ESnet
Energy Sciences Network

ESnet Services and Service Level Descriptions

Joseph Burrescia
Michael Collins
William Johnston
ESnet Staff

July 17, 2009
Version 4.0

Berkeley Lab

U.S. Department of Energy
Office of Science

Contents

Summary

Within the DOE Office of Science (SC) ESnet's[a] mission is to provide interoperable, effective, reliable, and high performance network communications infrastructure, and certain collaboration services, in support of SC's large-scale, collaborative science.

In summary, the services ESnet provides includes:

- Comprehensive physical and logical connectivity
    - High bandwidth access to DOE sites and DOE's primary science collaborators: Research and Education institutions in the US, Europe, Asia Pacific, and elsewhere
    - Full access to the global Internet for DOE Labs
- A full suite of network services
    - IPv4 and IPv6 routing and address space management
    - IPv4 multicast
    - Guaranteed bandwidth and virtual circuit services
    - Scavenger service so that certain types of bulk traffic will use all available bandwidth, but will give priority to all other traffic when it shows up
    - Reservable, guaranteed bandwidth using the ESnet On-demand Secure Circuits and Advance Reservation System (OSCARS)[b]

[a] ESnet is operated by Lawrence Berkeley National Lab for the U. S. Dept. of Energy, Office of Science.

[b] Available for Office of Science sites only.

- An architecture tailored to accommodate DOE's large-scale science
    - Move huge amounts of data between a small number of sites that are scattered all over the world
- Comprehensive user support, including "owning" all trouble tickets involving ESnet users (including problems at the far end of an ESnet connection) until they are resolved – 24x7x365 coverage
    - ESnet's mission is to enable the network based aspects of SC science, and that includes troubleshooting network problems wherever they occur
- Comprehensive monitoring and measuring of network resources
- Cybersecurity in the WAN environment
- An architecture and operations approach that is fault tolerant and supports a disaster recovery strategy for non-stop operation.
- Collaboration services and Grid middleware supporting collaborative science
    - Federated trust services with science oriented policy
    - Audio, video, and data conferencing
- A highly collaborative and interactive relationship with the DOE Labs and scientists for planning, configuration, and operation of the network
    - ESnet and its services evolve continuously in direct response to SC science needs

ESnet Services and Service Level Document Evolution

This document discusses the current services offered by ESnet. As such, it is constantly evolving as new services are offered. Sections are periodically updated to reflect significant changes.

1 Document Conventions

This document details each service offered by ESnet and a corresponding Service Level Description in the form:

- **Service Description:** What service is provided to the ESnet user community.
- **Service Level:** What are the characteristics of the service? (What ESnet does and does not do, regarding: reliability, coverage, security, privacy, etc.)

2 Network Services and Service Level Descriptions

2.1 Network Layer Connectivity

2.1.1 Service

ESnet provides comprehensive and high quality network access to Office of Science Labs and other DOE sites through a national circuit infrastructure that directly connects approx. 40 sites to a high-speed, national core that , through peering and routing, provides ESnet Sites global Internet access with redundant paths to most major networks.

