ESnet Update

Steve Cotter  steve@es.net
ASCAC Meeting - March 4, 2009
American Geophysical Union, Wash DC
ESnet4 Update
Network Services
• OSCARS
• perfSONAR
• Federated Trust
### 2008 Hub & Site Installs

#### JUL
- 1st set of Juniper MX’s arrived at LBNL mid-June

#### AUG
- 2nd set of Juniper MX’s arrived at LBNL mid-Sept

#### SEP
- 14 10GE Internet2 waves installed/split & accepted
- 19 10GE Internet2 waves installed/split & accepted

#### OCT
- 6 10GE Internet2 waves installed/split & accepted
- 2 10GE Internet2 waves installed/split & accepted

#### NOV
- 1 OC12 LANT-SUNN
- 1 10GE Internet2 STAR-CHIC

#### DEC
- 1 10GE MAN-LAN #2
- 1 10GE NLR Temp WASH-STAR
- 1 10GE CIC-OMNiPop at STAR

**Total hub installs:**
- 6 MX480’s
- 19 MX960’s

**New site installs**
- 1 M120 PPPL
- 1 M10i LASV-HUB

---

**Waves installed/split & accepted:**
- 14 10GE Internet2 waves installed/split & accepted
- 19 10GE Internet2 waves installed/split & accepted
- 6 10GE Internet2 waves installed/split & accepted
- 2 10GE Internet2 waves installed/split & accepted
- 1 OC192 private peering with Internet2 CHIC
- 1 10GE NLR AOFA-WASH #192
- 1 ORNL-NASH 10G IP

---

**1st set of Juniper MX’s arrived at LBNL mid-June**

**2nd set of Juniper MX’s arrived at LBNL mid-Sept**
Current Hub Count: 21

- 32 AofA, NEWY*, WASH, ATLA, NASH, CLEV*, BOST*, CHIC, STAR, KANS*, HOUS*, ELPA*, DENV, ALBU, BOIS*, PNWG, SUNN, SNV(Qwest), LOSA*, SDSC, LASV(SwitchNap)*

*9 New hubs since July 2008

Current Backbone Wave Count:

- Internet2/Level3 10G waves:
  - IP: 17 new/split for a total of 25
  - SDN: 25 new/split for a total of 30

- NLR 10G waves:
  - 1 new wave for a total of 5
  - 1 temp wave (STAR-WASH) for used during NLR northern path upgrade
Additional new connections:

• 1 OC12 between SUNN-LASV (support the Las Vegas sites and future OC12 to GA)
• 1 10GE NLR AofA-WASH (Backup for SDN & USLHHCnet)
• 1 ORNL-NASH 10G IP (waiting on there 10GE SDN)
• 1 10G NLR FrameNet connection in WASH-HUB
• 1 10GE MANLAN connection at 32 AofA (this gives us 2 10G peering links)
• 1 10GE NRL Temp WASH-STAR during NLR network upgrade
• 1 10GE CIC-OMNIPOP at STAR
• 1 OC192 ESnet-Internet2 private peering in CHIC
• 1 10GE FRGP (upgrade from 1GE) DENV (Jan 2009)
• 3 10GE Lightower LIMAN#3, LIMAN#4 & AofA-NEWY dark fiber (Feb 2009)
• 1 10GE between BOST to MIT (Feb 2009)
### 2009 MAN & Site Upgrades Timeline

#### JAN

- LBL-MR2, SNV-MR2, SNLL-MR2, LLNL-MR2 & SLAC-MR2 (Completed on or before Jan 27th)

#### FEB

- FNAL-MR2 installed Feb 9th
- JGI-MR2 installed Feb 26th
- FNAL-MR2 installed Feb 20th

#### MAR

- NERSC-MR2 will be installed Mar 5th
- BNL MX’s to be installed on Mar 16th
- ANL MX’s to be installed on Mar 25th

**Footnotes:**
- 1 10GE FRPG (upgrade from 1GE) DENV
- 1 10GE LIMAN#3 AofA-BNL IP up Feb 2nd
- 1 10GE DF circuit between AofA-NEWY up on Feb 2nd
- 1 10GE LIMAN#4 NEWY-BNL up on Feb 9th
- 1 10GE BOST-MIT peering end-to-end
Future Installs

