Identity and Cloud Services: Case Studies from State Networks

Wednesday, Nov. 3, 2010
Internet2 Fall Member Meeting

- George Laskaris, NJEDge.Net
- Leslie Williamson, Merit Network, Inc.
- Mark Johnson, MCNC
- Pankaj Shah, OARnet
- Sujay Daniel, NJEDge.Net
- Timothy Lance, NYSERNet/University at Albany
Merit Network
Services Presentation

November 3, 2010

Leslie Williamson
Director of Services
Merit Network, Inc.
www.merit.edu
About Merit Network

- Merit is a 501(c)3 nonprofit organization founded in 1966 by University of Michigan, Wayne State University and Michigan State University and currently governed by 12 of Michigan’s public universities.

- Owns and operates one of the largest Regional Optical Networks (RON) in North America.

- Merit’s membership consists of public universities, private and community colleges, K12 schools, public libraries, state and local governments and other nonprofits.

- Merit’s members benefit from services that include dedicated network connections, professional learning and forums, BC/DR, and hosted services provided in Merit’s Community Cloud.
Merit Services Portfolio
Merit Services Overview

Is a hosted, “near sourced” email service based on the Zimbra open source platform. MeritMail 2.0 includes:

- A full set of collaboration features including group Calendars, shared Address Books, Documents and Briefcase.
- Ajax Webmail Interface; supports many clients including Outlook, Thunderbird, Eudora, others.
- Mobile Support for iPhones, Android OS phones, Windows Mobile smartphones, Palm Treos and more.
- Cisco IronPort AS/AV protection.
- Data Privacy and Security in Merit Data Centers.
- Unlimited Storage for each account!
MERIT MAIL

300,000+ Accounts

Wayne State University

Eastern Michigan University

Internet

Western Michigan University

Kansas State University

Macomb Community College

Lamar University

Central Michigan University

University of Michigan Museum of Art

ALPENA COMMUNITY COLLEGE

And many others!
Benefits of Near Sourcing

- Secure Merit Data Centers
- Personal 24/7 Support
- Flexible, Agile Implementation and Administration
- Assurance of Privacy/Integrity/Ownership Issues
- Regulatory Compliance
- No Hidden Costs
in a Nutshell

- Enterprise Class Storage Arrays
- Protocols: CIFS/SMB, HTTP/WEBDAV, NFS
- A “Fiber- Network Enhanced” Service; no additional bandwidth fees
- 3 Year Commitment
- Piloting through 2010
- Production in 1<sup>st</sup> Quarter 2011
Security – Two layers

- Member enforces security policies on their file server(s)/applications server(s) for applications and users

- Merit OpenFiler-to-Member connection protected by password and IP address

- Member has separate storage space on Merit’s storage array that is not shared with anyone else.
Piloting and In Development

• Virtual Machines
• Web Conferencing – ATT Connect as platform, no individual subscriptions
• MeritMail 3.0 – With web conferencing and DR options

Under Review
• Media Repository
Merit Network
Services
Presentation

Leslie Williamson
Director of Services
Services-Info@merit.edu
Placeholder Mark Johnson, MCNC
Cloud Computing and Identity Management

Nov. 3, 2010
Internet2 Fall Member Meeting

Pankaj Shah, Executive Director, OARnet
Shared Infrastructure (SI) Objectives

• Shared Infrastructure
  – Implementation of a hardware platform that will advance the development of an integrated technology infrastructure as outlined in 10-year Strategic Plan for Higher Education

• SI implementation will allow Ohio to lead the country in cloud computing on a statewide level as a higher education system
Shared Infrastructure (SI) Objectives

- Primary drivers for cloud computing
  - Consolidate services
  - Reduce costs
  - Provide better SLAs for compute services
  - Reduce duplication
  - Take advantage of economies of scale
  - Provide disaster recovery and business continuity to all state universities
  - Provide full array of hosting solutions
Shared Infrastructure Capacity

Capacity measured in four key areas:

1. Computing
   • Performance and utility blades
2. Storage
   • Tier 1 and Tier 2
3. Backup
4. Database
Compute Capacity – Performance Blades

Capacity of physical blades

- 14 of 22 blade servers in use
  - 9% Capacity expansion to VM cluster
    - 2 blades
  - 27% Available capacity (on order)
    - 6 blades
  - 64% Total compute capacity in use
    - 14 blades
    - 5 virtualized
    - 9 stand alone

Capacity of 280 virtualized machines based on 7 blades

- 1 blade, virtualized, equals 40 VM
- 120 of 280 virtual machines in use
  - 28.5% Expansion of 2 blades
    - 80 virtual machines
  - 28.5% Available capacity (from 2 of the blades on order)
    - 80 virtual machines
  - 43% Virtual capacity in use
    - 120 virtual machines
Compute Capacity – Utility Blades

- Of 11 utility blades, 6 available for future growth

- Total compute capacity in use: 5 blades (45.50%)

- Capacity expansion (on order): 4 blades (36.30%)

- Compute capacity currently available: 2 blades (18.20%)
Virtual Storage Capacity, October 2010

