June 9th 2010, US ATLAS Tier-3 Meeting

Jason Zurawski - Network Software Engineer & Research Liaison

Dale Finkelson – Sr. Program and Service Manager

Internet2 Overview

Introduction

- Internet2 is an advanced networking consortium led by members of the Research and Education (R&E) community.

- We promote the missions of our members, in part through the development and support of networking activities and related initiatives.

- We are committed to supporting scientific use of the network, including USATLAS.

  – Enabling large scale data transfers over a high capacity nationwide network

  – Dynamic circuit capability through the ION service

  – Performance Monitoring through perfSONAR

  – Support for the debugging of Network Performance, end to end.

INTERNET2

Outline

- R&E Networking
  - Explanation
  - Campus Connectivity
  - Benefits
- Information about Internet2
  - History
  - Network
  - Services
- Scientific Outreach
  - Mission
  - Communities
  - Challenges
- USATLAS Performance Monitoring
- Additional Information

INTERNET2

R&E Networking

- Many Schools/Research Centers maintain two types of internet connectivity:
  - Commodity (Commercial) networking
    - Connections purchased through a commercial venture (e.g. similar to a cable modem for home use)
    - Connects to the 'public' internet: commercial web sites, etc.
    - Typically support less raw connectivity speed due to cost assocaited
  - R&E Connectivity
    - Purchased through a non-profit national organization (Internet2, NLR), a regional network (CENIC, MAX, etc.) or a governmental agency (ESnet)
    - Purpose built to support R&E activities, next generation hardware and software services
    - Traditionally less traffic than the public internet
    - Higher connectivity speeds possible (1G, 10G, and above)

R&E Networking

- Typical Camps Network Scenario (Researcher's POV):
  - Data Center/Lab
    - Storage/Cluster equipment lives here.
    - Connectivity to the central part of campus may vary:
      - Copper or Fiber links.
      - 1G or 10G campus connectivity is becoming common. May be much less though…
  - Campus Core
    - Central campus networking/computing facility, also may vary between 1G and 10G
    - Manages the Wide Area (WAN) connections through the 'core', e.g. where the campus intersects with the rest of the world.
  - Regional Network
    - Provide WAN connectivity to several Campus/Research centers at different speeds.
    - Through here, each network will directly connect to Backbones/National networks. Occasionally will peer with International exchange points.
  - Backbone/National Network
    - ESnet, Internet2, NLR, etc.
    - Multiple Points of Presence (PoPs) on several redundant long haul segments
    - Many regional's may connect through a PoP, and PoPs may have access to an International exchange point.
  - International Peer (if Applicable)
    - Connections that span countries, continents, oceans.

INTERNET2

R&E Networking

- Value add from R&E networking
  - Access to measurement and monitoring data
    - All measurements (Netflow, SNMP, throughput, latency) are available or may be requested
    - Greatly simplifies requests to assist with performance problems
  - Latest equipment to keep up with technology
  - Ability to experiment with new technologies and protocols
    - Dynamic circuits, transport protocols beyond TCP
  - Connectivity to institutions of Interest
    - Other research institutions, national and international
    - Governmental agencies
  - Support
    - 24hr Network Operations Center (NOC) for technical issues
    - Research and Community Outreach
    - Application and Network Research support

INTERNET2

Information about Internet2

- Internet2 is an advanced networking consortium.
- Formed and led by the R&E community since 1996
- Promote the missions of our members by providing both network capabilities and services using the latest technologies
- Current Membership profile:
  - 214 Higher Education Members
  - 10 Industry Partners
  - 3 Industry Sponsors
  - 27 Industry Members
  - 59 Affiliate Members
  - 33 Regional Network Members
  - 56 International Partners

INTERNET2

Information about Internet2

- Internet2 Network and PoPs:

Information about Internet2

- Internet2 Products and Services
  - IP Network Connectivity
    - Built across a carrier-class infrastructure and supports major and leading edge networking protocols
    - Ability to increase backbone capacity as required
  - ION
    - Virtual Circuit network service that provides dedicated bandwidth for the most demanding applications.
  - perfSONAR
    - Infrastructure for network performance monitoring, making it easier to solve end-to-end performance problems on paths crossing several networks.
  - Incommon
    - Common framework for trustworthy shared management of access to on-line resources in support of education and research in the United States.
    - Built using software from the Shibboleth project.

