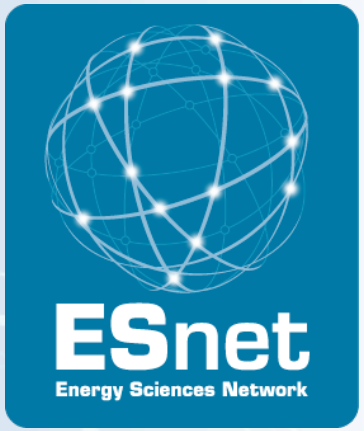


ESnet Update

Steve Cotter, Dept Head

Lawrence Berkeley National Lab

Winter 2011 Joint Techs

Clemson, SC

Feb 2, 2011

@ESnet: It's all about the Science

- More bandwidth to DOE facilities and Labs at lower costs
- Richer network services in support of distributed science
  - Develop 'network aware' integrated services that deliver 'end-to-end' high-performance data transfer, HPC/cloud computing, and science collaborative services
- Carrier-class network operations providing high network availability to all DOE facilities
  - Seamless network interoperability across multiple network domains
- Develop and deploy energy-aware and efficient networking infrastructure
- Provide a networking research testbed for DOE community
  - Conduct/enable groundbreaking research in new protocols/storage/energy efficient networking

ESnet4 Network

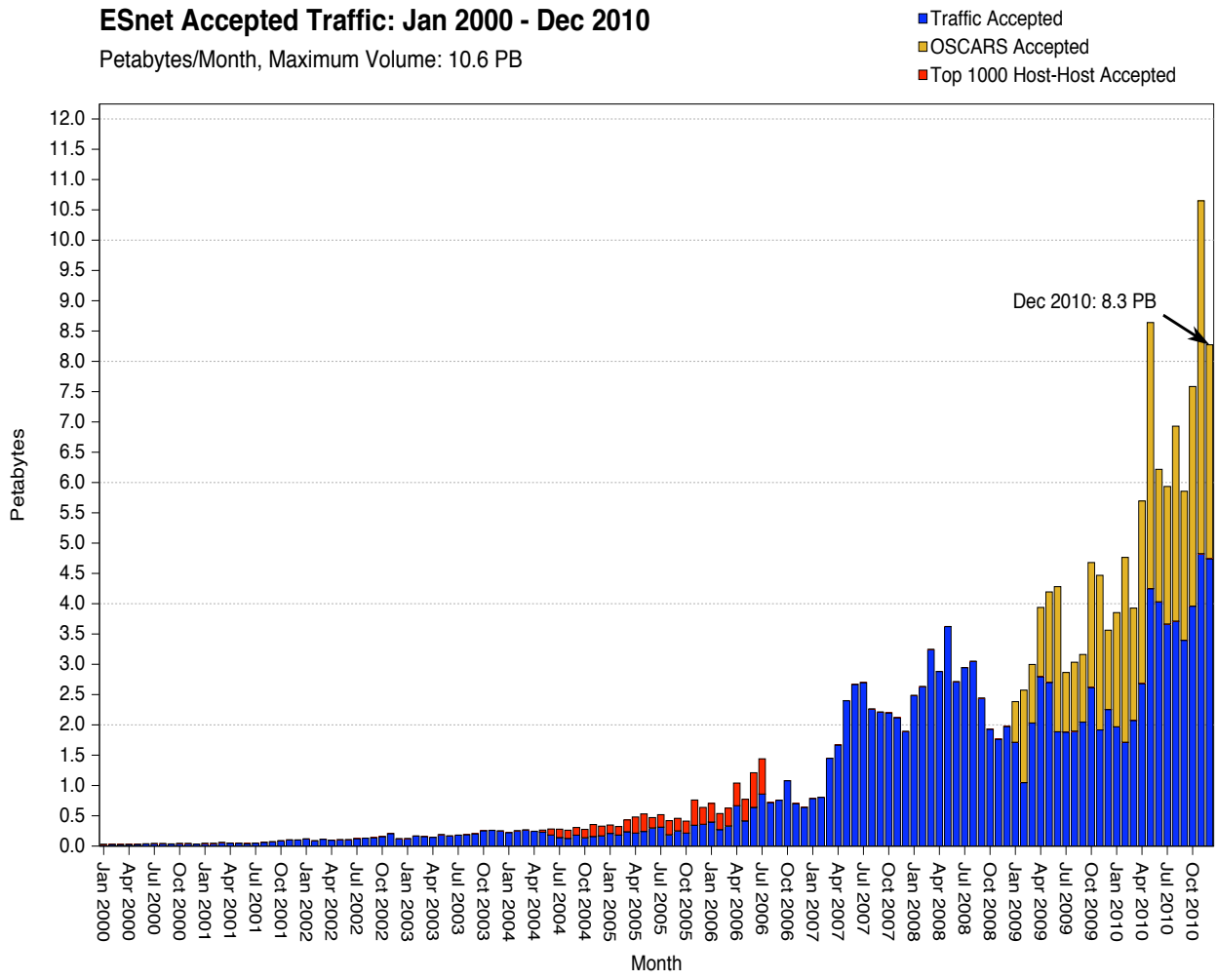
Peaked at 10.6 PB in Nov

>1PB of genomics traffic between JGI – NERSC that month alone

More than 50% continue to go over OSCARS circuits

Instrumenting with perfSONAR all connections 1G and up

99.9983% network availability to Office of Science labs



Long Island MAN

Argh! No more snow! It's impossible to find manholes under 15' snow

Light between 111 8th & 32 AofA waiting for portions of the ring to be constructed

- Southern: (was) First week of Feb
- Northern: Mid-March

Will handle testbed & production traffic (eventually) using Layer 1 VPN capability

Dramatically reduces our costs to reach BNL

OSCARS Development Status

PCE SDK (v01.11.11) is done

- SDK for development of PCEs for use in OSCASR (v0.6)
- ESnet has used the PCE SDK to develop bandwidth and VLAN selection PCEs