**Tier 1 Storage**
Provides Performance, 44 TB

- 40% 18 TB available
- 60% 26 TB in use

**Tier 2 Storage**
Provides Capacity, 300 TB

- 37% 110 TB available
- 63% 190 TB in use
Backup Capacity, October 2010

- Tivoli Storage Manager (TSM)
  - 700 TB total useable backup purchased
  - 42% in use (300TB of 700TB)
Database Capacity, October 2010

- Oracle database
  - Used for SI application
  - 3 node production cluster
  - 2 node test cluster
  - 25% of production capacity
Projected Capacity of the SI

100% utilization and full capacity projections, in months

- **VM cluster**: rate of growth = 25 machines/month. 100% utilization by Feb. 2011
- **Storage, Tier 1**: rate of growth = 2 TB/month. Full capacity by Aug. 2011
- **Storage, Tier 2**: rate of growth = 15 TB/month. Full capacity by Jan. 2011

Timeline from Aug. 2010 to 8 months.
USO’s Shared Infrastructure

- Active applications include:
  - DRC development site
  - OARnet webservice
  - Regents development webservice
  - OhioLINK NFS server
  - OhioLINK NAXOS
  - Tivoli storage manager
  - Oracle
  - Nagios monitoring
  - VMware ESX cluster
  - Shibboleth
  - ATC development server

Kinnear Road Center
Future SI Activities

- A sample of planned projects include:
  - eFolio
  - Student Portal Gateway
  - Ohio Tuition Trust Authority
  - Higher Education Information system
    - BOR-SAS – Two-year project to replace the HEI system
  - eTech Conference Tracking
  - OhioLINK QuickSearch2
  - Identity management project
  - Microsoft Exchange e-mail
  - OhioLINK Electric Journal Center (EJC) migration to Linux
Lessons Learned

- Current apps migrate easily.
- Legacy apps have a fairly steep entry curve.
  - Must first be ported
- It takes time to figure out governance.
- It takes time to develop relationships that allow organizations to take full advantage of the environment.
Federated Identity Management System Project

- Will support the University System of Ohio, which includes:
  - Ohio’s 14 public universities, 24 branch campuses and 23 community colleges
  - More than 500,000 students

- Will ultimately optimize resources for Ohio’s higher education by:
  - Strengthening articulation and transfer within the system
  - Growing its online offerings and student services
  - Saving significant administrative costs through institutional collaboration
  - Strengthening workforce development, research and technology transfer
Overall Approach & Drivers

- The Chancellor’s *Strategic Plan for Higher Education* calls for creation of a single, integrated technology infrastructure, including a federated identity management system of authentication.

- With federated identity management, students will be able to:
  - Apply for admission and register for courses at multiple University System of Ohio institutions and campuses.
  - Access and easily use online system to research courses at different schools, enroll and transfer credits, and complete necessary financial transactions.
  - Allow K-12 students and teachers to access the University System of Ohio online resources and services by using the same mechanisms they use to authenticate at their home school.
Next Steps – Federated Identity Management System

OARnet was tasked with developing implementation plan

Step 1
Created a subcommittee under OARnet ‘s CIO Advisory Board to ensure input regarding FIDMS role among Ohio educational institutions

Step 2
Secured commitment funding from OARnet members to create an FIDMS Project requirement and implementation plan

Step 3
Review InCommon as the standard for our Federated Identity Management System. Determine whether or not to recommend membership in InCommon to all Ohio Schools

Step 4
Hire a SME dedicated to the project

Step 5
Review implementation plan with the OARnet Governing Board and CIO Advisory Board for approval and funding
OARnet
1224 Kinnear Road
Columbus, Ohio 43212
614-292-9191
www.oar.net
pshah@oar.net

Special thanks to:
Paul Schopis, CTO
Denis Walsh, Chief Relationship Officer
James Jacob, Director of Systems
Susan Mantey, Outreach Manager
Cloud Services and Federated Access Management

Sujay Daniel
NJEDge.Net
New Jersey Research & Education Network
Cloud services

• NJVID – Digital Video Repository and streaming services
• Remote Storage Services (DR)
• Shibboleth Identity Provider VMs
• Vidyo portal

Future:
• Statewide Digital Object repository.
• Wireless across state campuses for visiting students
Thank you

Sujay Daniel
NJEdge.Net
New Jersey Research & Education Network
sujay@njedge.net
NJTrust – now and the future

- Trust federation initially created for NJVID
- LDAP for smaller schools without Enterprise Directories
- Plans to add K12, Libraries and govt. entities
- In the future we may include entities outside New Jersey
- Inter-federate with InCommon
NJTrust Participants

Identity Providers
- Berkeley College
- County College of Morris
- Fairleigh Dickinson University
- Middlesex County College
- Montclair State University
- New Jersey Institute of Technology
- Ramapo College of New Jersey
- Rowan University
- Rutgers, the State University of New Jersey
- Stevens Institute of Technology
- The College of New Jersey
- William Paterson University

Service Providers
NJTrust Participation

• Run your own Shibboleth Identity Provider (IdP)
  ▫ Recommended

• Participate in the Statewide Directory System
  Institutions without a directory server
  Administrators can manage their Institutional User data using web front-end
  For smaller organizations – K12s, Museums, Public Libraries