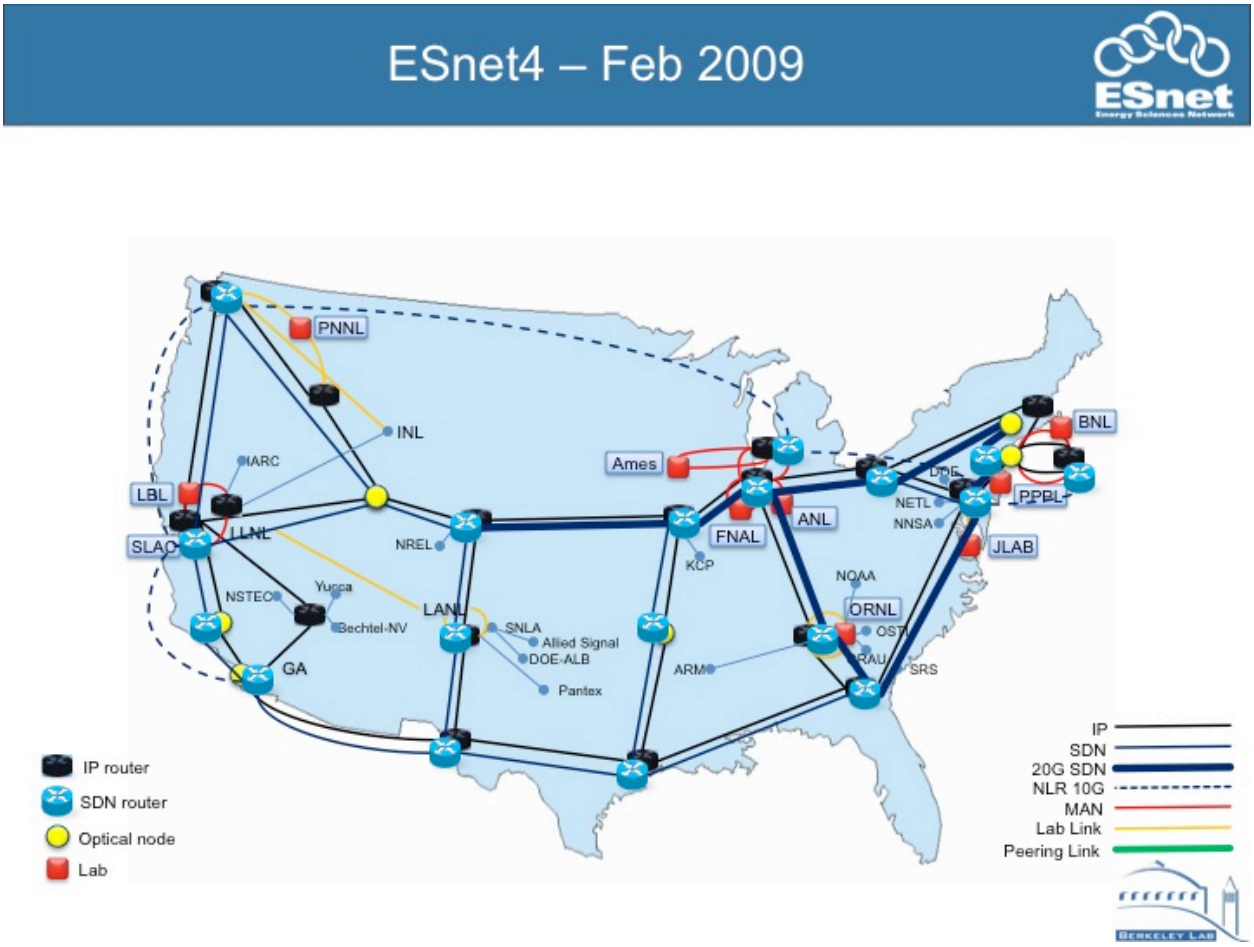

Figure 2-1 ESnet Network Layer Connectivity

ESnet carries the full set of Internet routes, allowing sites to use ESnet as their only Internet connection. The full set of network services ESnet provides is enumerated below.

2.1.2 Service Level

Unless noted, the following characteristics apply to all Network Layer Connectivity Services.

Latency

The maximum round trip latency between any two ESnet sites is less than 100ms. Latency in the backbone is continuously monitored.

Packet Loss

ESnet uses several criteria to accept new circuits. Recently, more formal methods have been used to determine acceptable loss rates based on a standard. For example, acceptance testing of the all MAN and backbone circuits were based on a loss rate based on a theoretical calculation of the maximum loss permitting an acceptable FTP transfer rate (see M. Mathis, J. Semke, J.

Mahdavi, T. Ott, "The Macroscopic Behavior of the TCP Congestion Avoidance Algorithm", Computer Communication Review, volume 27, number 3, pp. 67-82, July 1997).

Appendix A is the current ESnet Circuit Acceptance Criteria.

Throughput

The circuit shall be saturated with as demonstrated bandwidth over 95% of the link capacity for 5 minutes.

Appendix A is the current ESnet Circuit Acceptance Criteria.

Availability

ESnet is operated as a high-availability network on a 24hrs/day x 365 days/year basis. High-reliability and high-operational security are provided in both the network and in the ESnet infrastructure support – the systems that support the operation and management of the network and services - to ensure the continuous operation of the network. Availability is continuously monitored.

ESnet strives for 99.9% availability to sites, the actual availability, typically much greater, is monitored and posted on the Web. Please refer to http://calendar.es.net.

