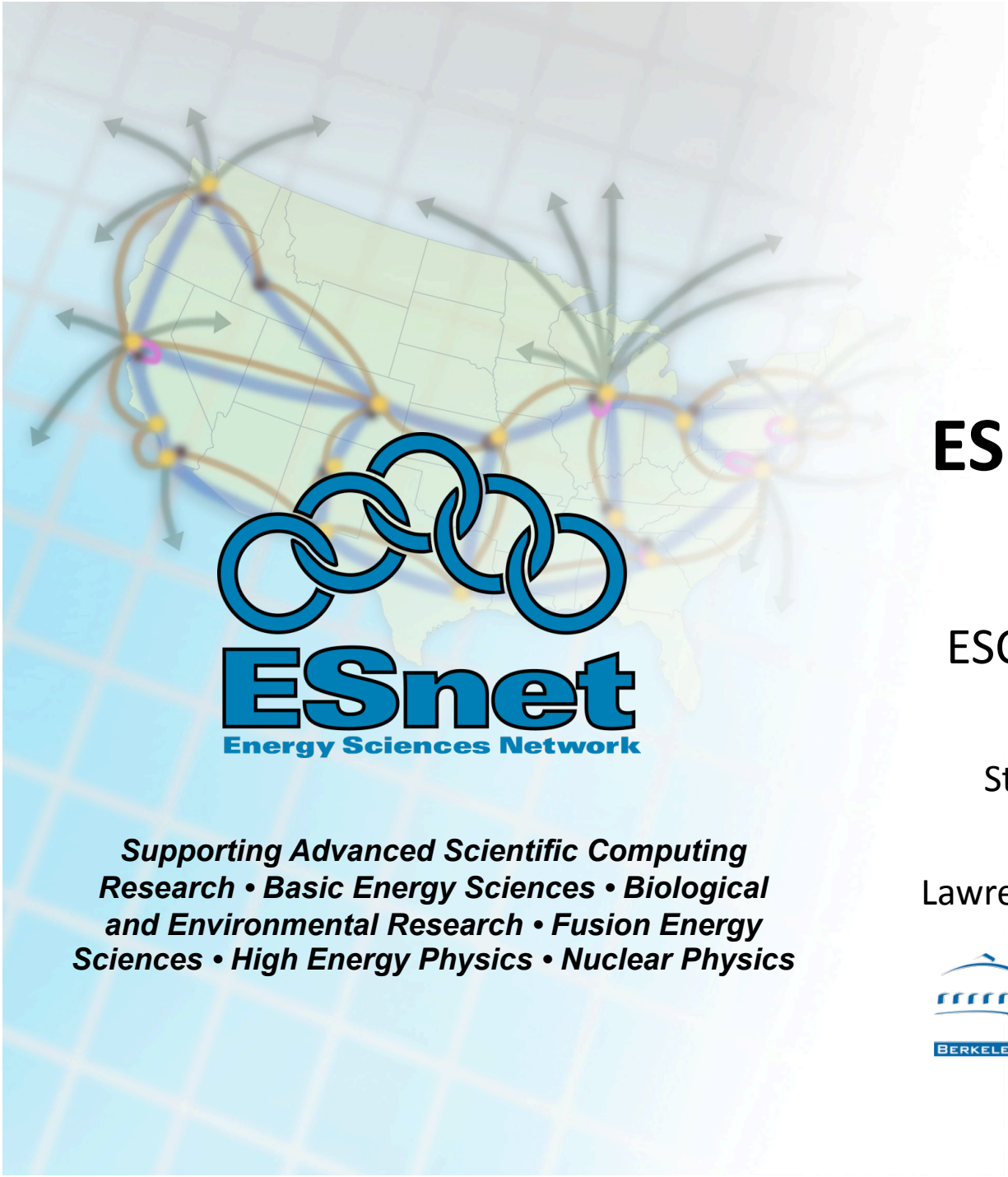


ESnet Update

Feb 3, 2009
ESCC, Salt Lake City

Steve Cotter, Dept Head
steve@es.net
Lawrence Berkeley National Lab

Supporting Advanced Scientific Computing Research • Basic Energy Sciences • Biological and Environmental Research • Fusion Energy Sciences • High Energy Physics • Nuclear Physics

BERKELEY LAB

U.S. DEPARTMENT OF ENERGY
Office of Science

Outline

- Staff Updates
- Network Update
- Advanced Networking Initiative
- ESnet Projects
- Infrastructure Projects
- Staff Projects

Staff Update

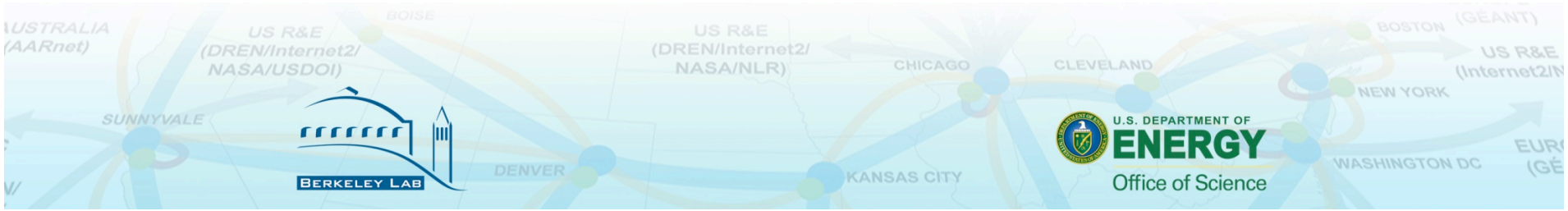
New hires:

- Hing Chow: Project Manager (ANI)
- Chris Tracy: Network / Software Engineer (ANI)
- Andy Lake: Software Engineer (ANI)
- Inder Monga: Network / Software Engineer (ANI)
- Josef Grosch: Sys Admin
- Positions posted:
  - Chief Information Strategist
  - Software Engineer

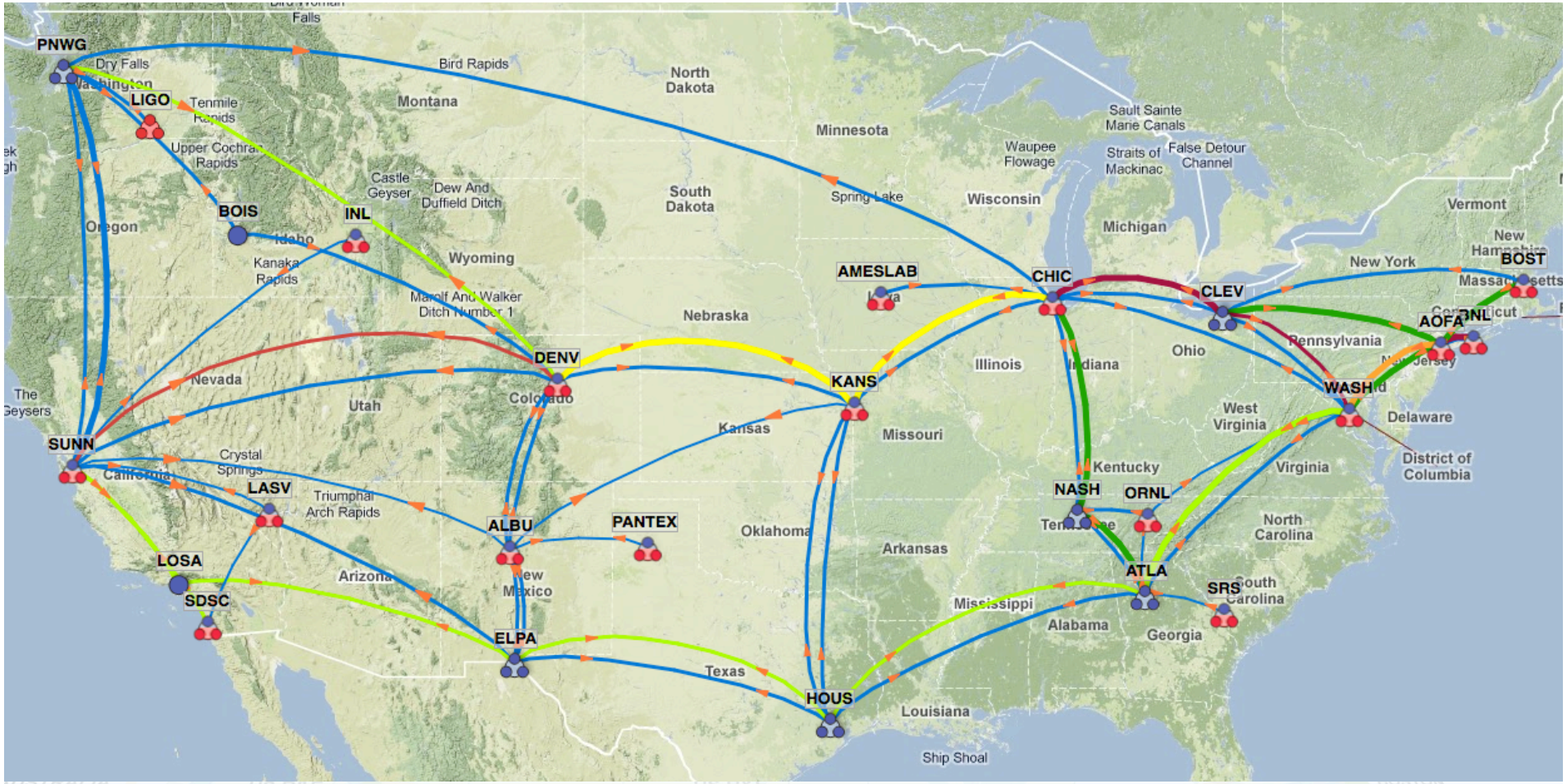
ESnet
Energy Sciences Network

Network Update

Current Status, Upgrades



ESnet4 Network



Equipment Upgrades / Installs

Peering upgrades:

- EQX-SJ: installed MX480 on Oct 15th
- EQX-ASH: installed MX480 on Nov 30th
- EQX-CHI: Pending MX480 install on Feb 18th

Site / hub upgrades:

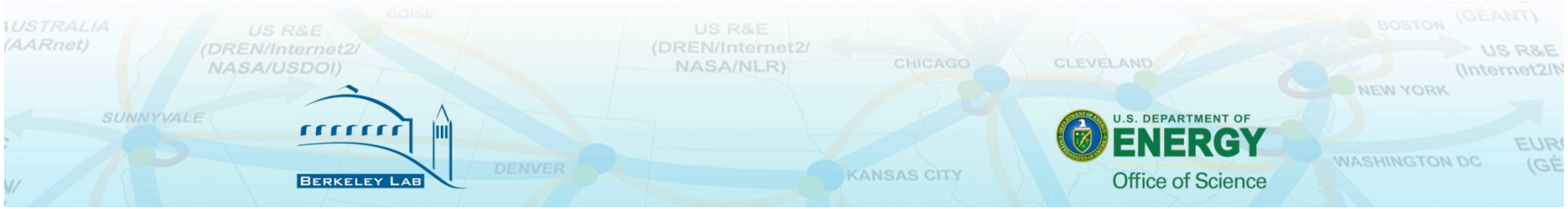
- ORNL M20: Upgraded to MX480 on Dec 17th
- AMESLAB M10: Upgraded to M10i on Jan 13th
- FORR 7206: Upgraded to M7i on Dec 22nd
- DOE-GTN 7206: Scheduled upgrade to M10i late Feb.
- DOE-NNSA M10: Scheduled upgrade to M7i late Feb.
- PPPL M120: Scheduled upgrade to MX480 mid-Mar/Apr
- GA M7i: Scheduled upgraded to MX480 mid-Mar/Apr

Circuit Installs

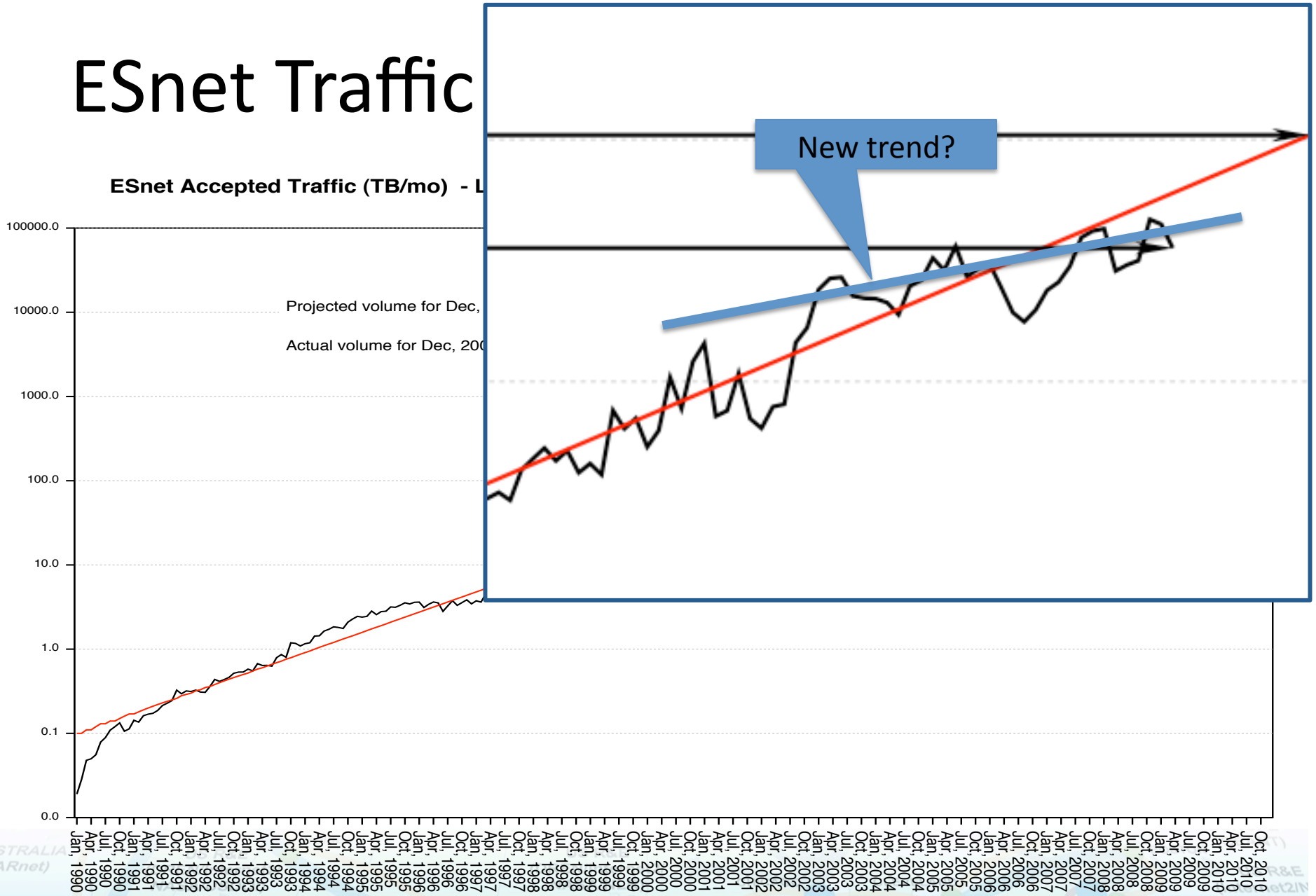
- 10G connection at BOIS with PNNL for backup peering Nov 10th
- 10G peering at PNWG-HUB with Korea (KSTAR & KISTI) Nov 11th
- Combine of LOSA-SUNN & ELPA-LOSA into new ELPA-SUNN SDN (prior to the decommission of LOSA-HUB) Dec. 3rd
- 10G Equinix ASH (DC2) fabric upgraded on Jan14th
- 10G Equinix SJ (SV1) fabric upgraded on Jan19th
- OC12 between DENV-HUB and Pantex Jan 28th