- Several DOE funded projects have identified interest in using the SDK
- PCE SDK is available for downloads to beta-user (contact chin@es.net)

OSCARS (v0.6 native) is close to code complete

- ~120,000 lines of code now (~130K when complete)
- Undergoing testing right now
- OSCARS v0.6 native (i.e. communicates with other OSCARS v0.6 servers) should be in beta in February
- Backward compatibility with v0.5.x is on target to be completed by April

OSCARS Deployment Status

OSCARS is now one of the core applications deployed by SCinet to support dynamic circuits

- First deployed in SCinet in SC09
- For SC10 over 150 dynamic circuits were configured in support of roughly 15 demonstrations

First deployments of OSCARS (v0.6) will be in 2Q2011

- Green field installations as part of I2's DYNES project

Proven use cases

- IaaS: extending site LAN over the WAN (Genomics: JGI/NERSC)
- Traffic engineering to avoid congestion points (Cloud testing: LBNL/Google)
- Guaranteeing network resources, including failover (LHC T0-T1, T1-T2)
- Deadline scheduling (Fusion DIII-D: GA/NERSC)

Fermi Challenge

The general ESnet philosophy has been:

- Move large and predictable flows to circuits over SDN
- 'Everything else' should fit in the 10G IP backbone

Issue:

- Fermi already using 15 OSCARS circuits to move bulk of traffic
- The remaining IP traffic was still approaching 10G from time to time.
  - Could not move it to Layer 2 OSCARS circuits because that requires a willing and able partner at the far end of the circuit!
  - Could not move it to Layer 3 OSCARS circuits because OSCARS Layer 3 circuits start and terminate on ESnet routers, and the bottleneck was on the 10GE link to ESnet on the FERMI egress router