Reliability

Reliability involves the reliability of the IP Core network and the reliability of site connections to the IP Core

- ESnet's multiple ring backbone topology insures that no single backbone circuit failure will cause an outage to a site. The internal routing protocols are configured to switch to a backup path within 2 seconds upon determining a backbone link has failed. ESnet has deployed its Science Data Network (SDN), which provides an independent backup to the current IP backbone. The SDN provides additional redundancy to insure maximum availability of network connectivity.

  Connections to a site (either to an ESnet router or to a Site router) are terminated in a Core Router. The core routers are industry standard systems housed in commercial telecom quality space with generator backed up power. The routers themselves are fully redundant with dual independent switching fabrics, dual route processors, and dual power supplies. The maintenance contract on these routers is next day response. ESnet has "smart hands" contracts at each core location to assist with module replacement etc.

- The ESnet router at a Site is attached to the core via a local loop. The standard level of service reliability for ESnet equipment located at a Site is generally that of parts replacement on a next business day basis, with the labor performed by the site. Higher than standard service reliability can be provided, if paid for by the site i.e. through the purchase an extended maintenance contract with the equipment vendor. If a local loop circuit fails any site access supported by the local loop will be unavailable until the loop is repaired.

- Sites attached via MAN. Sites in the San Francisco Bay area, Chicago, and Long Island are connected to the backbone via a Metropolitan Area Networks (MANs). MANs provide a ring topology offering the same protection against single circuit failures as the WAN. Routers on the MANs are also fully redundant, with dual supervisor cards, dual

power supplies, and separate line cards for the East and West ring interfaces. In addition, spare ports are available to switch to in case the primary interface to the site fails. Other MANs, or ring topologies, may be established as opportunities arise.

- Sites not attached to a MAN. Sites not within a region that cost effectively supports a MAN will be dually connected to the ESnet backbone as funding permits.
- ESnet equipment at sites. Depending on the sites needs, some sites may house ESnet equipment at their location. This equipment typically requires out-of-band management connectivity via a POTs, i.e. telephone, line. This connection enables remote ESnet personnel to access the ESnet equipment in the face of catastrophic primary circuit failure. It is up to the site to insure this connection is provided and to understand that restrictions placed on its ability to function (i.e. only manually enabling the circuit when ESnet requests access – perhaps done in response to a sites security rule) will hamper the ability to diagnose and repair ESnet equipment and adversely affect reliability.

Reservable, Guaranteed Bandwidth

The On-Demand Secure Circuits and Advance Reservation System (OSCARS) project provides a dynamically provisioned, guaranteed bandwidth secure circuits both within ESnet, and between ESnet and other network domains. OSCARS' circuits are mainly configured for large research data flows on the SDN network. The project is funded out of the DOE Office of Science. For current status of OSCARS development see http://www.es.net/OSCARS.

Privacy

ESnet is a federally funded network and DOE has the authority to request monitoring in order to assure that the use of the network is consistent with federal rules. While the architecture of ESnet makes unauthorized monitoring unlikely, ESnet does not take special care to protect against this. Users should have no expectation of privacy when using the default network. If privacy is required, then users can, and should, set up software or hardware end-to-end VPNs.

CALEA

ESnet is CALEA compliant. ESnet must and will respond to any legitimate warrant issued to it by law enforcement for data.

Release of Customer Data

ESnet only collects user data for statistical use on traffic patterns and diagnosing network problems. ESnet may also incidentally collect data in the form of logs from DNS zones ESnet provides secondary service for. ESnet will only release any of this data to the site coordinator whose site generated the data at the site coordinator's explicit request.