Planned Upgrades / Installs

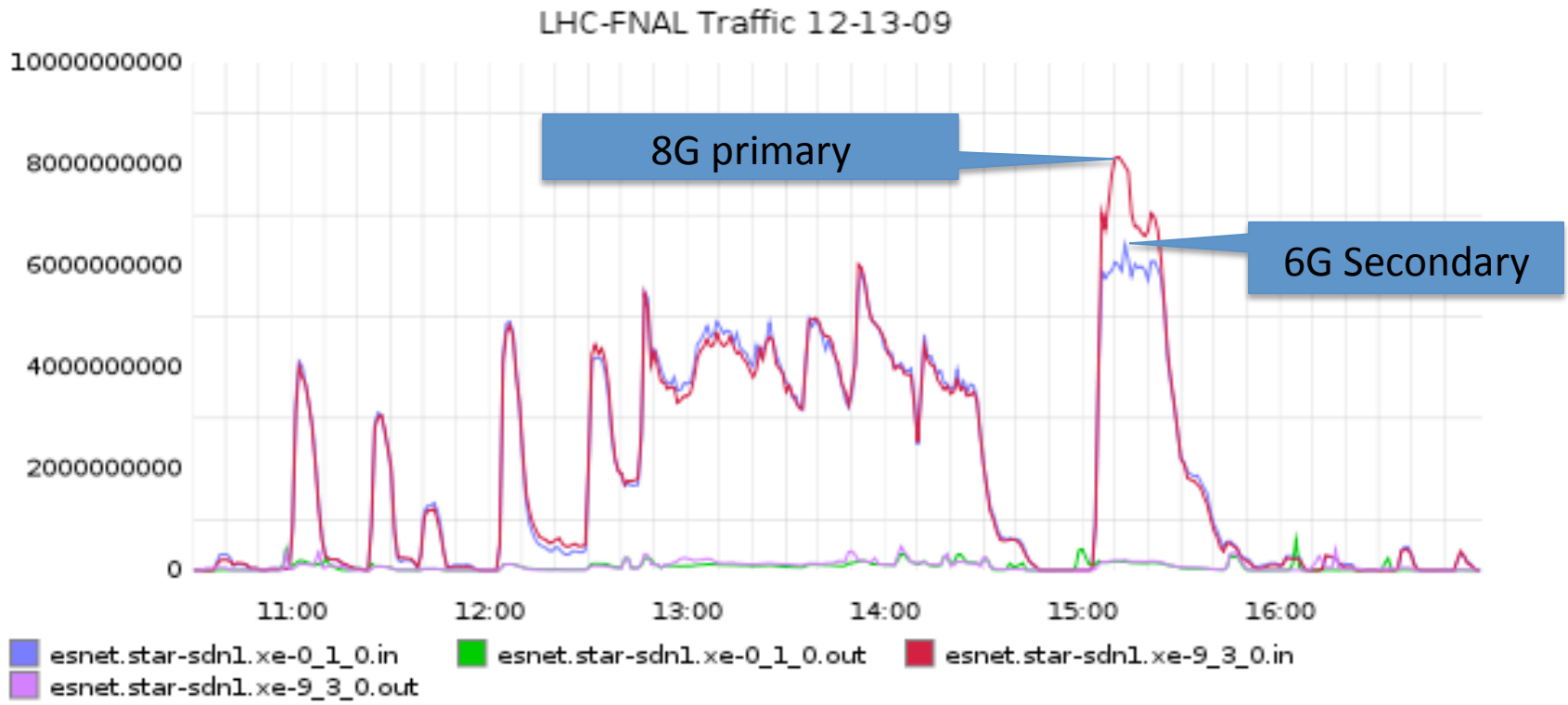
- DS3 between WASH-HUB and SNL-DC (now in test)
- 10G EQX-CHI fabric upgrade scheduled Feb 19th
- 1GE wave in BOIS to INL via IRON (TBD)
- 1GE links in D.C. Area for Germantown, IN to WASH-HUB (ordered)
- OC3c for NSO to LASV-HUB (ordered)
- 10G Wave between CHIC-HUB and EQX-CHI (ordered)
- 10G Wave between WASH-HUB and EQX-ASH (pending)
- Future 10G peering with MERIT @ Starlight
- Future additional 10G peering with GPN @ KANS-HUB



ESnet Traffic

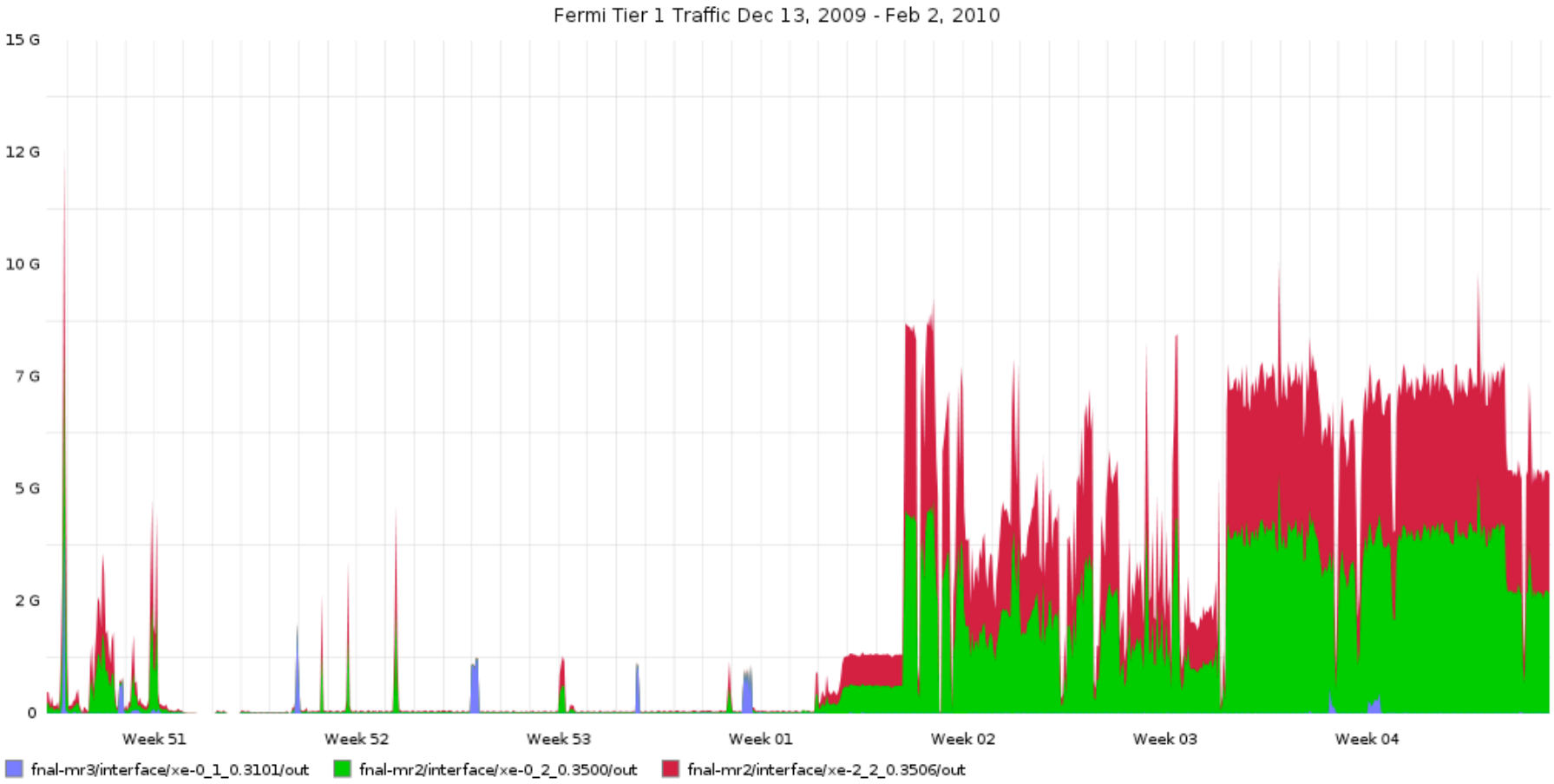


Dec 12 - 2.36 TeV LHC Run



Monitoring traffic on primary and secondary ESnet/USLHCnet interconnects at StarLight

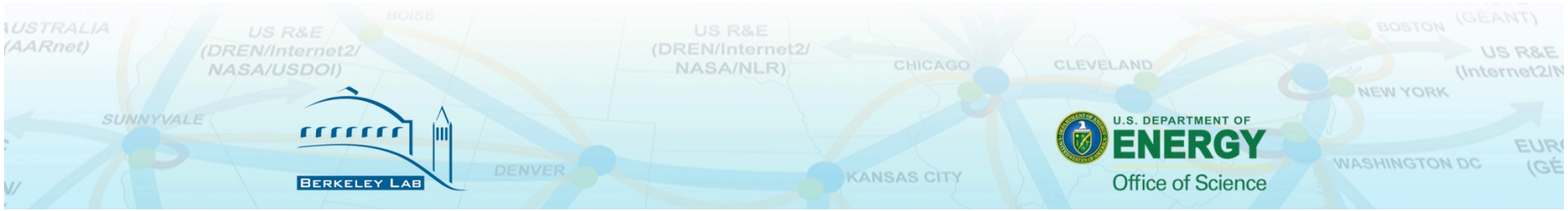
Fermi Traffic: Dec 13 – Feb 2



Monitoring traffic on primary and secondary ESnet/USLHCnet interconnects at StarLight