Scientific Outreach

- Research use of the network is pivotal to our mission to support our members

- Scientific outreach involves interacting and understanding network requirements from major scientific projects and Virtual Organizations (VOs)

- VO examples:
  - LHC (USCMS, USATLAS)
  - LSST (Astronomy)
  - NEES (Structural Engineering)
  - eVLBI (Astronomy)
  - NEON (Climate Research)
  - OOI/NEPTUNE (Oceanography)

INTERNET2

Scientific Outreach

- Typical requirements for science data vs other traffic:
  - Flow size is typically larger to mimic data sets that come off of an instrument.
  - Flow duration may grow as data size increases. May be uninterruptible (e.g. file structure is small number of large files vs large number of small files)
  - Destinations and Sources originate at research oriented facilities, normally with available R&E connectivity
- Technological Challenges
  - Scientific Instrument/Data Origin
  - Campus
  - Regional
  - Backbone

INTERNET 2

Scientific Outreach

- Scientific Instrument/Data Origin
  - Remote instruments (telescopes, colliders) may have dedicated local facilities to store experimental data
  - WAN connections to the major local and national networks
- Campus
  - Campus design favors the population of a campus, e.g. 1000s of laptops running windows. As such design favors 'protection' in the form of firewalls, limiters, and smaller local capacity.
  - Perhaps some thought to research, e.g. connections to data centers may be offered special treatment in the core.
- Regional
  - High capacity uplinks and the ability to engineer traffic
- Backbone
  - High capacity long haul links
  - Peerings to other networks and facilities