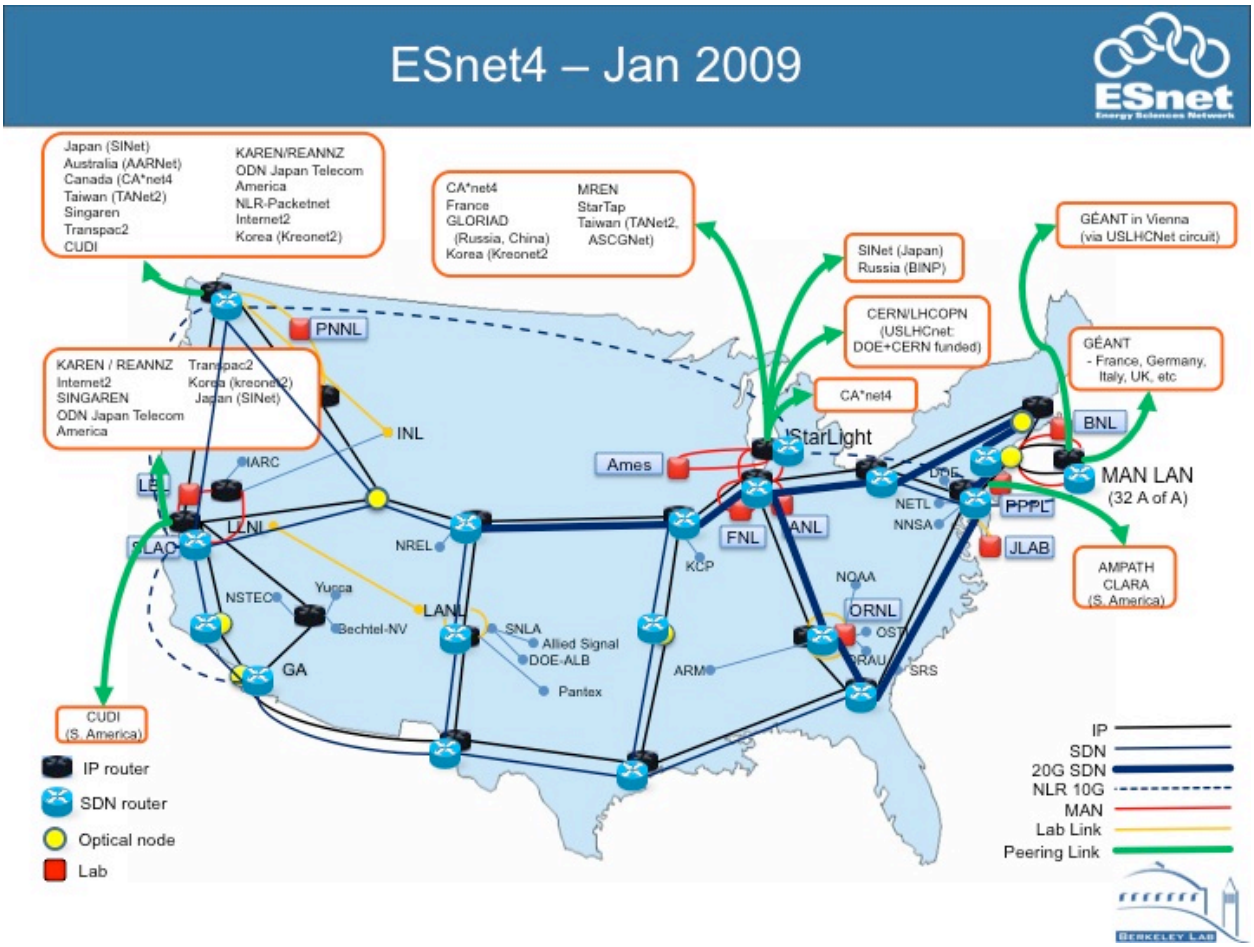

Figure 2-2 Peering Defines ESnet's Logical Infrastructure that Connects the DOE Community with its Collaborators

2.2 IPv4 Services

ESnet provides all IPv4 services necessary for ESnet sites to fully participate in the Global Internet. Please refer to Figure 2-2.

2.2.1 Routing and Peering With Commercial Networks

2.2.1.1 Service

ESnet is a Tier 1 network that provides DOE scientists access to all Internet sites by managing a full complement of Global Internet routes (about 283,000 IPv4 routes filtered from about 660,000 from 271 peers) at 17 general peering points. ESnet is present at the commercial peering points referenced in Table 2-1 ESnet Commercial Peering Points

2.2.1.2 Service Level:

ESnet's connection to the global Internet is resilient and provides optimal routes by decentralizing commercial peering both by location and multiplicity of peers. ESnet's routing infrastructure selects optimal routes and filters sub-optimal routes, keeping them in reserve for

backup connectivity. ESnet personnel monitor the routes and, in case of problems, work with all tiers of ISPs to resolve issues.

The vast majority of commercial peering ESnet is involved with are settlement free, that is, ESnet does not pay an ISP for connectivity to their network. The reason for this is two-fold:

1) ESnet's has been peering with commercial networks since before the breakup of the original NSFnet and is a well known and respected ISP.

2) ESnet carries many Site networks that can only be reached by peering with ESnet.