Advanced Networking Initiative

Prototype Network and Testbed



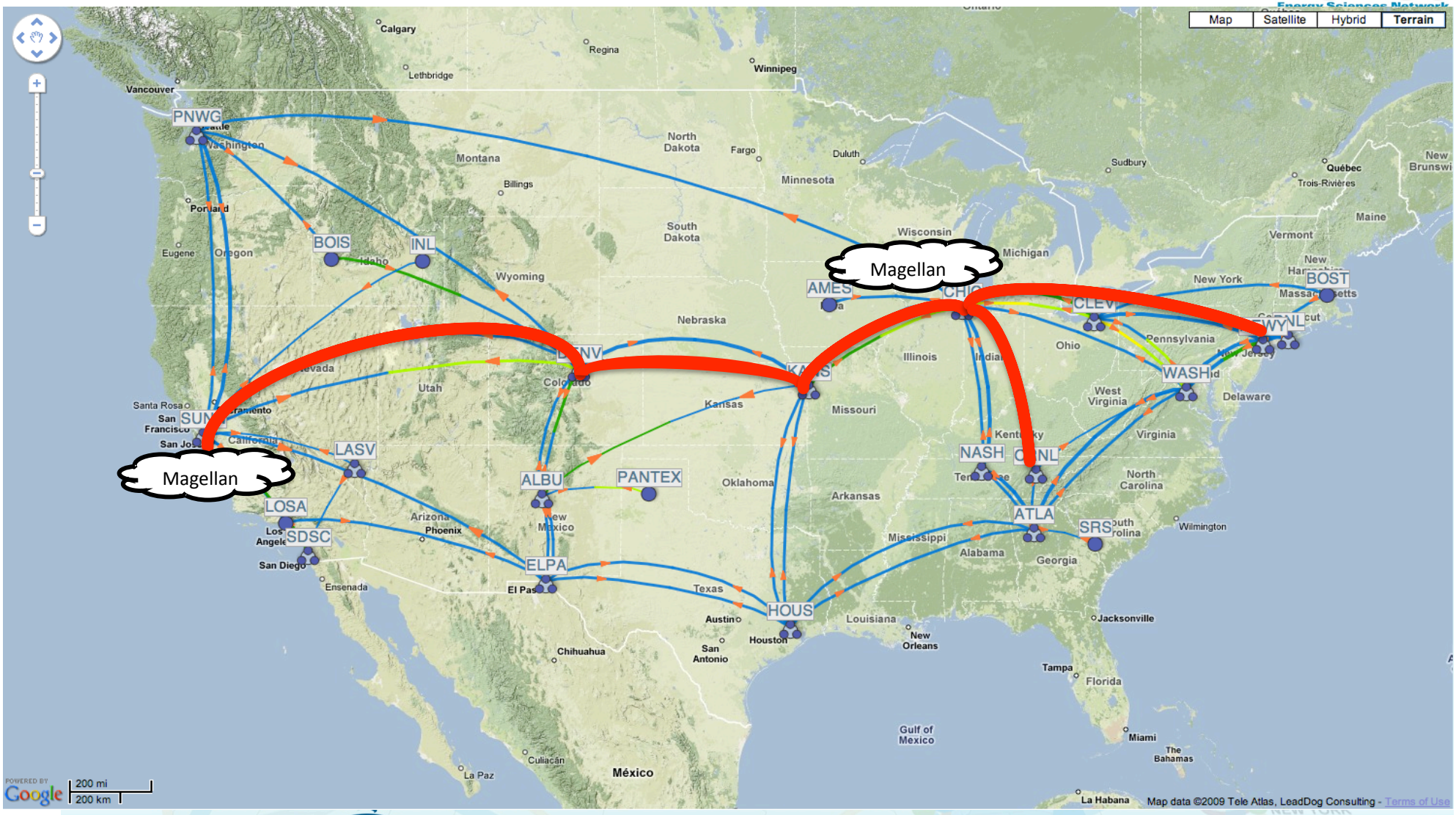
Advanced Networking Initiative

- ANI Project scope:
  - Build end-to-end 100 Gbps prototype network between DOE supercomputers and MANLAN
  - Build a network testbed facility for researchers and industry
    - DOE has funded an additional $5M in network research that will use the testbed facility
- Magellan:
  - Separate DOE-funded nationwide scientific mid-range distributed computing and data analysis testbed to explore whether cloud computing can help meet the overwhelming demand for scientific computing
  - NERSC / LBNL & ALCF / ANL configured with multiple 10's of teraflops and multiple petabytes of storage, as well as appropriate cloud software

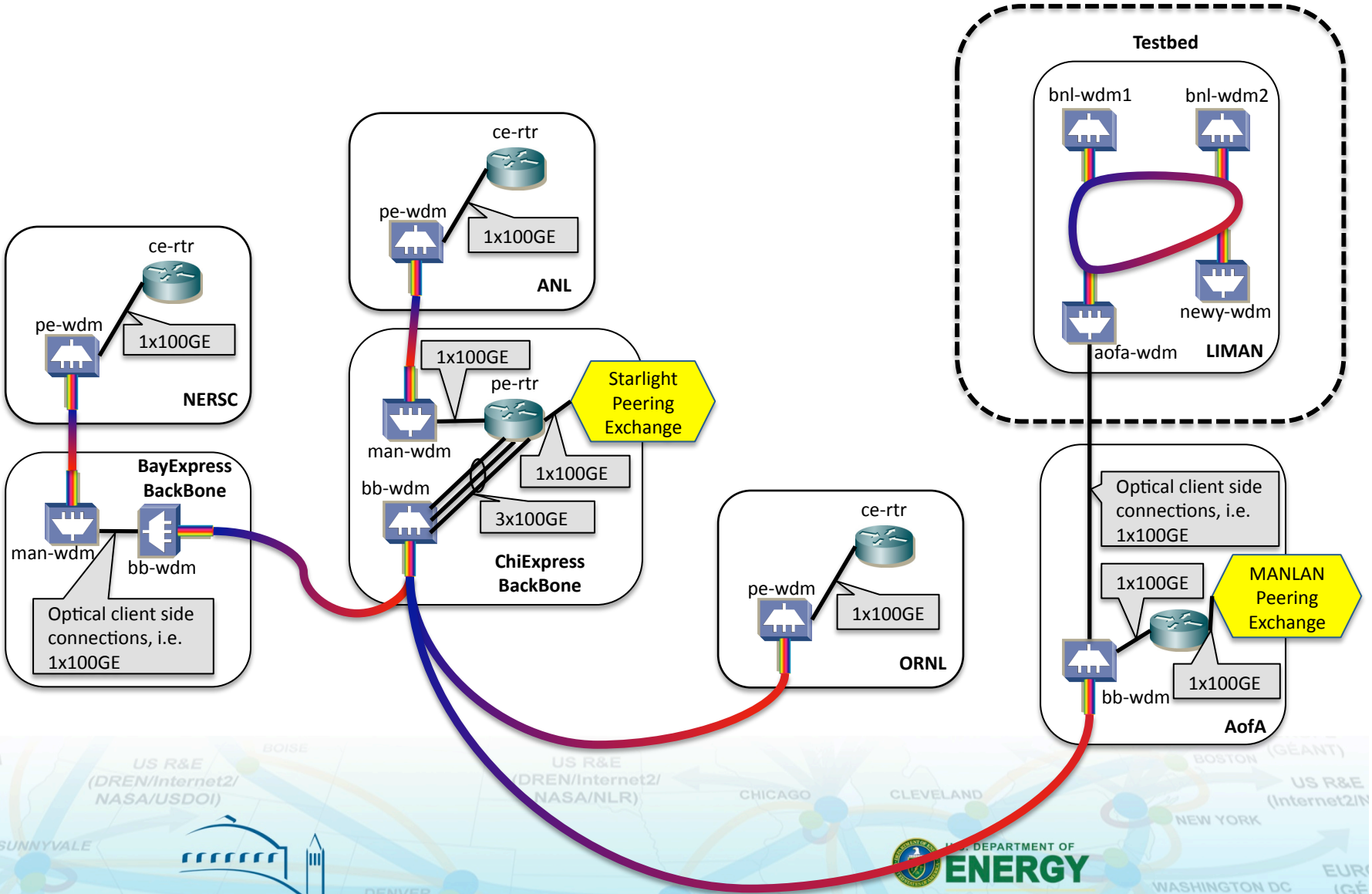
ANI Project Goals

- Prototype network:
  - Accelerate the deployment of 100 Gbps technologies
  - Build a persistent infrastructure that will transition to the production network ~2012
    - Key step toward DOE's vision of a 1-Terabit network linking DOE supercomputing centers and experimental facilities
- Testbed:
  - Build an experimental network research environment at sufficient scale to usefully test experimental approaches to next generation networks
    - Funded for 3 years, then roll into the ESnet program
    - Breakable, reserveable, configurable, resettable
    - Enable R&D at speeds up to 100 Gbps

