduOrg implemented

Interrealm authentication and authorization tools

Federation

Activities in the marketplace

Digital Identity is the center of attention

http://www.digitalidworld.com/

(note the article on Shibboleth: Identity the Internet Way)

Microsoft Passport lurches towards federation

Liberty Alliance grows large and issues first specs

http://projectliberty.org

If we wave our hands fast enough, will pigs fly?

Interrealm authorization: current approaches

Lots of ad hoc, non-scalable, difficult to maintain, and restrictive approaches

Content providers limit access by IP address, leaving campus users on cable modems at home frustrated…

Users get new userids and passwords in each realm, and then set all their passwords to be the same…

Campuses operate proxy services that inconvenience users and present performance bottlenecks.

Campuses load user identities into content provider databases, incurring additional cost, stale data, and the potential for privacy violations

Shared passwords are distributed, perhaps widely, presenting significant security risks

Shibboleth Basics

“Interrealm Attribute-based Authorization for Web Services”

An initiative to develop an architecture, policy framework, and practical technologies to support inter-institutional sharing of resources

Provides the secure exchange of interoperable attributes which can be used in access control decisions

Controlled dissemination of attribute information, based on administrative defaults and user preferences

Shifts the model from passive privacy towards active privacy

Based on a federated administration trust framework

Vendor participation - IBM/Tivoli

Standards Alignment - OASIS/SAML

Open solution(protocols and messages documented rfc-style, open source implementation available)

Stage 1 - Addressing Three Scenario's

Member of campus community accessing licensed resource

- Anonymity required

Member of a course accessing remotely controlled resource

- Anonymity required

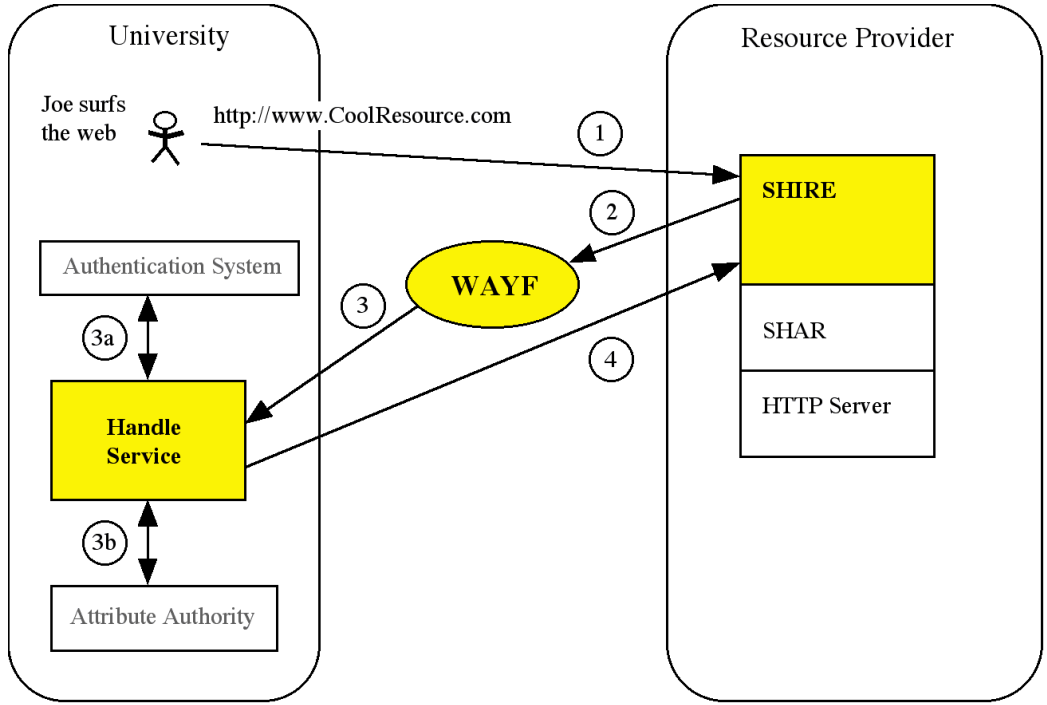
Member of a workgroup accessing controlled resources

- Controlled by unique identifiers (e.g. name)

