
June 10th 2010, Transatlantic Networking for LHC Experiments

Eric Boyd, Internet2 Deputy CTO

# An Overview of Architectural Directions and Advanced Services

# Outline

- Internet2 Architectural Directions
  - Hardware
  - Bandwidth
  - Services
- Internet2 Network and Advanced Services Update
- ARRA and Stimulus Update
- Scientific Involvement and Outreach
  - USCMS
  - USATLAS
- Performance Software Development

# 2010 Internet2 Architectural Directions

- As Internet2 discussed with the community about potential responses to ARRA funding, Internet2 staff in collaboration with the Indiana University GRNOC prepared a set of architectural statements that seek to clarify the Internet2 architectural direction as it relates to stimulus funding.
- The contents of this document are meant to be broad in scope and are not indicative of a committed set of directions.
- They are intended to spur community discussion of this vision.
- Many of the core principles are taken directly from recommendations collected during the ARRA outreach calls in summer of 2009.

# Architectural Principles

- **Multiples of 10GigE will be the primary transport to Regional and State Networks over the next 3-5 years.**
  - 10G cost low compared to the cost of 40 or 100G
  - Multiple large sub-10G flows the norm
- **Internet2 Network access will be divorced from physical interface speeds and available for apportionment across the network**
  - Flexibility for connectors an important success factor.

# Architectural Principles

- **Native 100GigE at the optical layer is an important technology to adopt today**
  - Take advantage of current opportunities to lay the groundwork for future expansion.
- **Collapsing Layer2 and Layer3 services onto a single delivery platform is an important step toward the hybridization of the network**
  - Reduce overall operating expenses to the Connectors
  - Candidate technologies include MPLS L2 VPNs, Layer2 Ethernet VLANs and Virtual Private LAN Service (VPLS).

# Architectural Principles

- **The Internet2 IP and Layer2 Networks need a migration path to 40G and 100G in the next few years**
  - Backbone must be able to efficiently handle multiple simultaneous 7-10 G flows and individual flows >10Gbps
- **The Internet2 Network emphasis should be on additional services and technologies that will drive transport bandwidth requirements**
  - The use case for the network drives the technology of the network.

# Architectural Principles

- **Internet2 will coordinate with the Regional and State Network partners to determine the most optimal node quantity and locations**
  - Offer a flexible partnership with the connectors.
  - Create more options for connections.
- **As mission-critical applications become more integral to the Regional cost-recovery model, the Internet2 Network must focus on enhanced redundancy where needed**
  - Many recent services and uses of the network require increasingly reliable/redundant/resilient connectivity

# Architectural Principles

- **The Internet2 Network will continue to be instrumented and operated in a transparent fashion that supports the end-to-end model**
  - The more information that is available about the network the better everyone understands the need for and requirements of the network.

# Outline

- Internet2 Architectural Directions
  - Hardware
  - Bandwidth
  - Services
- Internet2 Network and Advanced Services Update
- ARRA and Stimulus Update
- Scientific Involvement and Outreach
  - USCMS
  - USATLAS
- Performance Software Development

# Internet2 Network Upgrades

- Backbone capacity growth
  - Entire backbone now at 20G.
  - Selectively going to 30G on busy segments summer 2010

- Juniper MX960 Transition
  - All T640's upgraded to MX960's in Atlanta, Washington, New York, Kansas City, Houston, Los Angeles, Salt Lake City and Seattle
  - Chicago will remain a T1600

# Evolving the Internet2 ION Service

- Dynamic circuit network services called Internet2 ION launched as a production service in Fall 2009. Next step is hybridization:
  - ION and IP service on the same platform brings new use cases and development opportunities
  - ION/IP services on a Universal Hybrid Port(s) provide easier interconnect and regional flexibility
  - More efficient leverage of member fees toward growing both services at the same time
- NTAC concluded that ION services over MPLS tunnels could co-exist on the same platform as the IP Network
- AOAC endorsed plans to migrate ION to the IP Network
  - ION services provided by MPLS tunnels on the MX960
  - Transition timeline being worked out
  - Goal is complete transition by July 31st, 2010

# New Connector Fee Options

- Fee options developed in coordination with and community input and has been Board reviewed
- Benefits:
  - provides a pathway for true hybrid network capabilities
  - substantially lower cost per-bit
  - more headroom for research
  - more diversity
- New options include:
  - 2x10G $525/yr
  - 2x5G $375/yr
- More than half the connectors have already upgraded to dual ports or intend to in 2010