Fermi Challenge

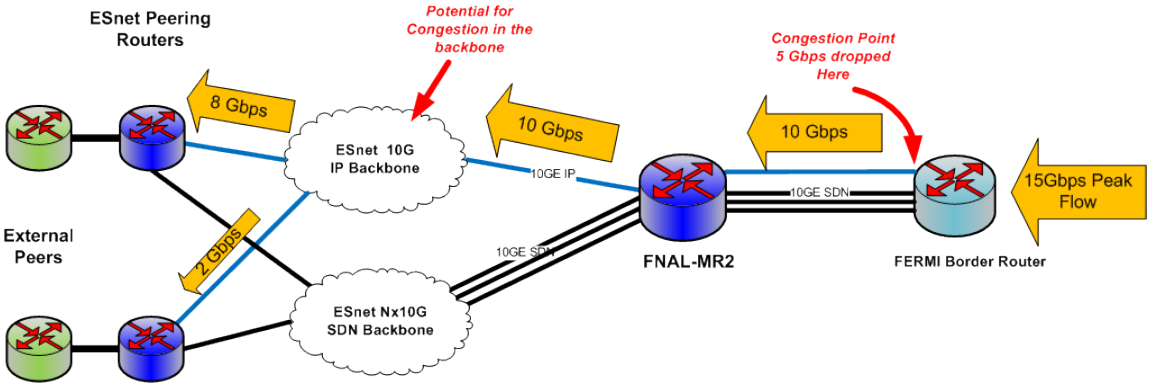
Moved FERMI 'IP' DMZ to a 2x10GE aggregated connection

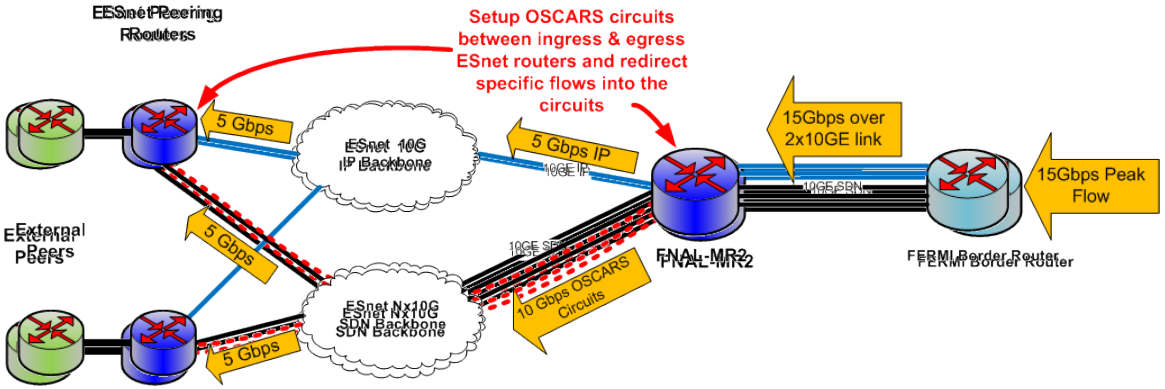
Deploy a minor enhancement to OSCARS

- Allow Layer 3 circuits to dump packets into the standard forwarding table if the MPLS LSP supporting the circuit fails.

Re-route several large FERMI 'IP' flows onto SDN across the backbone

- Note that re-routing a flow into a circuit is hit less!