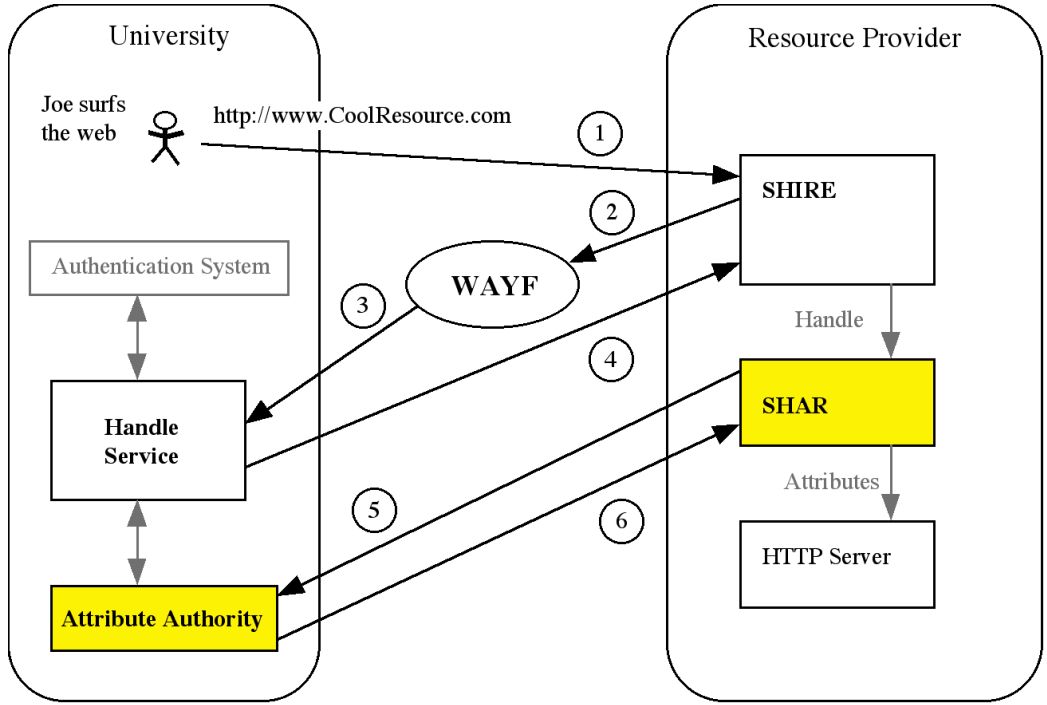
Taken individually, each of these situations can be solved in a variety of straightforward ways.

Taken together, they present the challenge of meeting the user's reasonable expectations for protection of their personal privacy.