# Internet2 Commodity Peering Service (CPS)

- Mid January Traffic Snapshot (based on peak 2 hr averages)
  - 30Gbps toward Connectors
  - 9.5Gbps from Connectors
- TransitRail Peering
  - Internet2 CPS and TransitRail now peering in four locations: Los Angeles, Seattle, Chicago, Washington DC
  - Each network sending a subset of its *peer* routes to the other
  - Each network sending the other network's *customer* routes to a subset of its peers
  - Majority TR peers are residential broadband networks
  - Full consolidation of services proceeding to move commodity on to a dedicated backbone wave with routers at exchange points
- Over 157,000 Commodity Peer routes
  - Mostly high-value content providers
  - Amazon S3 and EC2 service now reachable via CPS

# Open Network Management Initiative

- Understand how network links are being used and performing.
- The initiative aims to:
  - Use the perfSONAR infrastructure to structure the data
  - Develop an open set of tools that address specific performance questions and issues
  - Develop initial use cases and prototypes
  - Develop an Open Network Management platform incorporating tools
  - Provide assistance in the use of these tools and understanding their results
- Community input and partnerships needed to help further this development
- Contact Dale Finkelson (dmf@internt2.edu) for more information

# Cisco Telepresence

- Internet2 and NLR have signed an MOU to present a unified telepresence service offering for the US R&E community. The goal is to leverage the strengths of each organization to make Cisco TelePresence more widely available.
- Plans include:
  - Work jointly with on a Cisco TelePresence strategy
  - Jointly engage commercial TelePresence providers
  - Initially focused on Cisco TelePresence, but intended to be interoperable with other forms of videoconferencing
- Many details remain to be worked out, but far enough along to permit demonstration / proof of concept connections.
- For input or questions, please email: cisco-telepresence-request@internet2.edu or cisco-telepresence-request@nlr.net
- Mailing list for discussion, news, etc: cisco-telepresence-info@internet2.edu or cisco-telepresence-info@nlr.net

# Internet2 DNSSEC Deployment

- The US Department of Homeland Security (DHS) funded Shinkuro (shinkuro.com) to promote pilots for DNSSEC in the U,S.
- The Internet2 DNSSEC advisory group was formed in cooperation with Shinkuro, focused on the R&E community. Group supports:
  - Monthly conference calls held to support those who want to deploy
    - Participation on the calls from representatives from Educause, NIST, and NASA.
  - Biannual BoFs at member meetings
  - ~100 subscribers on the email list
- Participation in Educause beta test effort for .edu signing
- internet2.edu & ucaid.edu zones signed in production
- Deployed validating resolvers to be deployed internally (this week)

# Outline

- Internet2 Architectural Directions
  - Hardware
  - Bandwidth
  - Services
- Internet2 Network and Advanced Services Update
- ARRA and Stimulus Update
- Scientific Involvement and Outreach
  - USCMS
  - USATLAS
- Performance Software Development

# Recent Grant Activity

- Internet2 is acting as project lead for CI Days (sponsored by NSF) in conjunction with CI Consortium (PI: Russ Hobby)
- Internet2 is acting as project lead for perfSONAR Workshop (sponsored by NSF, DOE, and NITRD)
- Internet2 is involved in several GENI-related efforts, including a project to deploy OpenFlow on the network backbone
- Internet2 is acting as project lead to develop perfSONAR package for IRNC links (sponsored by NSF)
- Internet2 is acting as project lead to develop OSCARS/DRAGON package for some IRNC links (sponsored by NSF)
- Other proposals have been recently submitted or may be close to getting funded

# perfSONAR Workshop

- First Workshop on the perfSONAR Network Measurement Infrastructure
  - Bring together researchers, applications developers, network operators, network managers, and others with an interest in network research and network performance monitoring and measurement.
- July 7th, 8th, 9th 2010 – Arlington VA (Crystal Gateway Marriott)
- Sponsored By:
  - **Networking and Information Technology Research and Development Program**
    Large Scale Networking (LSN) Coordinating Group
  - **National Science Foundation**
    Directorate for Computer and Information Science and Engineering
  - **Department of Energy - Office of Science**
    Advanced Scientific Computing Research
- http://www.internet2.edu/workshops/perfSONAR/

# Grant Proposal - IRIS

- Objective: provide the infrastructure necessary for the identification, diagnosis and eventual correction of network performance problems for paths traversing IRNC links
- IRIS will provide a software framework to enable performance monitoring services on IRNC links
  - Produce a set of easy-to-install, tailored software packages of the perfSONAR-PS software suite for use by the IRNC link operators
  - Develop new functionality specific to international exchange points
  - Work with the IRNC link awardees to help them deploy the software