Replace site 6509s (FNAL, ANL & BNL) with MX’s

- BNL (MX960 & MX480) to ship on Mar 11th, install on Mar 16th
- ANL (MX960) ship on Mar 19th, install on Mar 25th

Replace Baman 6509s with MX’s

- JGI-MR2 installed on Feb 26th
- NERSC-MR2 install on Mar 5th (racked up on Feb 19th)

Future circuit installs

- OC12 between LASV-HUB and General Atomic
- DS3 back-up for ORAU to WASH-HUB (awaiting XC at McLean - est. Mar)
- OC12 between DENV-HUB and Pantex (TBD)
- 1GE wave in BOIS to INL via IRON (TBD)
- 10GE SDN wave between PNWG-HUB to PNNL (TBD)
- 10GE SDN wave between NASH-HUB to ORNL (TBD)
- 1GE links in D.C. Area for Germantown, NGA/IN to WASH-HUB
ESnet4 Metro Area Rings

- LI MAN expansion, BNL diverse entry
- FNAL and BNL dual ESnet connection
- Upgraded Bay Area MAN switches
ESnet4: 12 Month Circuit Availability 1/2009
ESnet4: Increasing Site Availability

Site Availability 2/2006 to 1/2007

Site Availability 2/2008 to 1/2009
Historical ESnet Traffic (Tby/mo)

ESnet traffic increases 10X roughly every 47 months

Top 1000 Flows are Beginning to Dominate Network Traffic
Historical ESnet Traffic Growth

Log Plot of ESnet Monthly Accepted Traffic, January 1990 – December 2008

- Aug 1990: 100 MBy/mo.
- Oct 1993: 1 TBy/mo.
- Jul 1998: 10 TBy/mo.
- Nov 2001: 100 TBy/mo.
- Apr 2006: 1 PBy/mo.
- July 2010: 10 PBy/mo.

Observation, 1990-2008

- 38 months
- 57 months
- 53 months
- 40 months

Terabytes / month
ESnet traffic

HEP exp. data

ESnet capacity

Climate modeling data

ESnet traffic

HEP experiment data

ESnet capacity roadmap

Climate model data

All Four Data Series are Normalized to “1” at Jan. 1990

(HEP data courtesy of Harvey Newman, Caltech, and Richard Mount, SLAC. Climate data courtesy Dean Williams, LLNL, and the Earth Systems Grid Development Team.)
Science Drivers:

• LHC is expected to send ~50Gb/s to the two US Tier 1 Data Centers within two years. This data is then redistributed to the US Tier2 Centers (mostly at universities), effectively doubling the bandwidth requirements for ESnet.
  – In the event of outages on some paths, others will have to take up at least some of the stranded traffic. This motivates 100G connections between the ESnet core and the Tier 1 Centers and then to Internet2 for university access.

• SC Supercomputer Centers at NERSC, ORNL, and ANL for Climate simulation

• Others like Bio-Energy centers – Berkeley, Ann Arbor, Madison

Technology Planning:

• ESnet4 planning assumes technology advances will provide 100 Gb/s optical waves (they are 10 Gb/s now)

• The ESnet4 SDN switching/routing platform (Juniper MX960) is designed to support new 100 Gb/s network interfaces

• With capacity planning based on the ESnet 2010 wave count, we can probably assume some fraction of the core network capacity by 2012 will require 100 Gb/s interfaces

ESnet is involved in a collaboration with Internet2, Juniper Networks (core routers), Infinera (DWDM), and Level3 (network support) to accelerate its deployment and help drive down the cost of 100G components
ESnet4 Update
Network Services
• OSCARS
• perfSONAR
• Federated Trust
IP Network:
• Best-effort routing is simplistic, opportunistic, and resilient. However it provides no assurances, consistency, or predictability.

Science Data Network:
• OSCARS provides predictable and specific network service performance that applications can demand and will be reliably provided through automated agents.
• The network is no longer just a cloud over which the user has little or no control, but a cyber-resource that can and should be directed by the application just as other resources such as compute/server cycles, storage resources, or workflow scheduling.

perfSONAR end-to-end monitoring service:
• Providing useful, comprehensive, and meaningful information on the state of end-to-end paths.
• Supports regularly scheduled tests & archiving of results, acting as an intermediate layer between the performance measurement tools and the diagnostic or visualization applications.
OSCARS service requirements:

- Guaranteed bandwidth with resiliency
  - User specified bandwidth - requested and managed in a Web Services framework
  - Explicit backup paths can be requested
- Traffic isolation
  - Allows for high-performance, non-standard transport mechanisms that cannot co-exist with commodity TCP-based transport
- Traffic engineering (for ESnet operations)
  - Enables the engineering of explicit paths to meet specific requirements
    - e.g. bypass congested links; using higher bandwidth, lower latency paths; etc.
- Secure connections
  - The circuits are “secure” to the edges of the network (the site boundary) because they are managed by the control plane of the network which is highly secure and isolated from general traffic
- End-to-end, cross-domain connections between Labs and collaborating institutions
Phase 1
- Proof of concept
- Intra-domain virtual circuit (VC) services

Phase 2
- Inter-domain interoperability
- Pre-production ESnet VC services

Phase 3
- Productionalizing OSCARS (IDC)

Phase 4
- Extending service offerings
- Framework for research
Community approach to supporting end-to-end virtual circuits in the R&E environment is coordinated by the DICE (Dante, Internet2, Caltech, ESnet) working group

- Each organization potentially has their own InterDomain Controller approach (though the ESnet/Internet2 OSCARS code base is used by several organizations (flagged OSCARS/DCN)
- The DICE group has developed a standardized InterDomain Control Protocol (IDCP) for specifying the set up of segments of end-to-end VCs
  - While there are several very different InterDomain Controller implementations, they all speak IDCP and support compatible data plane connections
- The following organizations have implemented/deployed systems which are compatible with the DICE IDCP:
  - Internet2 Dynamic Circuit Network (OSCARS/DCN) – LHCNet (OSCARS/DCN)
  - ESNet Science Data Network (OSCARS/SDN) – LEARN (Texas RON) (OSCARS/DCN)
  - GÉANT2 AutoBahn System – LONI (OSCARS/DCN)
  - Nortel (via a wrapper on top of their commercial DRAC System) – Northrop Grumman (OSCARS/DCN)
  - Surfnet (via use of above Nortel solution) (OSCARS/DCN) – Nysernet (New York RON)
  - University of Amsterdam (OSCARS/DCN) – DRAGON (U. Maryland/MAX) Network
- The following "higher level service applications" have adapted their existing systems to communicate via the user request side of the IDCP:
  - LambdaStation (FermiLab)
  - TeraPaths (Brookhaven)
  - Phoebus (UMd)
OSCARS 0.5 Architecture (1Q09)

- **Tomcat**
- **Web Service Interface**
- **Web Browser User Interface**
- **Web Service Interface**
- **OSCARS**
  - **RMI**
  - **Core**
    - Resource Scheduler
    - Path Computation Eng
    - Path Setup Modules
  - **BSS DB**
- **AAA**
  - **RMI**
  - **Core**
    - Manage Subscriptions
    - Forward Notifications
  - **AAA DB**
- **Notification Broker**
  - **RMI**
  - **Core**
    - Authentication
    - Authorization
    - Auditing
  - **Notify DB**
Modular PCE Function
Modifications required by FNAL and BNL

- Changed the reservation workflow, added a notification callback system, and added some parameters to the OSCARS API to improve interoperability with automated provisioning agents such as LambdaStation, Terapaths and Phoebus.

Operational VC support

- As of 12/2/08, there were 16 long-term production VCs instantiated, all of which support HEP
  - 4 VCs terminate at BNL
  - 2 VCs support LHC T0-T1 (primary and backup)
  - 12 VCs terminate at FNAL
  - 2 VCs support LHC T0-T1 (primary and backup)
  - For BNL and FNAL LHC T0-T1 VCs, except for the ESnet PE router at BNL (bnl-mr1.es.net) and FNAL (fnal-mr1-es.net), there are no other common nodes (router), ports (interfaces), or links between the primary and backup VC.

- Short-term dynamic VCs
  - Between 1/1/08 and 12/2/08, there were roughly 2650 successful HEP centric VCs reservations
    - 1950 reservations initiated by BNL using Terapaths
    - 1700 reservations initiated by FNAL using LambdaStation
OSCARS is a Production Service

OSCARS generated and managed virtual circuits at FNAL – one of the US LHC Tier 1 data centers. This circuit map (minus the yellow callouts that explain the diagram) is automatically generated by an OSCARS tool and assists the connected sites with keeping track of what circuits exist and where they terminate.
Spectrum Now Monitors OSCARS Circuits
Tools in the perfSONAR software suite:

- SNMP Measurement Archive
- Lookup Service
- Topology Service
- Circuit Status Measurement Archive
- Status Measurement Archive
- perfSONAR-BUOY
- PingER Services

Visualization

- Allow ESnet user community to better understand our network & its capabilities.
- Allow ESnet users to understand how their use impacts the backbone.

