Summary of the NMI Advanced CAMP on Virtual Organizations
Denver, June 29 – July 1, 2005

1. Background:

As part of their NSF award, the NMI-EDIT Team, (consisting of Internet2 and Educause and partners) conducts periodic meetings called CAMPs, where developers and interested parties can assemble to promote the creation and deployment of campus middleware. An Advanced CAMP with a focus on virtual organizations and enterprise support was held in Denver, June 29-July 1, 2005.

2. Model and Terms:

Virtual organizations are a description for the collections of scientists, researchers, and scholars who collaborate and share meaningful resources. They are differentiated from brick and mortar organizations because they are not typically a primary employer or equipment, network and identity provider for the participants. They are differentiated from simple affinity groups, such as blogs, usenix groups, IM buddy lists, etc (though VOs frequently use such tools), because virtual organization share substantive resources such as computers, scientific instruments, data, etc. and may be subject to audit and compliance requirements.

VOs tend to cluster around notable sites, such as CERN and SLAC, that provide data and computational science resources. VOs may also form around distinctive scholarly resources, such as the Library of Congress and the British Museum. They could also comprise a small group of collaborators working with shared resources.

The enterprise-leveraged model being considered to supply support services to VOs includes several components:

The user. It is assumed that the user has several persona – their lives in their enterprise, the virtual organization, and a personal set of electronic identities which they may want to integrate to some degree.

The enterprise, and in many cases, a federation (and its peer federations) that the enterprise belongs to. The enterprise is assumed to be the primary employer of the collaborator/researcher and provisions most of the user’s basic IT services (network connectivity, workstation and basic software, identity management, etc.) Across the enterprise, the federation and its peers, support services from trust to bandwidth can be administered.

The virtual organization. The primary purpose of the VO is to conduct scholarly work, often through a shared set of facilities, researchers, grants, etc. Besides managing a set of resources, the organizational members spend considerable effort in collaboration, from sharing of documents, code and data to messaging, videoconferencing, listprocs and other communications.
A virtual organization service center. Such a centrally operated facility serves several purposes. It acts a registry for sharing of virtual organization metadata among enterprises and federations. It provides a place for user support in those instances where the user’s home enterprise is unable or unwilling to provide virtual organization assistance. The service center can also coordinate training and support for both virtual organization leadership and enterprise IT staff.

2.1 Caveats

Though there was rough consensus around the model described above, not everyone agreed with the approach. In particular, there is a community that feels that virtual organizations should stand apart from campus infrastructure as much as possible.

Much research in some disciplines, such as the humanities, is done by an individual or a small group of collaborators.

The model assumes a collective interest (federal agency, institution, researcher and virtual organization) in improving the sustainability and scalability of virtual organizations. That may not the case, and even within an area, such as an enterprise, interests of institutional participants may conflict.

There was a strong collective sense that virtual doesn't work unless flesh and blood meets flesh and blood, hopefully early in the collaborative process.

3. Requirements:

Much of the meeting was devoted to identifying use cases and their resulting requirements for components of the model. It is possible to sort the applications into several broad categories.

3.1 Domain Specific Software

Particularly in the sciences, virtual organizations have several critical sets of specific software that they share. That software sorts into several categories:

  - Code-sharing. Chemists, astrophysics, economists, etc share a common code set of dedicated functionality, or a set of libraries that service many of their domain needs.
  
  - Data-sharing. Many VOs share large and complex data sets. They may be developed by the VO or imported from external sources, such as Census Data and geospatial sets.
  
  - Distributed computing. Globus (multi-domain specific), Condor and one or two other packages serve a wide variety of VOs by providing a computationally oriented, cohesive computing environment that couples resources across autonomous systems.