# Grant Proposal - DyGIR

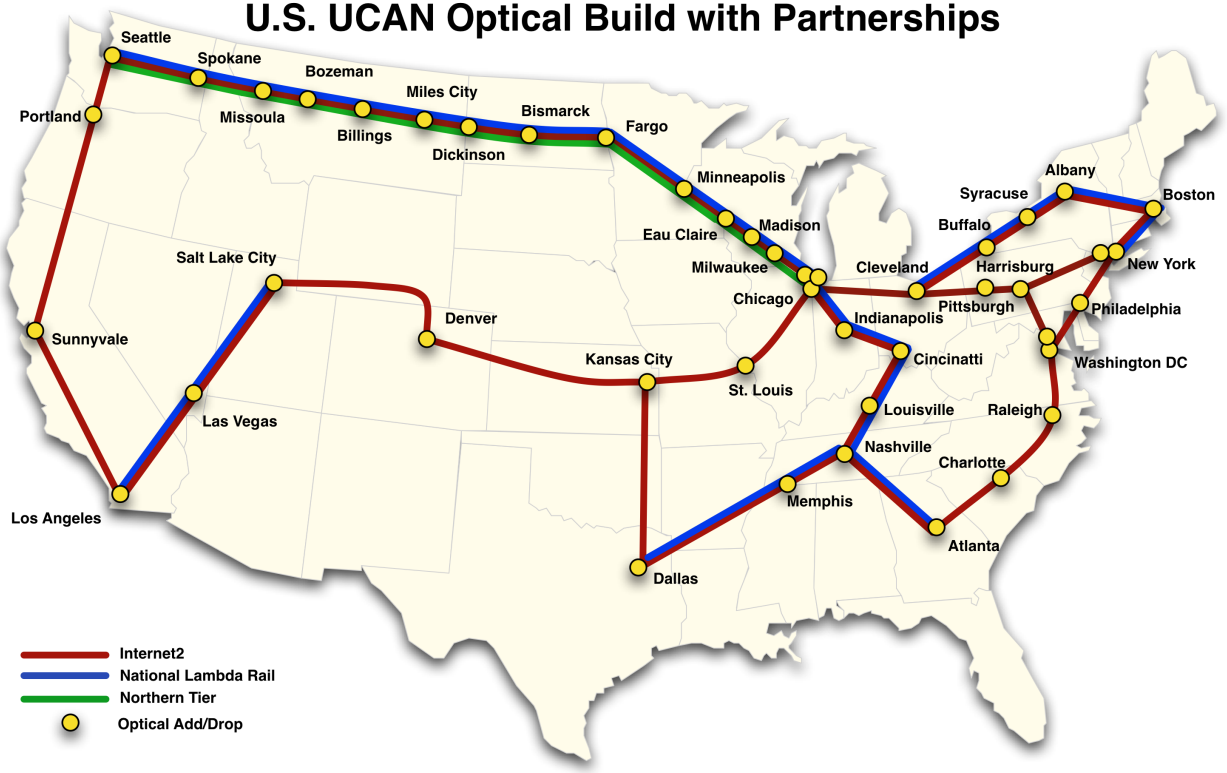
- Objective: enable researchers to reserve dedicated bandwidth over IRNC links
  - Allow for more distributed collaboration
- DyGIR will provide a prototype software framework to enable dynamic circuit services for a pair of IRNC links
  - Produce a set of easy-to-install, tailored software packages for the OSCARS software suite
  - Prototype new functionality specific to international exchange points
  - Work with two of the IRNC link awardees to help them deploy the prototype software

# Internet2 and ARRA Stimulus

- Many connectors are looking to expand through NTIA BTOP program.
- Internet2 is exploring ways to upgrade/expand capabilities to match up with the expected growth of regionals and ensure fees remain the same or potentially reduced.
- Internet2 has submitted a Round 2 Proposal to the ARRA-funded Broadband Technologies Opportunities Program (BTOP) as funded by the NTIA
  - Seeks to acquire nationwide dark fiber, optical equipment to light the fiber at 100G speeds, and an upgraded IP network delivering 100GigE to the Internet2 Community