ANI Topology



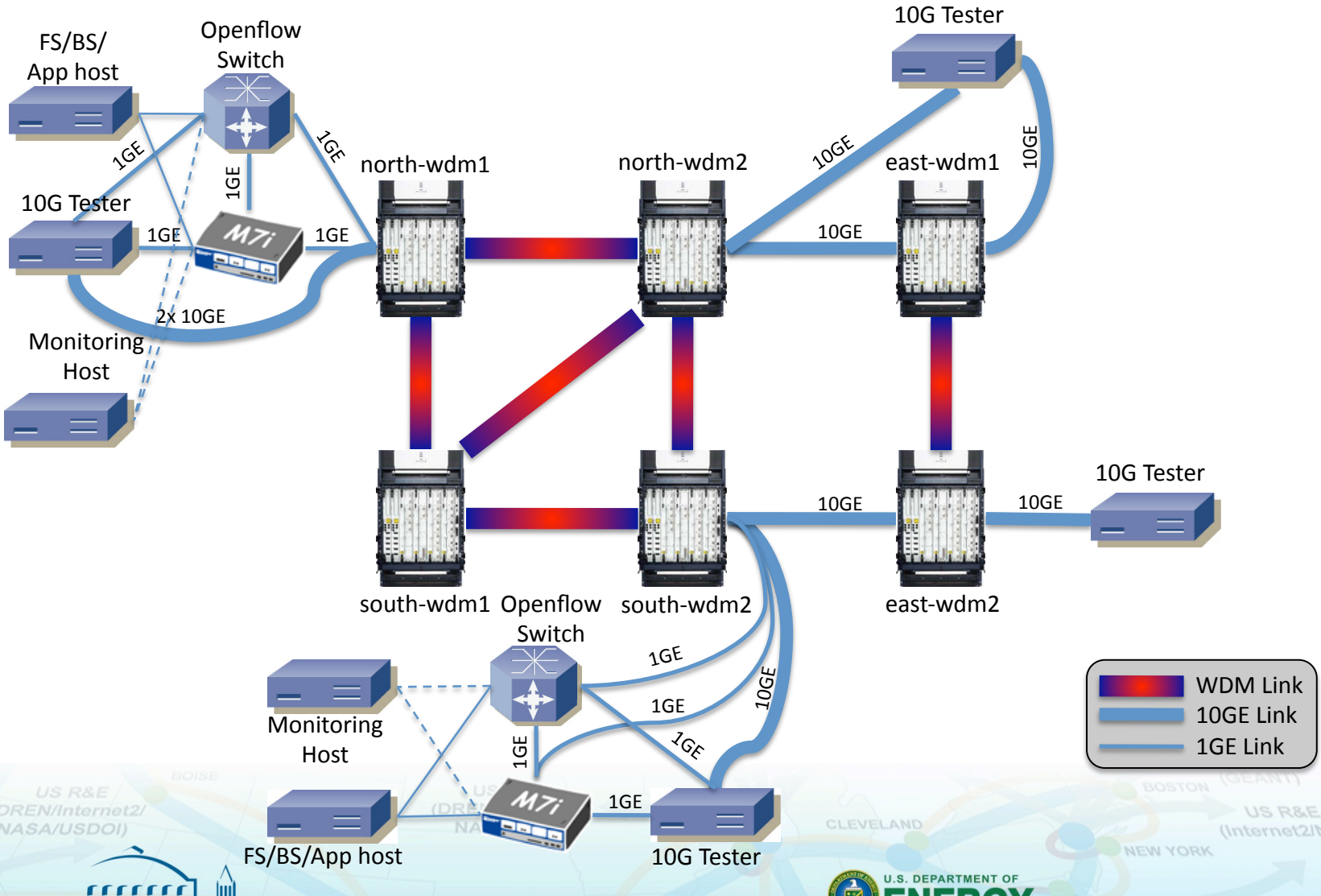
ANI Baseline Design



Testbed Overview

- Progression:
  - Start out as a tabletop testbed, then move out to the wide-area when 100 Gbps available
- Capabilities:
  - Ability to support end-to-end networking, middleware and application experiments, including interoperability testing of multi-vendor 100 Gbps network components
  - Dynamic network provisioning
  - Plan to acquire dark fiber on a portion of testbed footprint to enable hybrid (layer 0-3) network research
  - Use Virtual Machine technology to support protocol and middleware research
  - Detailed monitoring so researchers will have access to all possible monitoring data from the network devices

Tabletop Testbed Design

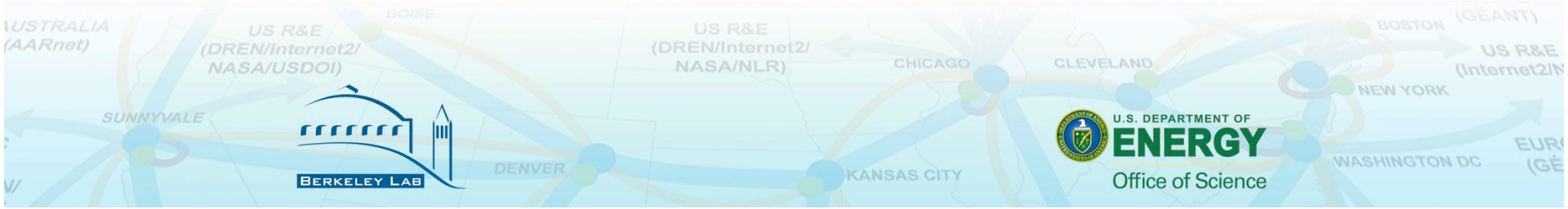


ANI Progress to Date

- Tabletop testbed equipment – ordered, racked at LBL
- 100 Gbps technology research & evaluation phase - ongoing
  - Meetings / briefings with vendors
  - Equipment in ESnet lab
- Transport RFP written – going through reviews
  - Acquire 100 Gbps wave service from a carrier
  - Don't need to own / control optical gear
  - Plan to run OSCARS layer 2 / 3 services across network
  - Dark fiber is part of DOE's long-term research agenda
- Routing / Switch RFP – Summer 2010
  - ESnet will purchase this equipment
  - Will conduct testing / evaluation as part of selection process

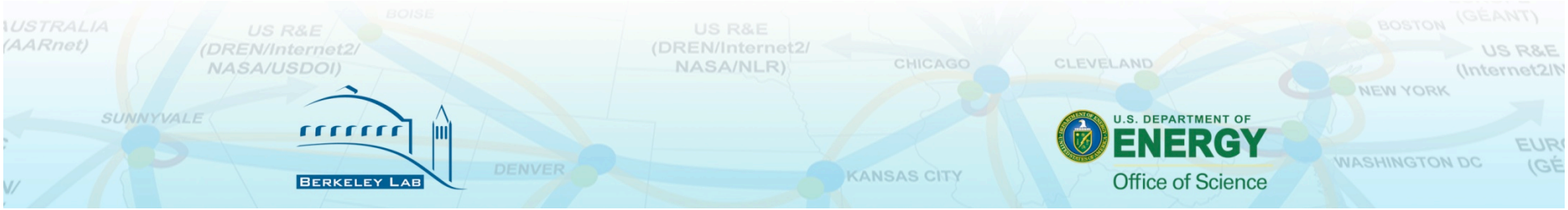
ESnet Projects

perfSONAR, OSCARS, Fenius, etc.



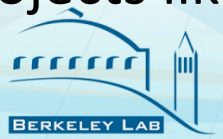
perfSONAR

- ESnet is a key member of the perfSONAR collaboration – http://www.perfsonar.net
- Numerous test hosts deployed, automated tests are run regularly (http://stats1.es.net)
- Test hosts are available to ESnet sites and R&E collaborators for bwctl/iperf tests
- Test and measurement is very helpful in locating the cause of network performance problems