  - Instrumentation management and data acquiring. Many VOs are associated with a set of physical instrument sites, such as polar stations, undersea and seismic sensors, satellites, etc.
3.2 Collaborative Tools

A common characteristic across all VOs is the need for a collaboration suite of tools. Among other taxonomies, these tools can be grouped as follows:

- End-user asynchronous tools – Wikis, Flickr, Webdav, RSS, Blogs, etc.
- End-user real-time tools – IM, desktop video, IP telephony, etc.
- VO-aware enterprise tools – calendaring, presence, etc. A common use case entails allowing some VO participants to view certain parts of a colleague’s campus-based calendar.

Some of these are deployed using commercial services and/or proprietary products. Others are open source and usually deployed on a VO or enterprise basis. Many tools are implemented in peer to peer or ad hoc fashion.

It was the sense of the meeting that such collaborative applications are blossoming, and the growth in their use was a testimony to user needs and the individual tools relative ease of use. However, several key elements are missing from the overall environment that together placed a high barrier for scalable deployments. These missing elements include:

- Consistency in their approaches to authentication and authorization.
- Consistent user interface and presentation of common concepts (e.g., groups) across the tools in similar ways.
- Overlapping functionalities. Some tools tightly couple additional functionality that may be better delivered by another tool, and most tools present different feature sets without the ability of the user to evaluate useful features from clutter.
- Non-interoperability among competing standards and competing vendors in key collaboration tools. This is most visible, and seemingly most hopeless, in desktop/room video (between SIP and H.323). The IM space is also difficult.
- Support for service initiation and troubleshooting.

3.3 Integration and Management

Given the number of applications (domain-specific, collaborative, institutional, etc) being used by a typical member of a VO, integration of management and diagnostics across those applications become key issues. These issues become even more considerable when the inter-realm and international issues implicit in virtual organizations are introduced. Complex distributed systems, particularly across autonomous realms, need management tools, both for the end-user and for system-admins; they are simply not there.

Across the applications, across the user’s contexts, several functions were quickly identified as needing integration or common management. These include calendaring, workflow, privilege management, diagnostics, etc. For some of these, implementations may not permit the loose linkages like those typically facilitated by Shibboleth, and so a provisioning tool in particular is important for collaboration so that the accounts can be created flexibly. Event calendaring seems to be a particularly useful instantiation of calendaring functionality for this sort of organization.
4. Operational Issues

a. Enterprise

Considerable thought is required to architect appropriate ways to implement the VO support at the enterprise level, from arrangement of directory infrastructure to higher levels of identity management to accommodate high-security VO needs. Once these approaches are engineered, they could be captured and disseminated to institutions interested in supporting their researchers.

There was extensive discussion of the economics of supporting VOs at the enterprise level. To begin with, many universities do not have recent history of working with the research community. Staffing and funding both seem problematic: there are a number of funding models that have inherent disincentives for well-managed service changes. Even if the VO support is set up on an automated self-service and no staff would be required, there's still a need for machines, networks, infrastructure, connectivity, and troubleshooting. It's much cheaper than staffing, but even that must come from somewhere. If VO support is established by the enterprise on a fee-for-service basis, the grant itself will need to include some provision for the funding of this. There was strong agreement around:

- Cost of labor in such activities is inappropriately discounted. The mentality that these people are already hired and it's a sub-cost can easily set in.
- Cost of policy formation is inappropriately discounted
- The true cost of not doing this is underestimated

Staffing and managing the VO office is a related issue. The breadth of tools and the deep vertical knowledge required to support some of the domain software, suggests that a VO support office itself might be part virtual, coupling the skills of several distributed staff together.

Some concerns were also expressed about the wide variety of demands that the full breadth of virtual organizations could present. One attendee summed it up by expressing how stunned he was by the cheap cost of buying all the boxes, components, and software necessary and the huge amount of money and man hours that go into the negotiation of the policies surrounding this technology. On the other hand, there are strong institutional motives for the investments, from the “lock-in” that the university gets on the activities, to the reduction of exposure that poor VO security represents to the coupling of research and education.