# New Network Builds in Proposal

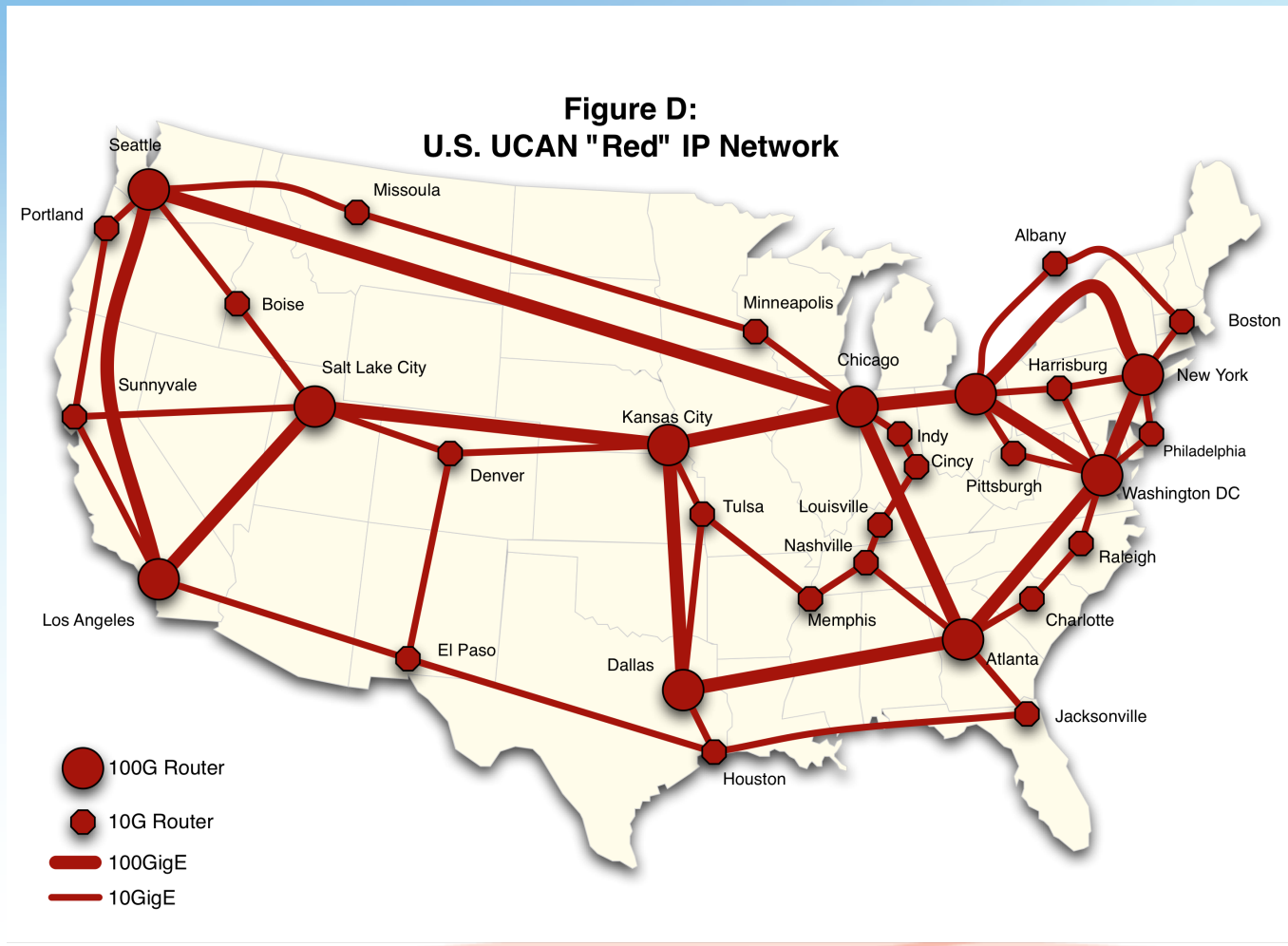
**Figure C:**
**U.S. UCAN Optical Build with Partnerships**



# Combined US UCAN System Capability

**Figure B:**
**U.S. UCAN Combined Optical System Capabilities**

# Upgraded IP Backbone



# Outline

- Internet2 Architectural Directions
  - Hardware
  - Bandwidth
  - Services
- Internet2 Network and Advanced Services Update
- ARRA and Stimulus Update
- Scientific Involvement and Outreach
  - USCMS
  - USATLAS
- Performance Software Development

# Scientific Outreach

- Internet2 remains committed to the advancement of scientific research
- Scientific Communities/Projects:
  - Physics
    - LHC Community, LIGO
  - Astronomy
    - eVLBI, LSST, SDSS
  - Climate
    - NEON, OOI, NCAR, NOAA
  - Biology
    - JGI
  - Structural Engineering
    - NEES