Advanced Networking Initiative

$62M ARRA funded 100G prototype network and testbed

- Leveraging this infusion of funds to lower long-term networking costs
  - Dark fiber in metro areas
  - Scalable, cost-effective 100G network
  - National dark fiber testbed
- Testbed already in use
  - Table-top being used by first round of research projects
  - Disassembled and shipped to Long Island as LI MAN fiber build completes
  - Second round of research projects selected by Advisory Committee comprising of R&E, lab, and commercial members

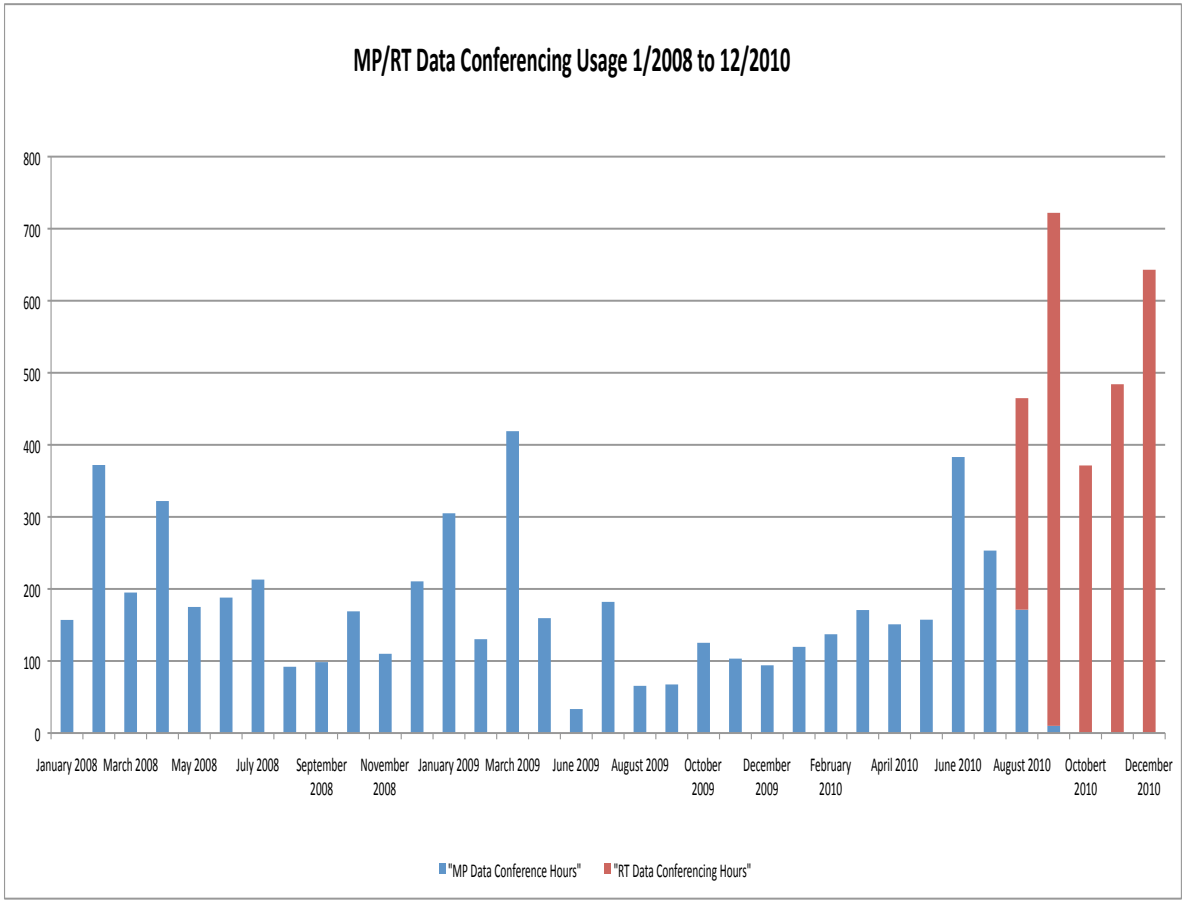
ESnet Collaboration Services

Transitioned to commercial audio/data conferencing system

- Improved capabilities for less than half the cost
  - Support for international calls
- Greater utilization of service
  - ~600 accounts, ~200 active

Testing new video conferencing service for late Feb launch

- Physically diverse, redundant Tandberg MSE 8000s
- 80 SD, 20 HD ports each
- Telepresence capable



New ESnet Website

More engaging, relevant content

- Easier/cheaper to maintain

Platform for more efficient/effective communications with our community

‘My ESnet’ portal

- Leveraging federated identity: InCommon/OpenID enabled
- Customizable ‘widgets’
- Access to network tools/ maintenance calendar/stats/etc.