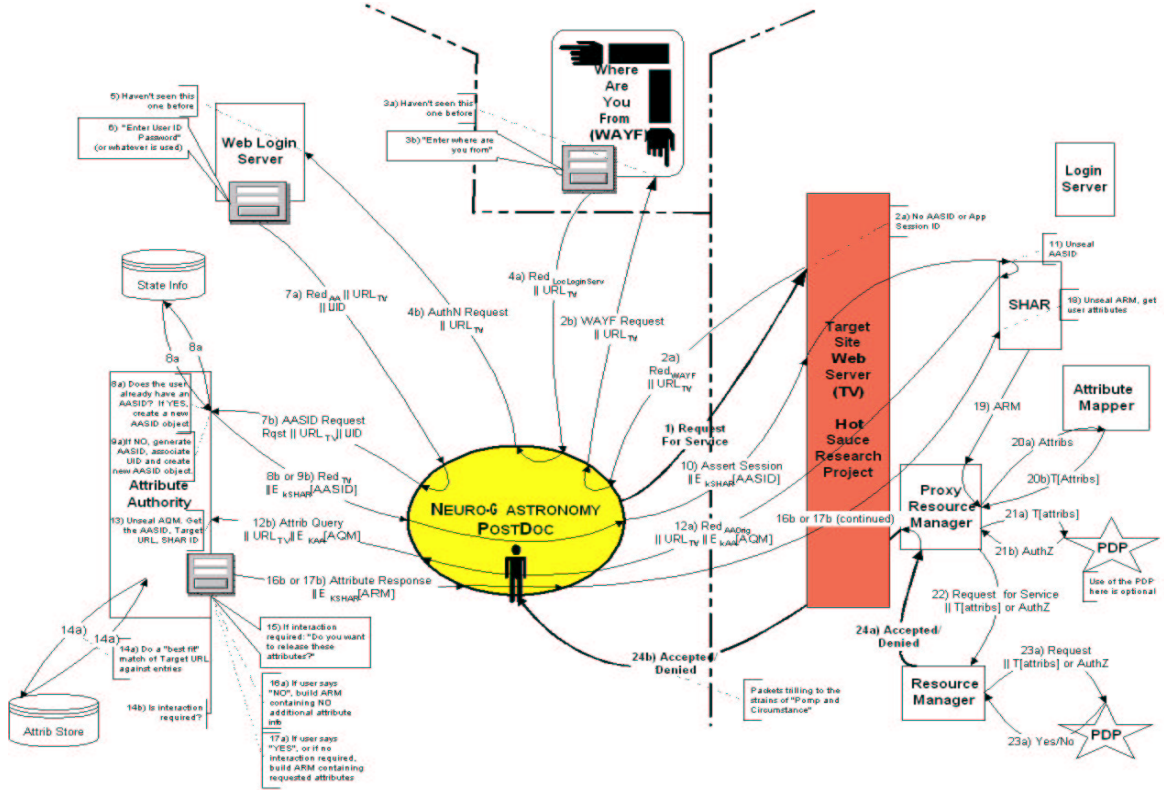
Establishing a User Context



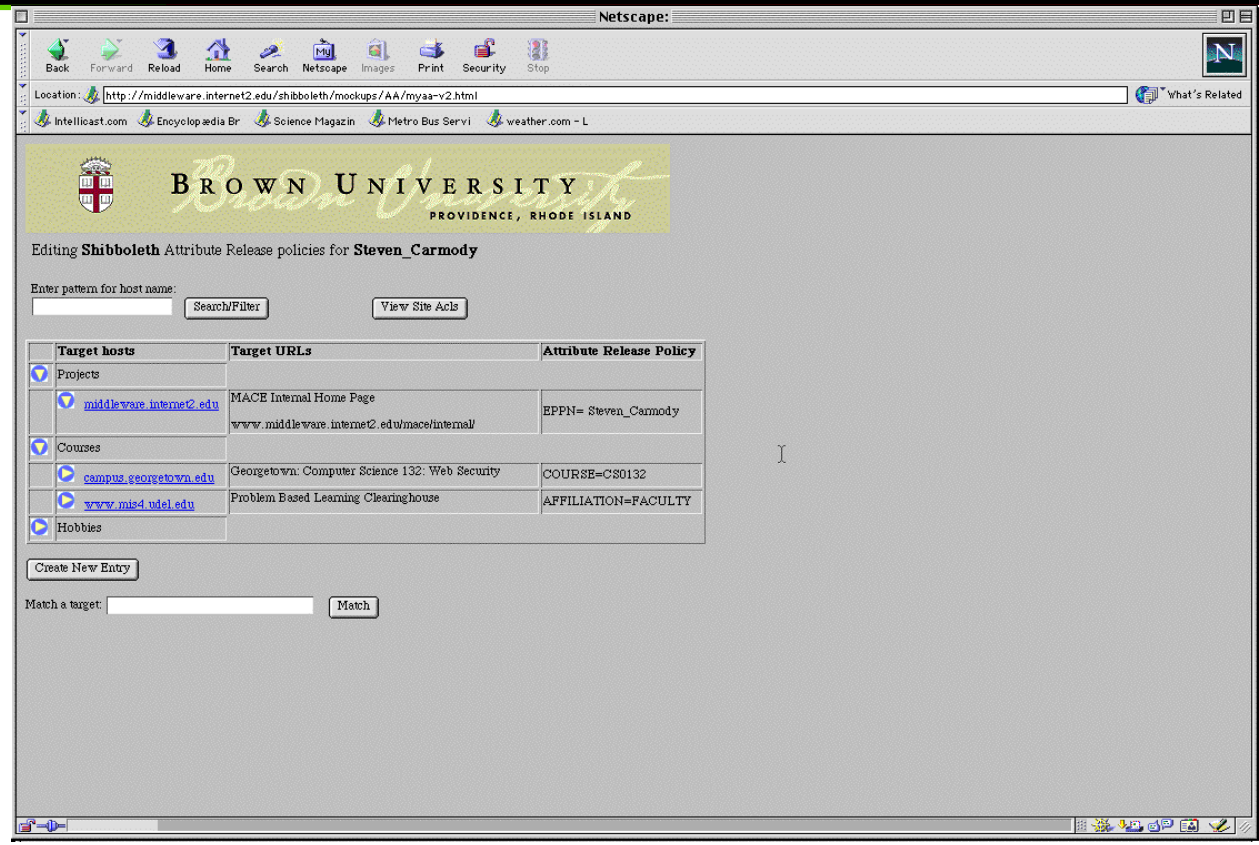
Getting Attributes and Determining Access



Shibboleth Flows Draft



User Attribute Management



Milestones

Project formation - Feb 2000 Stone Soup

Process - began late summer 2000 with bi-weekly calls to develop scenario, requirements and architecture.

Linkages to SAML established Dec 2000

Architecture and protocol completion - Aug 2001

Design - Oct 2001

Coding began - Nov 2001

Alpha-1 release – April 24, 2002

OpenSAML release – July 15, 2002

Alpha-2 release – July 20, 2002

Alpha-2.5 release – Aug 19, 2002

Library pilot

EBSCO
ProQuest
OCLC
Elsevier
SFX

California
Colorado
Columbia
EDINA
Georgetown
London School of Economics
Michigan
Ohio State
Penn State
Univ of Washington
etc

The next three months and then…

Pilots commence around Oct 1 (some already underway)

Beta-1 code (the real stuff) Sept 15 – CMU, OSU, MIT

Beta-2 code (the real stuff, fixed) Oct

Shib 1.0 released as part of NMI – R2 - Oct 27,2002

- *Core Shib subsystem 1.0*
- *Resource Managers 0.5*
- *Attribute Release Managers 0.5*

Post 1.0

- *Shib 1.1 and 1.2*
- *ARM and RM development*
- *FDRM development*
- *Video…*

Next Steps

Follow the work

http://middleware.internet2.edu/

http://www.nsf-middleware.org

the marketplace in digital identity and privacy

Begin the community discussions

what is the community? What standards does it need? What external standards should it adopt? With what other communities will you federate? What processes should govern and manage these developments?