# Scientific Outreach – LHC Specific

- Driving network requirements
  - Capacity increases
  - Traffic engineering for scientific traffic
- Advancement of software to facilitate scientific research
  - ION Service
    - Bandwidth on demand for time sensitive or intensive network use
    - Integration into existing physics software (LambdaStation, TeraPaths)
    - Encouraging use for institutions of all sizes (Tier1s through Tier3s)
  - perfSONAR
    - Open platform for the collection, storage, and exchange of network measurements in a multi-domain fashion
    - Development and advancement of the pS Performance Toolkit – a set of measurement tools and supporting control framework to ensure network performance.

# Scientific Outreach– LHC Specific

- USATLAS
  - Continued support of the pSPT for Tier1 and Tier2 centers
    - Latest Release is 3.1.3 (http://psps.perfsonar.net/toolkit)
  - Network debugging
    - Supporting role in using the pSPT to solve network problems
    - Ongoing exercise at a small number of Tier2s and Tier3s
- USCMS
  - No formal relationship on use of the pSPT
    - Tier1 and several Tier2s utilize pSPTs on a diagnostic basis
  - The Research and Education Data Depot network (REDDnet) is an NSF funded initiative to study and investigate data storage along network paths of interest
    - REDDnet has adopted perfSONAR tools to monitor the health and status between locations
    - REDDnet will include monitoring tools on all future software distributions

# USATLAS Use Case

- 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

# USATLAS Use Case

- 2009 – All sites still on original hardware, running 3rd generation (3.1 and 3.1.1) 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

# USATLAS Use Case

- 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

# USATLAS Use Case

- 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).
  - Weekly calls (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 to debug perceived problems

# USATLAS Use Case

- 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

# USATLAS Use Case

- 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

# USATLAS Use Case

- Future
  - Worrying about Tier3s – evaluating new ISO release before recommending
  - Tier3s may not want 2 servers
    - Purchase at least one for diagnostics – occasional testing
  - Tier3s could not do a full mesh of testing with Tier2s and Tier3s (too much traffic)
  - KOI machines may be replaced with a comparable piece of hardware

# Outline

- Internet2 Architectural Directions
  - Hardware
  - Bandwidth
  - Services
- Internet2 Network and Advanced Services Update
- ARRA and Stimulus Update
- Scientific Involvement and Outreach
  - USCMS
  - USATLAS
- Performance Software Development

# Performance Software Development

- Internet2 Performance Initiative
  - Development of software to support monitoring and diagnostic activities for end to end performance problems
  - Active and Passive measurement, along with scalable and extensible control frameworks
- Software
  - BWCTL
  - NDT
  - OWAMP
  - perfSONAR
- For more information please visit:
  - http://www.internet2.edu/performance

# perfSONAR Adoption

- *perfSONAR* is gaining traction as an interoperable and extensible monitoring solution
- Adoption has progressed in the following areas:
  - R&E networks including backbone, regional, and exchange points
  - Universities on an international basis
  - Federal labs and agencies in the United States (e.g. *JET* nets)
  - Scientific Virtual Organizations, notably the LHC project
- Recent interest has also accrued from:
  - International R&E network partners and exchange points
  - Commercial Providers in the United States
  - Hardware manufactures

# perfSONAR Software Update

- perfSONAR-PS Release 3.1
  - Formally released in Sept 2009
    - Subsequent releases available as needed
  - Available as source code and RPM packages for several architectures and platforms
  - Installation through dependency management systems (e.g. YUM, APT-RPM) is available.
    - http://software.internet2.edu
  - For more information please visit:
    - http://psps.perfsonar.net

# perfSONAR Software Update

- pS Performance Toolkit Release 3.1
  - Formally released in Oct 2009
    - Generating updates to meet security and stability concerns quarterly
    - Release 3.1.3 completed in April 2010
  - Network monitoring solution on a single CD
    - Ability to run diagnostic tools
    - Ability to perform regular measurement according to a schedule
      - Latency
      - Bandwidth
    - perfSONAR-enabled to facilitate data sharing
- pS Performance Toolkit 3.2
  - Expected release July 2010
  - Next generation based on CentosOS Linux
- For more information please visit:
  - http://psps.internet2.edu/toolkit


# An Overview of Architectural Directions and Advanced Services

June 10th 2010, Transatlantic Networking for LHC Experiments

Eric Boyd, Internet2 Deputy CTO

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