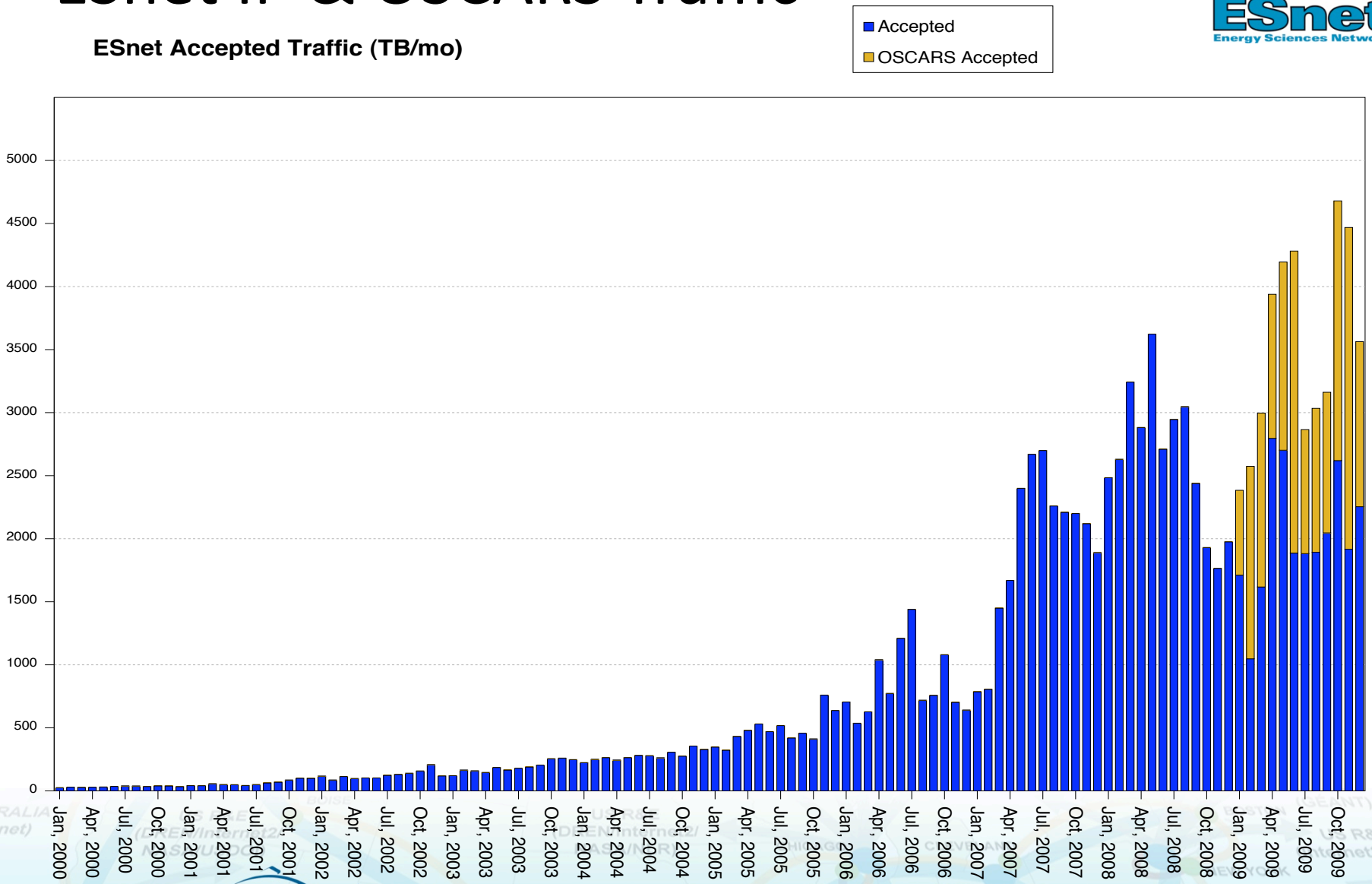


OSCARS: "Multi-Domain, Virtual Circuits" as a Service

- Successfully deployed within ESnet SDN
- OSCARS Software is Open-source (oscars-idc.googlecode.com)
  - A resource for the community
  - Example: Internet2 ION leverages OSCARS
- Ongoing challenge: Build dual-purpose software
  - Enable researchers to innovate using this framework
  - Provide robust product-grade software
  - Take advantage of new innovations and research in this field
- Direction forward: Build critical mass around the open-source effort
  - Collaborate with like-minded researchers and open-source projects like Open-DRAC and Openflow



ESnet IP & OSCARS Traffic

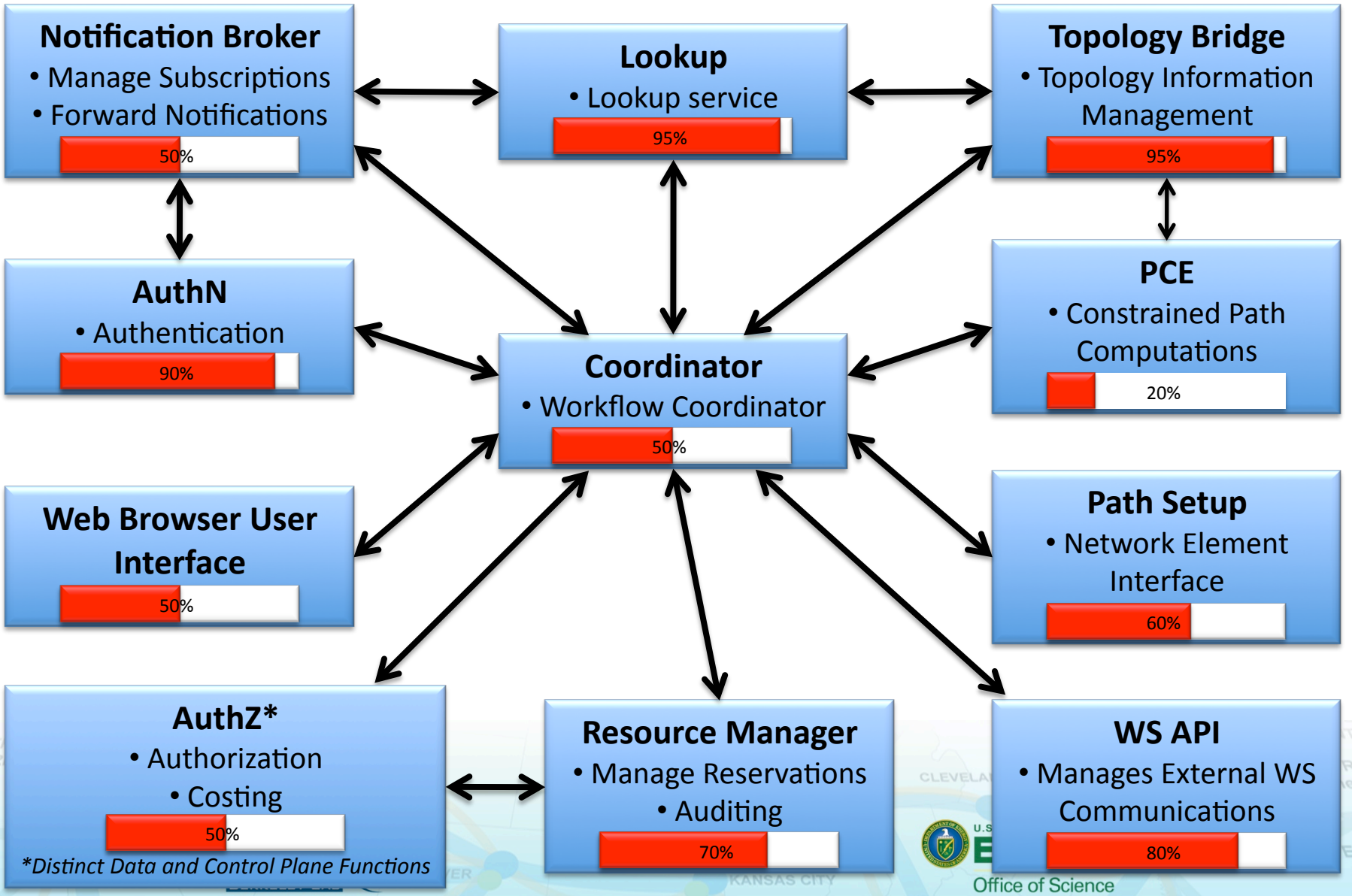


ESnet
Energy Sciences Network

BERKELEY LAB

U.S. DEPARTMENT OF ENERGY
Office of Science

OSCARS 0.6 – Target Release 3/10



Fenius

- At the 9th Annual Global LambdaGrid Workshop in Daejeon, Korea.
  - ESnet, KISTI, AIST and the EU-funded Phosphorus project successfully demonstrated interoperability between their network resource scheduling systems
- Coordinated within the activities of the GLIF consortium GNI API Task Force
  - Developed specialized software to enable the different network scheduling services to be used and monitored through one common interface.
- Demonstrated again at Supercomputing 2009

Site Outreach Program

- Started Jan 1, led by Eli Dart
- Goal is to increase effective use of networks for science
  - Leverage ESnet's experience in helping sites solve problems and increase performance
  - Understand site network infrastructure, drivers, and long-term plans
  - Help sites and disciplines build networks well-matched to their needs

Site Outreach Program

- Pilot underway: SLAC
- Looking at:
  - Network architecture
    - Impact of “converged networks” on high bandwidth data transfers – and possible need for separation of Science and Enterprise networks
    - Adequate buffering on switches and routers
  - Host and system configuration
    - Dedicated hosts for wide area data transfer
    - Proper TCP tuning
  - Test and measurement infrastructure (e.g. perfSONAR)