INTERNET2

Scientific Outreach

- Internet2 Performance Initiative has the following goals:
  - Create a predictable, and well-supported networking environment
  - Network users have routinely successful experiences in their development and use applications
- Focusing resources and efforts on improving performance problem detection and resolution throughout campus, regional, and national networking infrastructures.
- perfSONAR
  - Infrastructure for network performance monitoring
  - Make it easier to solve end-to-end performance problems on paths crossing several networks
  - Joint effort between ESnet, GÉANT, Internet2 and RNP
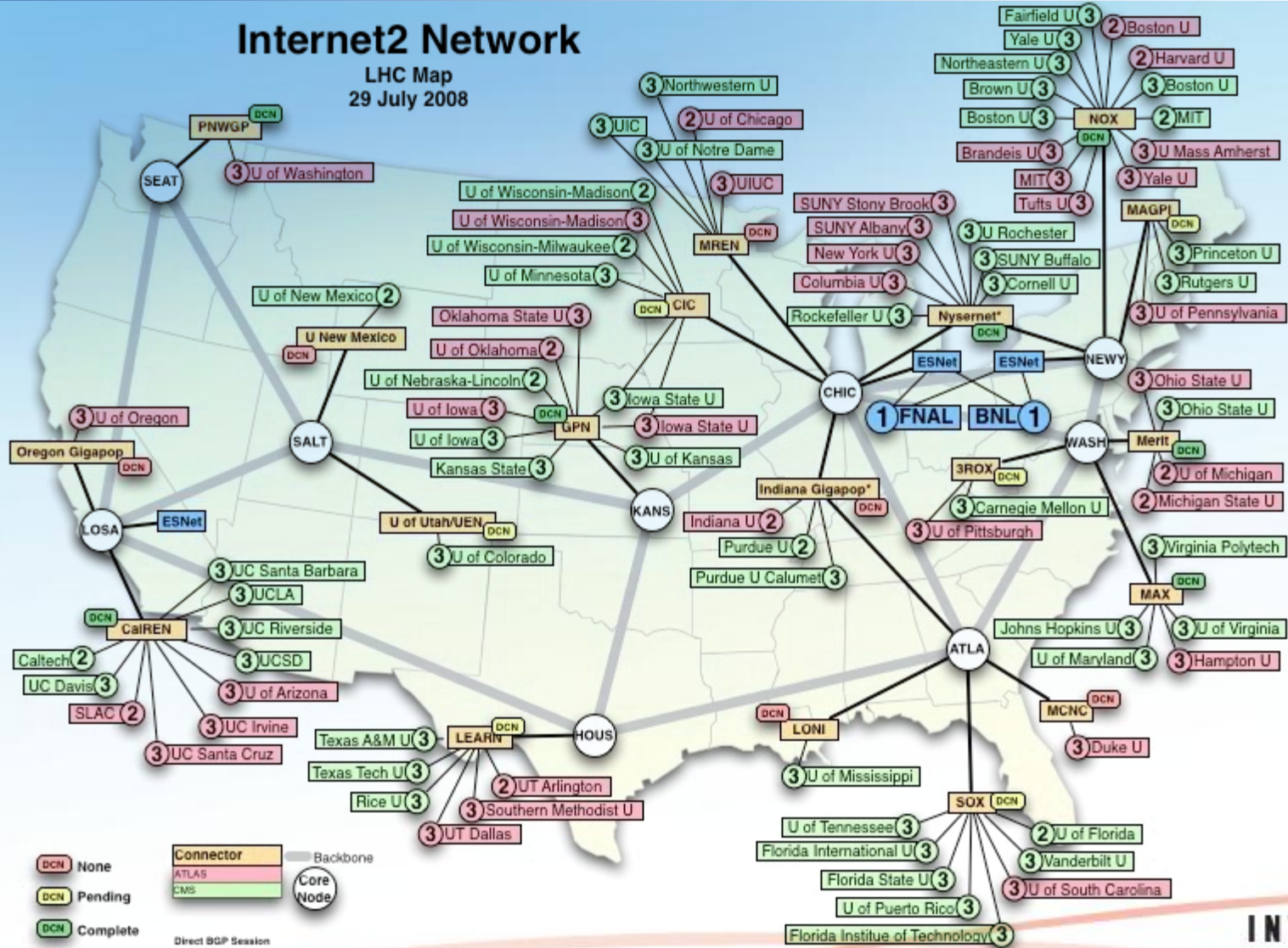 - http://www.perfsonar.net

INTERNET2

Scientific Outreach

- Provisions for evaluating/addressing network performance
  - Encourage use of instrumentation and tools
    - The pS Performance Toolkit (http://psps.perfsonar.net/toolkit/)
    - Two modes of use:
      - Regular Monitoring using performance tools
      - Diagnostic mode – use on demand when something seems 'wrong'
    - Used at all USATLAS Tier2s and the Tier1
    - Suggested use at Tier3s (Rik, Doug, Shawn and Jason working on drafting a recommendation of hardware and software use).
  - Help in debugging
    - Identification, tracking, and resolution of performance problems, end to end if necessary
    - Internet2 staff can help the correct people get together, and encourage the use of systematic debugging using the proper tools
  - Help in altering network design when applicable
    - Working with campus and regional staff to suggest changes
    - Use of the ION service for added capacity as needed

Internet2 LHC Project Connectivity

USATLAS Performance Monitoring

- 2007 – USATLAS decided as a group to evaluate 2nd generation perfSONAR CD (e.g. NPToolkit) as a testing and monitoring framework

- Each Tier2 facility and the Tier1 Purchased 2 servers
  - "***Koi Computers***" – 1U Chasis
  - Dual Core Pentium 2.2GHz Processor
  - 2GB Ram
  - 160GB Hard Drive
  - ~$1200 for both