Alarming

- Automated analysis of regularly scheduled measurements to raise alerts.
Currently deploying the hardware across the network to support adhoc measurements for debugging

- OWAMP servers
- BWCTL servers
- Topology service
- Utilization service

**perfSONAR Buoy deployment**
- Between ESnet systems
- To Internet2 & GEANT
- To/From ESnet sites

**Hardens the infrastructure**
- Continuous monitoring of servers & services
- Centralized management of OS & services configuration
- Performance tuning & verifying everything is working as designed
perfSONAR R&D Activities

Scaling & robustness enhancements

Visualization tools

• Single domain tools
  – Utilization browser
  – Topology browser
  – Latency & bandwidth browser

• Advanced tools
  – Looking across multiple domains
  – Looking at correlations between different types of measurements
  – Application or user community specific views

Integrating OSCARS circuits

• Topology
• Utilization
• Active measurements across them

Alarming
DOEGrids Certification Authority

- New Logo and ID Mark
- Operations
  - Vista – IE browser support in development
    - Also beginning testing IE 8 browser
  - ESnet 2-factor
    - Support ESnet 2-factor authentication token project
    - Add ESnet RA to list of official RAs in DOEGrids CA
  - Cloning and Geographical Dispersion
DOEGrids CA (one of several CAs) Usage Statistics

User Certificates: 9259
Total No. of Revoked Certificates: 2056
Host & Service Certificates: 21043
Total No. of Expired Certificates: 19452
Total No. of Requests: 35629
Total No. of Certificates Issued: 30331
Total No. of Active Certificates: 8823
ESnet SSL Server CA Certificates: 50
FusionGRID CA certificates: 113

* Report as of Jan 29, 2009
DOEGrids CA (Active Certificates) Usage Statistics

* Report as of Jan 29, 2009
Active DOEGrids CA Breakdown

**DOEGrids CA Statistics (8823)**

- OSG: 67.57%
- FNAL: 24.92%
- LCG: 0.83%
- PNNL: 0.01%
- ORNL: 0.60%
- NERSC: 1.29%
- LBNL: 0.73%
- FusionGRID: 0.17%
- ESnet: 1.67%
- ESG: 0.67%
- ANL: 1.53%

**OSG** Includes (BNL, CDF, CIGI, CMS, CompBioGrid, DES, DOSAR, DZero, Engage, Fermilab, fMRI, GADU, geant4, GLOW, GPN, GRASE, GridEx, GUGrid, i2u2, ILC, JLAB, LIGO, mariachi, MIS, nanoHUB, NWICG, NYSGrid, OSG, OSGEDU, SBGrid, SDSS, SLAC, STAR & USATLAS)
DOEGrids CA and its key management hardware will be cloned and dispersed around the US

• Improve Continuity of Operations and disaster recovery issues (ESnet requirements)
• Improve availability to customers
• Provision for future, robust services
• Current status: Testing and configuration of netHSM hardware, and project planning
Advances in security at ESnet over the last 6 months:

• Implemented Two-factor authentication for ESnet network engineers requesting privileged access to the network management plane. Reviewed and re-defined access to network management plane.

• Upgraded Bro Intrusion Detection System

**ESnet Security Peer Review – Feb 11-12**

• Fed/R&E/Commercial experts reviewed ESnet security practices and procedures

**Disaster recovery improvements**

• Deployed Government Emergency Telecommunications Service (GETS) numbers to key personnel

• Deploying full replication of the NOC databases and servers and Science Services databases in the NYC Qwest carrier hub
Goals

• Better organization of information, easier navigation, searchable (not everything in pdfs) but don’t want it to all be ‘push’

• Collaborative tool – upload best practices, video from conference, community calendar, staff pages

• Integration of business processes into site
  – “My ESnet” portal for site coordinators / users
  – Exploring Google Earth or similar network visualization
    • IP / SDN / MAN representation
    • perfSONAR performance data
    • OSCARS virtual circuit status
  – Looking for ideas/input/suggestions.