Go Live – early March

- Moving content, finalizing information architecture with contractor

Check out fasterdata.es.net

New Ticketing System

- Moving to Service-Now.com SaaS (Software-as-a-Service) solution for ticketing
  - Replacing legacy (circa 1995!) ticketing system
- Same vendor as UC Berkeley, NERSC, LBNL IT Division
- Benefits: reduced support costs, easier to open/query/update tickets
- On our roadmap: integration with new website
  - ticketing portal + knowledge base for the entire community
- We're interested in your experience: have you adopted a new ticketing system recently?
  - please share your experience with greg@es.net

Science Identity Federation (SIF)

Focus: Get labs into InCommon

- Blanket agreement in place
  - 3 year memberships in InCommon for DOE labs, sites, user facilities
  - Primary: SC sites - but there should be sufficient funding for some non-SC
- Sites sign a participation agreement with InCommon
- Berkeley Lab will manage the procurement (billing)
- Agreement web sites will be available soon (maybe this week)

Future focus: applications; interoperability with DOE – ICAM

(Contact Mike Helm helm@fionn.es.net for more information)

Network Analytics

In-house developed system built using open-source tools in 2004

2 new commercial systems deployed

- Arbor Networks & Packet Design

Uses:

- BGP analytics
- Peerings
  - What if we added a new peering location?
  - Is it worth peering directly? Impact on traffic?
- Looking at cloud services
- Security: Is that really a DDoS or just 'slashdot effect'?

Experimental Capacity Planning Tool

Background: Complex network, variable traffic patterns make it difficult to determine when is a link saturated

- There is no widely held idea of what percentage of utilization indicates that a circuit is becoming saturated. A thread on NANOG has ranges from 40% to 95% http://nanog.markmail.org/thread/vmkjrldzvsrbz3en

Approach: Try to take some of the 'art' out of deciding on when to upgrade

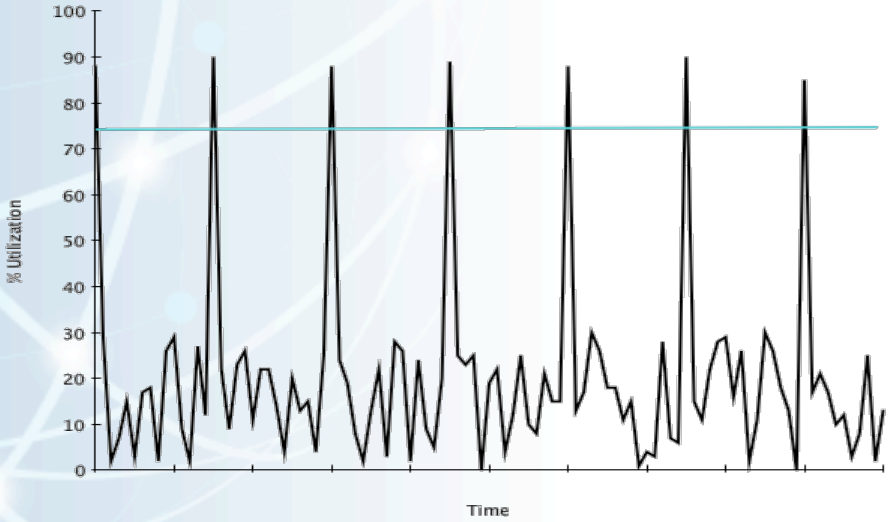
- Define metrics, analyze circuits and assign a score for each metric, set thresholds

Defining Metrics

Peak Metric

The number of sampling intervals where the utilization is over some threshold T. In this example:

- T = 75%
- The value of the metric is 7 (there is a peak at time 0)

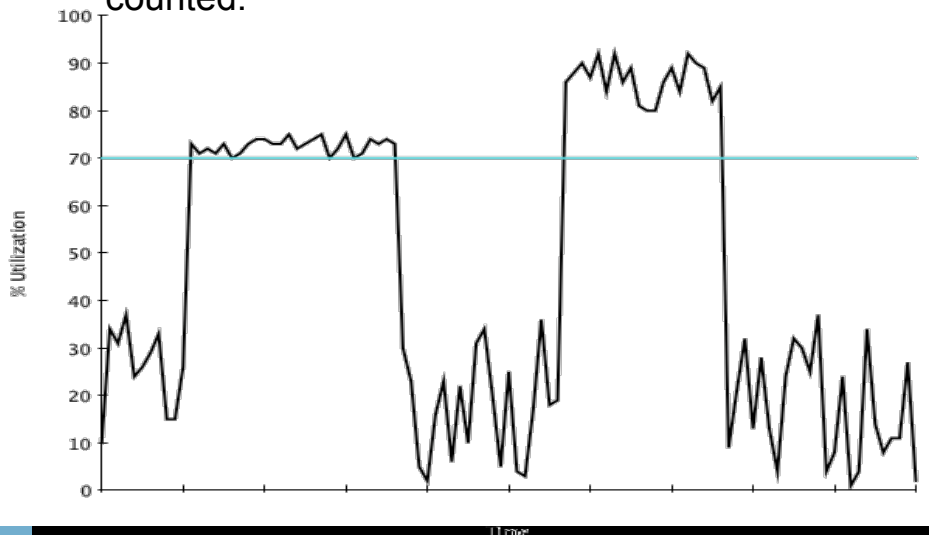


Plateau Metric

The number of sampling intervals where the utilization is over some threshold T for at least t intervals. In this example:

T is 70%, t is 20

Metric value is 48 (left plateau is 27 time intervals wide, the right is 21 time intervals wide). Peaks which persist for less than t intervals are NOT counted.



Current Status And Limitations

Status

Peaks and Plateaus metrics are both implemented

In the process of finding appropriate thresholds for each metric

Need more data regarding past upgrades to feed the logistic regression

The peaks and plateaus metrics when properly tuned may provide enough filtering to allow a human to evaluate the remaining circuits

Limitations of the regression

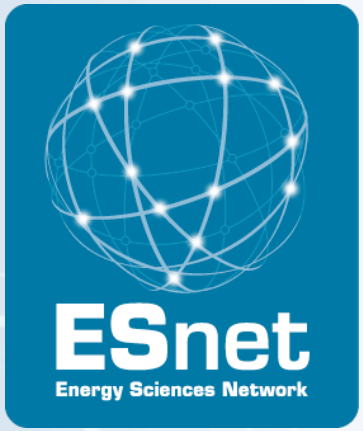
- Topology (both physical and routing) changes can complicate things
- The past may or may not be a good predictor of the future
- Additional metrics would be helpful

Contact Jon Dugan (jdugan@es.net) with questions, comments or suggestions!

How do we become the Dept of (Less) Energy?

Ideas we're contemplating:

- Begin to understand how to measure the energy cost of moving data using existing infrastructure
  - Extend to 100G network
- Expand http://weathermap.es.net/ to include energy costs
- Test movement of VMs in terms of energy costs
- What is more energy efficient – computing in the cloud or locally? How can this decision be made?
- Use path computation engine SDK to create virtual networks on demand that satisfy not just BW allocation costs but also min energy cost, or max green energy costs
- Analyze the idleness of the paths – is there a chance for power management there? Does it make sense to purposefully create idle paths at a slight performance cost?
- Switch to move flows dynamically to create idleness along a path



Twitter: ESnetUpdates

Blog: http://esnetupdates.wordpress.com/

Thank you