The technology may make it easier for some bad actors to attack this infrastructure, but even if this weren't in place there would still be the need to consider attacks on resources as a matter of course.

b. Virtual organization
Accommodation by VOs of people from institutions with limited resources and expertise is an additional problem. One of the ways to approach the problem is to have these users hosted by an IdP and enterprise directory supported by a third party that is authorized by the federation and service providers to issue assertions on behalf of the appropriate domain. The VO service center, as described above, can also perform this function.

The set of collaborative tools that are used by a VO depends on two significant variables that can change over time:

- Project scale. Choices made at the inception of a project may not scale well as the community grows or changes.
- Available technologies. Choices come to depend on what's deployed by the constituency of the VO.

It was noted several times that the unanticipated use of a tool often turns out to be the most compelling one. This implies the need for flexibility in recommending collaborative tool sets and the value of sandbox-type experimentation by virtual organizations to push new uses.

These tools have to be developed with a degree of simplicity of use that allows them to be used intuitively and easily by ordinary faculty members and students. The people in virtual organizations using these tools are experts in other disciplines, not in technology management. Specifically, there can't be a need to schedule use of facilities, individuals, or a steep learning curve in order to get a tool to work. Those barriers will limit the uptake of these tools: it must be as easy as turning on the lights.

It was noted some virtual organizations might need to be created rapidly in dynamic situations, such as the SARS outbreak or the Tsunami defense situation. In these instances there will be the need to rapidly provision VO services. Other factors, such as extensive training and support, would present major obstacles to rapid VO deployment.

File sharing and a discussion page are basic, and once there have been plenty of open-source tools created for this purpose, it will be modular. Make it very easy.

Part of the problem with the creation of support mechanisms for virtual organizations is that the individuals who encounter problems will generally do so at the resource, and at some point these individuals need to interface with a real support system which is generally located in a traditional organization. Making this integration financially feasible, properly routed, and sufficiently seamless is important.

In addition to the support, ease-of-use and tool flexibility, and startup time of VOs, policies that govern resource access may be different for the VO and for the local enterprise that physically hosts it. Negotiating which one takes precedence could be as easy as the owner sets the rules. However, if the VO is the steward and the resource poses a local security threat, what recourse does the enterprise have? This implies a tight line of
communication among the VO and participating enterprises at the policy and security levels and a method of resolving conflicting priorities.

A follow-on topic entails policy and operational details of the storage of attributes that govern the access of a resource. If Campus A physically houses instrumentation for a VO, but doesn’t own it or control access to it, where do the attributes come from with which it uses to make authorizations decisions? The resource owner at another institution grants access to a new collaborator on a third campus and that is done in the database on Campus A or on the collaborator’s campus or in another datastore that pushes the relevant information to the instrument data store and the collaborators directory?

5. Next Steps

One of the followup items emerging from the workshop was the compilation and lightweight evaluation of collaborative tools by their functionality and their degree of integration with campus infrastructure and other tools.

There was concern about how to motivate the investments, both on the campus and VO levels. The development work and the early deployment will need to be far enough long, and spread widely enough, to motivate significant VOs to shift to a new infrastructure. Ideally, having VO-enterprise integration case studies and member researchers speaking about the benefits will go a long way towards helping their colleagues understand the value.

It appeared reasonable to the attendees that there would be institutions that would be willing to spend significant resources on the bet that supporting virtual organizations by leveraging the enterprise will become cost effective and commonplace, but there needs to be a vision of a model that brings sufficient value for the price in time.

Demonstrations of value in a theoretical or outwardly projected idea will invoke risk avoidance. All these ideas sound like bad ideas and potential threats to the infrastructure of the university today, because professors and staff already have enough to do and we don’t have a concrete, value-add story to offer them instead. Only once we can demonstrate people deeply invested in it, demonstrate the returns, thus invoke the machinery of public relations and appropriate rewards will it bloom at large.

Proceedings

Proceedings of the meeting can be found at http://www.educause.edu/camp053