INTERNET 2

USATLAS Performance Monitoring

- 2010 – All sites still on original hardware, running 3rd generation (3.1.3) of the pS Performance Toolkit
- Testing
  - BWCTL
    - Test in a "*full mesh*" to all Tier2s and the Tier1
    - 20 Second Throughput Tests, once every 4 Hours
    - May adjust schedule based on how much of traffic is observed to be measurements
  - OWAMP
    - Test in a "*full mesh*" to all Tier2s and the Tier1
    - Continuous stream of 1 minute OWAMP tests (10 packets per second – 600 total per minute).
    - Determine min/max latency, loss, and "*jitter*" (delay variation)
  - PingER
    - Not mandatory – but should test to "full mesh" of Tier2s and to the Tier1

INTERNET 2

USATLAS Performance Monitoring

- Machine Allocation
  - 2 Per site
  - Placed near other Tier2 equipment (e.g. temperature controlled and close to the WAN gear)
  - Latency Testing Host
    - OWAMP Tests
    - PingER Tests
    - SNMP Monitoring
  - Bandwidth Testing Host
    - BWCTL Tests
    - NDT Diagnostics
    - NPAD Diagnostics

INTERNET2

USATLAS Performance Monitoring

- Procedures
  - Takes about 2 weeks to upgrade when there is a new ISO
  - 1 – 2 Weeks to establish the initial testing parameters and set up regular tests
    - Set up boxes first so they can be "*found*" by the perfSONAR tools
    - Set up the testing meshes (each site tests to all others).
  - Bi-weekly calls (Shawn's 'Throughput' Group - most times with an Internet2/ESnet engineer) to evaluate the performance they are seeing and request any enhancements + report bugs regarding the ISO
  - Each site will coordinate with others (and if necessary an Internet2/ESnet Engineer) to debug perceived problems

INTERNET 2

USATLAS Performance Monitoring

- Uses
  - Regular BWCTL/OWAMP data is viewed daily by site and USAtlas admins for abnormalities
  - Used in conjunction with GridFTP data and other forms of throughput testing
  - Diagnostic tools (NPAD/NDT) are used by Tier2 and Tier3 participants to diagnose problems from end site to USAtlas data repositories

INTERNET2

USATLAS Performance Monitoring

- Problems Found
  - Throughput problem between Tier1 and Tier2
    - Couldn't exceed 1 Gbps across a 10GE end to end path that included 5 administrative domains
    - Used perfSONAR tools to localize problem
    - Identified problem device
      - An unrelated domain had leaked a full routing table to the router for a short time causing FIB corruption. The routing problem was fixed, but router started process switching some flows after that.
    - Fixed it
      - Rebooting device fixed the symptoms of the problem
      - Better BGP filters on that peer will prevent reoccurrence (of 1 cause of this particular class of soft faults)
  - Loss events inbound to a particular Tier2
    - Gave a quick reason to longstanding bandwidth problem
    - Corrected quickly once there was proof of loss

INTERNET 2

USATLAS Performance Monitoring

- This Audience (Tier3s) will provide new challenges:
  - Measurement scalability – a full mesh between Tier2s and Tier3s is impossible
    - Measurement traffic would dominate real network traffic
    - Considering smaller test sets, or on demand testing (opposed to scheduled measurements)
  - Hardware is being revaluated
    - Dell hardware in testing
    - Single box solution (vs 2 KOIs) is being considered
  - Cost must be justifiable for all participants
  - Audience – Tier3 centers may have limited resources to monitor the monitoring. Needs to be 'set it and forget it', and allow for alarms when things are not correct.

INTERNET 2

Additional Information

- Additional Information
  - USATLAS Liaison
    - Jason Zurawski (zurawski@internet2.edu)
  - Internet2 Membership questions
    - membership@internet2.edu
  - Internet2 Network
    - network@internet2.edu
  - Internet2 NOC (report network related problems, outages)
    - noc@net.internet2.edu

INTERNET2

Internet2 Overview

June 9th 2010, US ATLAS Tier-3 Meeting
Jason Zurawski - Network Software Engineer, Research Liason
Dale Finkelson - Sr Program and Service Manager

For more information, visit http://www.internet2.edu