• NJEDge hosts Shibboleth IdP instances (VMs)

• How to handle Public Libraries with Patron records?
Cloud Supercomputing

The challenges of distributed supercomputing

Tim Lance
NYSERNet
In the beginning was the research

- With industry (especially IBM) and NYSTAR (NY office supporting S&T), convened two meetings about tackling grand challenge problems
- Focused on Energy and Environment, Health Care, Government, Financial, Nano Materials and Electronics
- Asked how to balance and coordinate three orthogonal sets of requirements
In the beginning was the research
Conclusions of these meetings:

- Absolute requirement to cooperate across sectors: academe, industry, government
- As a component, finding ways to share expensive technology such as computing and instrumentation
- Critical need for “bridging” people: experts in the technology to be shared who can bring the researchers to their productive use
- Creation of HPC\(^2\) consortium
Building on the capabilities of previous investments at three New York computational centers, world class networking, and leveraging some of the best intellectual capital in the world, New York State is a leader in high performance computation and is poised to transition those capabilities to a wide variety of industries throughout the state.
To realize this potential, ongoing support is needed for researchers and technologists as they move to the more complex environment of massively parallel computation. This will include specific efforts to foster collaboration between industry and university researchers.
A key challenge is solving the “last mile of computing problem” - developing portals to enable simulation-based science and engineering.

Tom Furlani, PhD
Center for Computational Research
University at Buffalo, SUNY
"Old School" Computing

```
[bono:~]$ mkdir d_game$$
[bono:~]$ cd d_game$$
[bono:~/d_game$$]$ sftp myworkstation.buffalo.edu
[bono:~/d_game$$]$ is thymine.inp
[bono:~/d_game$$]$ vi thymine.inp
```

**Secure file transfer**

**Upload input data file**

**Input File**

**Secure login to front-end machine**

**Create subdirectory**

**Upload input data file**

**Add keywords to input file**

**Secure file transfer**

**Identify keywords for model**

**Edit input file**

**Create PBS script file**

**Edit file**

**Set number of processors**

**PBS format and syntax**

**Set path and variables**

**Submit job to queue**

**Set run time and queue**

**PBS commands**

**Monitor job**

**Submit job to queue**

**Application command line**

**Monitor and queue**

**PBS format and syntax**

*This job was run on a R16/600 model 450 (Oct 96)*

```bash
$cat thymine.inp
   run command
   $system mem=448000 tlimit=99999.0
   $basis g Basis 4 3
   $guess g
   $end
   $dir=FALSE, $end
   $dir=TRUE, $end
   $data
   $end
   thymine...one of the bases found in nucleic acids...RF/6-31G(d,p)/RF/6-31G(d,p)
```

```
1. 1.471469
2. 1.3706594 1.17.0271076
3. 1.3069220 1.12.7937358 1.32.007350 2.31.007350 2
4. 1.4576970 1.12.0647771 2.31.007350 2
5. 1.3958360 1.90.7794685 2.35.734770
6. 1.2708041 1.11.4849403 1.37.6723857
7. 1.3783592 1.11.6919987 3.37.6723857
8. 1.639509 1.12.5281856 4.37.6723857
9. 1.196342 1.11.7915522 1.37.6723857
10. 1.187159 1.11.9788448 1.37.6723857
11. 1.159735 1.11.6237097 1.37.6723857
12. 1.187159 1.11.7915522 1.37.6723857
```

```
1. 1.471469
2. 1.3706594 1.17.0271076
3. 1.3069220 1.12.7937358 1.32.007350 2.31.007350 2
4. 1.4576970 1.12.0647771 2.31.007350 2
5. 1.3958360 1.90.7794685 2.35.734770
6. 1.2708041 1.11.4849403 1.37.6723857
7. 1.3783592 1.11.6919987 3.37.6723857
8. 1.639509 1.12.5281856 4.37.6723857
9. 1.196342 1.11.7915522 1.37.6723857
10. 1.187159 1.11.9788448 1.37.6723857
11. 1.159735 1.11.6237097 1.37.6723857
12. 1.187159 1.11.7915522 1.37.6723857
```
Goal: Portal Driven Computing

- Secure login to web portal
- Upload input data file
- Select model and run job
- Monitor job
- View Output

WebMO Login
Version: 8.0.011e
Center For Computational Research

Input File

Configure Gamess Job Options

Open Browser Select Model Monitor Jobs

View Output in Browser
Catalyst for Collaboration

HPC Portal

Improved Simulation Through Collaboration

High Performance Computing Consortium Web Portal

Getting Started
Computing Software
Documentation and Training
Research
Lessons learned (so far)

- Long before the portal, very people intensive
- Depended on years of trust from networking efforts
- Start with low hanging fruit
- That fruit isn’t all that low
- Industry is a key partner
- Identity management still needs to be built in
Thank You

NYSERNet, Inc.
100 S. Salina St., Suite 300
Syracuse, NY 13202
315-413-0345

NYSERNet, Inc.
385 Jordan Rd.
Troy, New York 12180
518-283-3584