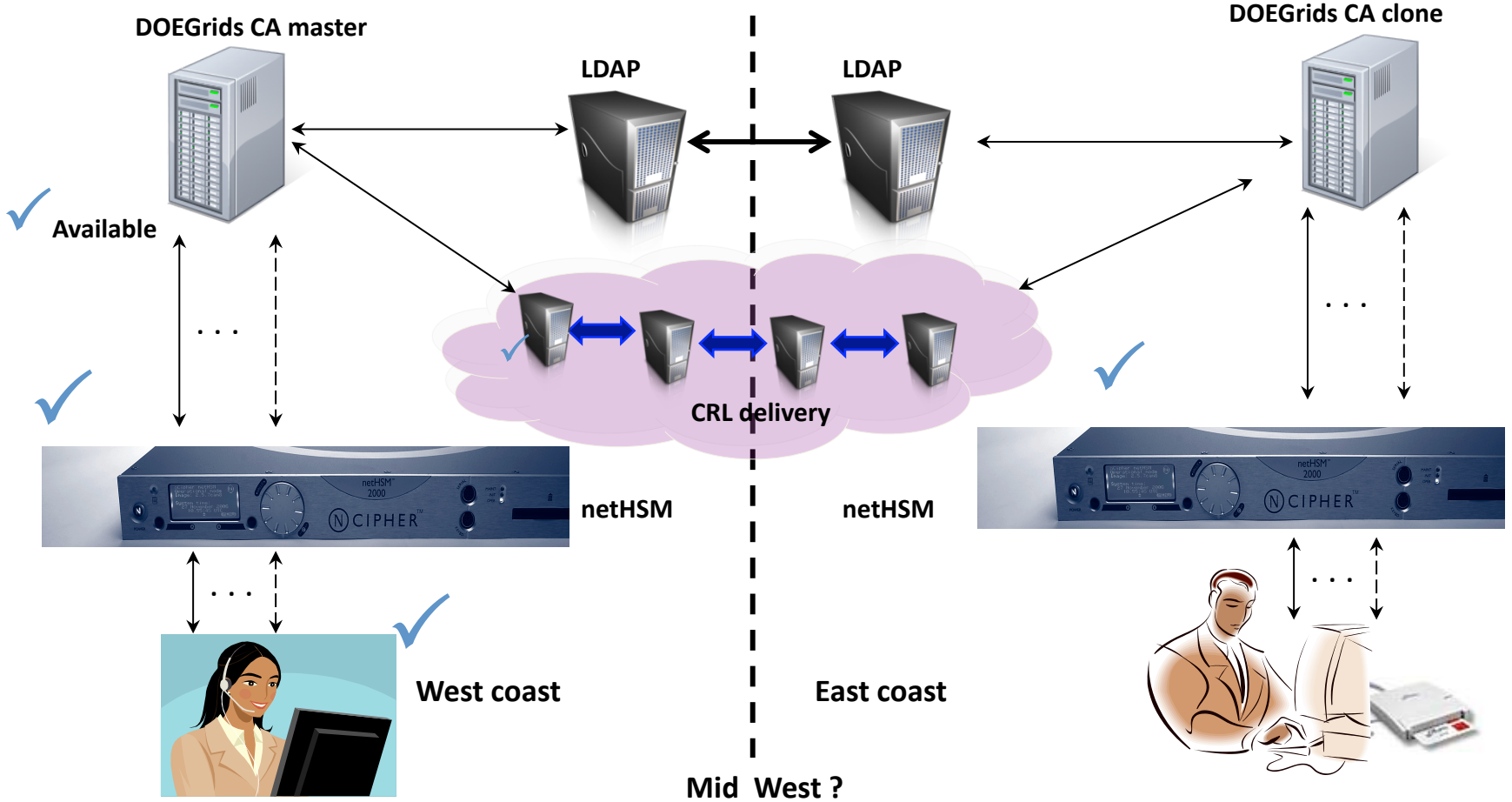
Long Island MAN

RFP responses due: Feb 22nd

Provide High Availability CA/PKI Service

- Replication & geographic dispersion of key management
  - netHSM (network enabled Hardware Security Module) are used to manage the key pairs that the CA uses
- Replication & geographic dispersion of CA (signing initiator, UI, database)
  - RedHat Certificate System
- Replication & geographic dispersion of CRL publishing
  - ANYCAST
  - CRL Distribution from the Cloud – eg Amazon Cloudfront
- Remote operator & geographic dispersion of operator
  - Remote operator service
- Nagios monitoring system
  - A new development for this project

DOEGrids CA with High Availability



Science Identity Federation

- Interoperable Identity for DOE labs

  … based on the well-known

- Shibboleth authentication & authorization software

  … so that labs can

- Federate with InCommon
  - Which is US Higher Education Shibboleth Federation: see incommonfederation.org
  - And other federations as needed

Science Identity Federation Program

- Training – get early adopters up to speed on Shibboleth IDP and integration with their home service
  - Shibfest Mar 30-31 at FNAL – details to come soon
- Write a minimal charter for activities
  - Applications and services will define the mission ultimately
  - Need to look at DOE services, user facilities, supercomputer centers &c for good integration candidates
  - Look at specific lab needs and interests – expressing attributes like “LoA” or “citizenship”
- Provide demonstration services
  - Confluence (a CMS with wiki-like features)
  - A demo Grid credential CA
- Activities to come:
  - Exploring interoperation with other services
    - CILogon, other SAML activities, EU federations
  - Alternative technology: OAuth, OpenID
  - TFPAP & ICAM – alternatives to federation

Science Identity Federation Contact

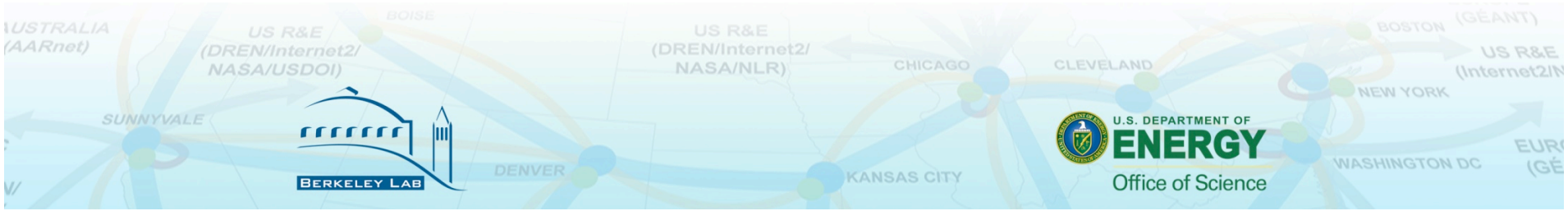
- If you're interested, and have some relationship to the DOE lab community or projects:

  http://groups.google.com/group/science-federation

  We're using this "private public" group to bootstrap;
- Or email Michael Helm helm@es.net

Infrastructure Projects

OpenDevNet, DNSSEC, Spectrum, etc.



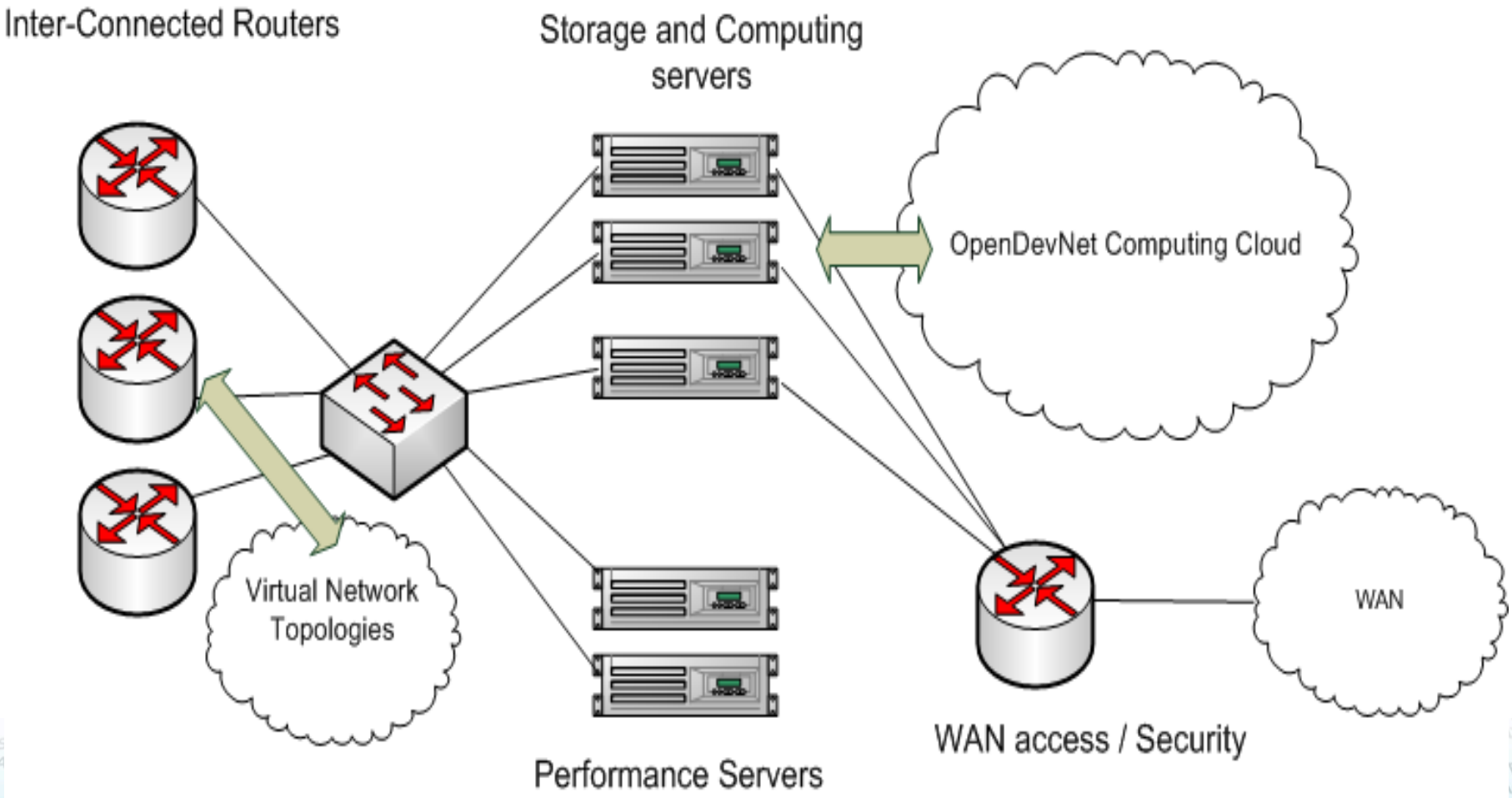
ESnet OpenDevNet

Testing tool:

- Software lifecycle: build, debug, QA
  - Platform to build software-as-service, ready for deployment on production servers
- 'Slice' model: made of virtual resources (VM, virtual network topologies) and physical resources (routers, performance nodes, WAN access)
- Testing platform for third party technologies: sandbox, demo, testing.
- Deployment of services required for experiments
- Deployment of user-specified network topologies
- Testing platform for other testbeds

Contact: Eric Pouyoul lomax@es.net

OpenDevNet Architecture



DNSSEC

Domain Name System Security Extensions (DNSSEC) provide authentication and ensure the integrity of the DNS through the use of cryptographic signatures generated with public key technology

- ESnet using DNS Signer, a dedicated appliance from Secure64 Corp
- Completed in December – ahead of the mandate from the U.S. Office of Management and Budget (OMB)
  - Top-level .gov domains had to be signed by February 2009
  - Those immediately under the .gov domain had to implement DNSSEC by the end of 2009

Spectrum Upgrade

- Spectrum r9.1:
  - All Juniper routers are being polled with and send traps using IPv6
  - The Spectrum MPLS transport manager auto-discovers the OSCARS circuit topology.
  - Additional thresholding alarms for interface utilization, errors and route engine temperature.
  - OSCARS LSP alarms have now been integrated into the daily outages on the Planned Maintenance Calendar (PMC)
    - These measurements will provide a basis for OSCARS availability metrics

Cfengine Installation

- Needed to provide automated configuration and maintenance of servers, from a policy specification
  - Every measurement host (60+) must be maintained in a known good state since OS & software differences affect measurement results
  - Software systems running (perfSONAR, OSCARS) are under active development & must be upgraded frequently
- Deployment and configuration underway for Linux and FreeBSD hosts
  - Automate new OS installations on a large set of hosts
  - Automated patching
  - Configuration management verification & reporting

Blade Server Deployment

- Needed a way to effectively manage growing number of servers (250+)
  - Increasing complexity, cost, power consumption, demands on staff
- Phased deployment over next year
- Benefits:
  - Reduce rack space consumed up to 80%
  - Reduce power consumption up to 65%
  - Eliminates near-term need for HVAC and power upgrades in ESnet datacenter (estimated at $2M+)
  - Automated management reduces FTE costs

Nagios & OpsView

- An industry standard monitoring system that enables organizations to identify and resolve IT infrastructure problems before they affect critical business processes
  - Needed a way to monitor servers and other equipment (i.e. video MCUs)
  - Implementation is automated and fully redundant
    - Automatic reporting of outages, availability, health, response, etc.

Community Engagement

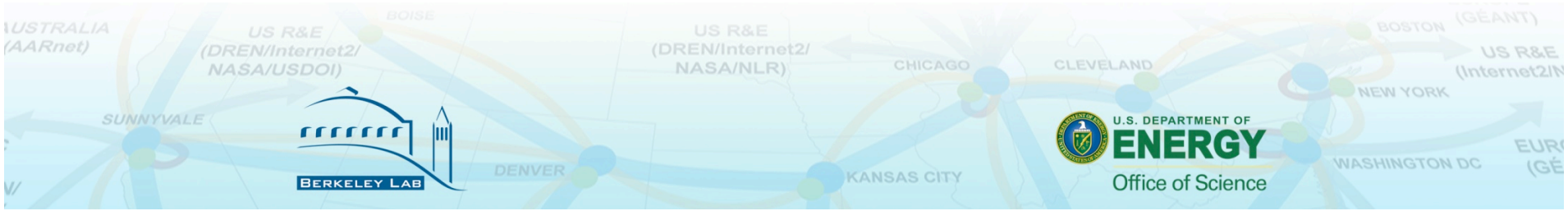
- Network Requirements Workshops
  - Advanced Scientific Computing
    - Published Dec 2: http://www.es.net/hypertext/requirements.html
  - High Energy Physics
    - Workshop Completed: Aug 28, 2009
    - Report undergoing final review
  - BER / BES scheduled for 2010
- DICE (DANTE, Internet2, Canarie, ESnet)
  - Continue to collaborate on virtual circuits, perfSONAR
  - DICE Framework agreement
- Technical discussions with other networks
  - Surfnet, CERN, Nordunet, RNP, AIST, etc.

Advancing Standards

- Contributions to relevant public standard bodies
  - Leverage operational and network research experiences
- Current focus and participation
  - GLIF (Global Lambda Integrated Facility)
    - Fenius
    - Advanced GOLE initiative
  - OGF (Open Grid Forum)
    - NM: Network Measurements
    - NMC: Network Measurement and Control
    - NSI: Network Service Interface
    - NML: Network Markup Language
    - GHPN: Grid High Performance Network Research

Staff Projects

PMC, Twitter, Weathermap, View, etc.



Planned Maintenance Calendar

- Developed in-house in response to a DOE requirement to categorize planned vs. unplanned network service outages
  - Improves the quality of ESnet notifications and saves ESnet operators significant time and effort
- Uses Spectrum Network Management System
  - Provides detailed outage data
  - Maintenance event scheduling and correlation
  - Path-based service availability
  - Maintenance impact prediction
  - Targeted notification
  - Topology mapping

Contact: Mike O'Connor moc@es.net

Passive Network Measurement

- ESnet has released two portions of it's SNMP data collection system as open source software: TSDB and ESxSNMP
  - TSDB (Time Series Data Base) is a database optimized for storing large amounts of time series data
  - ESxSNMP (ESnet eXtensible SNMP system) is a flexible SNMP polling system which has been designed for high reliability and minimal upkeep. ESxSNMP uses TSDB to store its time series counter data.
  - Both are available on Google Code

Contact: Jon Dugan jdugan@es.net

Twitter Feeds

- Exploring how to provide immediate status messages directly from ESnet alert systems as they occur
- Followers of these topical community feeds receive real-time concise and meaningful messages about service impact and the state of the network.
- In the spirit of Twitter, messages will be relevant at the instant they are sent, not tomorrow or next week, right now
- Potential Twitter Feeds:
  - Customer service impact – per site
  - Carrier circuit outages
  - OSCARS virtual circuit outages
  - Peering outages

Contact: Mike O'Connor moc@es.net

ESnet Weathermap

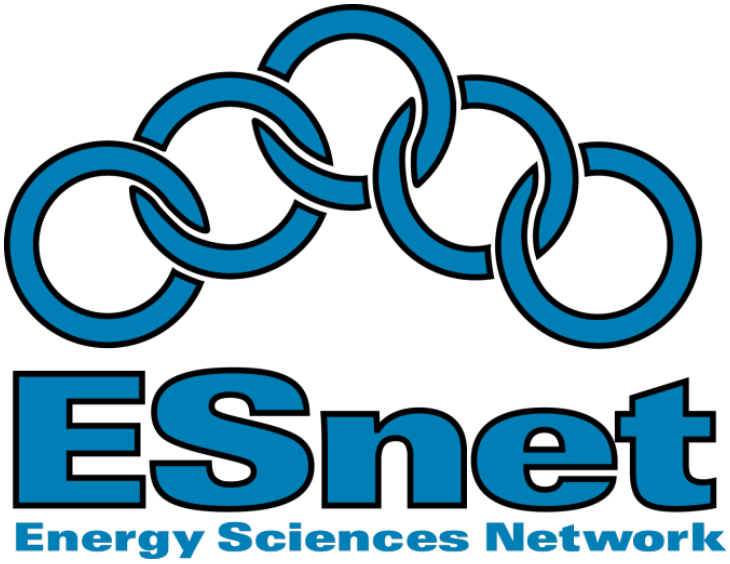
- By collaborating with other R&E partners, ESnet has constructed a display that captures bandwidth utilization of each circuit and graphically displays this information for ease of viewing
- Network traffic statistics by interface, network (IP & SDN), peering connection
  - Shows average and maximum usage over period from 30 secs to 60 mins
  - Shows OSCARS circuit topology and reserved bandwidth
- 24 hr playback capability
- http://www.weathermap.es.net/test

Contact: David Robertson dwrobertson@lbl.gov

ESnet View

- Google Earth-based network visualization tool
  - Layered network information
    - Fiber plant, optical layer, routing/switch layer
  - Links to ESnet databases
- Had many different tools developed over the years
  - Accessing information required many steps, process slow
  - Needed better information management for operational efficiency

Contact: Mark Redman redman@es.net


ESnet
Energy Sciences Network
U.S. DEPARTMENT OF
ENERGY
Office of
Science