EQUINIX	Ashburn VA
EQUINIX	Chicago IL
EQUINIX	San Jose CA
PAIX	Palo Alto CA

Table 2-1 ESnet Commercial Peering Points

2.2.2 Routing and Peering with Research and Education Networks

2.2.2.1 Service

ESnet provides sites with high-speed connectivity to the US, European, and Asia Pacific research and education (R&E) communities through multiple high-speed (1.0 -10 Gb/s) peering points (see Table 2-3) ESnet currently peers with the Research and Education Networks noted in Table 2-5Error! Reference source not found..

2.2.2.2 Service Level

The high-speed R&E peerings are continuously monitored for connectivity. To enhance redundancy and optimize connectivity, ESnet may peer with the same R&E network in multiple places.

Peering Point	Location	Speed
ALBUQUERQUE GigaPop (ABQG)	Albuquerque, NM	1G
CIC-OMNIPOP	Chicago, IL	10G
FRONT RANGE GigaPOP (FRGP)	Denver, CO	10G
ILLINOIS Inter CAMPUS COMMUNICATION NETWORK (ICCN)	Chicago, IL	10G
INDIANA GigaPOP (IU GigaPOP)	Chicago, IL	10G

MANHATTAN LANDING (MANLAN)	New York City, NY	10G
MID-ATLANTIC CROSSROADS (MAX)	College Park, MD	10G
NGIX-AMES	Sunnyvale, CA	10G
ALBUQUERQUE GigaPop (ABQG)	Albuquerque, NM	1G
CIC-OMNIPOP	Chicago, IL	10G
PACIFIC WAVE	Seattle WA	10G
PACIFIC WAVE	Sunnyvale CA, Los Angeles, CA	10G
PAIX PALO ALTO (PAIX-PA)	Palo Alto, CA	1G
SOUTHERN-CROSSROADS (SOX)	Atlanta, GA	10G
STARLIGHT	Chicago, CA	10G

Table 2-3 Research and Education Peering Points

AARNET	Australian Academic and Research Network (AARNet)
AMPATH	AMericasPATH
ASCC	Academia Sinica Computing Center
ASGC	Academia Sinica Grid Computing Center
CANARIE-CA*NET	Canarie/CA*net
CENIC	Corporation for Education Network Initiatives
CERN	CERN European Laboratory for Particle Physics
CLARA	Cooperación Latino Americana de Redes Avanzadas
CUDI	Corporación Universitaria para el Desarrollo de Internet A. C.
DREN	High Performance Computing Modernization Program
GEANT/DANTE	GEANT IP Service
GLORIAD	GLORIAD (Previously RBNET)
KEK-JAPAN	KEK High Energy Accelerators Research
KREONET2	KREONET2
INDIANAGIGAPOP	Indiana GIGAPOP
ICN	Illinois Century Network
INTERNET2	Internet2
LOSNETTOS	Los Nettos Regional Network
MREN	Metropolitan Research and Education Network
NASA-AMES	NASA Ames
MAX	Mid-Atlantic Crossroads Gigapop

NYSERNet	New York State Education and Research Network
NORTHWESTERN-UNIV	Northwestern University
NOX	Harvard University
NREN	NREN
PacketNet	National LambdaRail
PSC	Pittsburgh Supercomputing Center
PNW-GIGAPOP	PNW-GIGAPOP in SEATTLE
REANNZ	Research and Education Advanced Network New Zealan
ROUTE-VIEWS	Oregon Exchange
SDSC - SDNAP	San Diego Supercomputer Center
SINET	Science Information Network
SINGAREN	Singapore Advanced Research and Education Network
SOX	Southern Crossroads
STARTAP	Science,Technology & Research Transit Access Point
TANET-I2	Taiwan Network
TransPac2	
UIUC	University of Illinois at Urbana-Champaign
ULTRALIGHT	UltraLight
USLHCNET	American LHC Network
UWISC	University of Wisconsin
WISCNET	Wisconsin's education, research and public service network

Table 2-5 Research and Education Peering Partners

2.2.3 ESnet IPv4 Address and Space Management

2.2.3.1 Service

ESnet has been allocated IPv4 address space by the American Registry for Internet Numbers (ARIN). ARIN manages all Internet address space allocation in North America). ESnet manages its address space and uses it for allocating new address space to sites, the ESnet backbone and associated devices.

ESnet maintains redundant IPv4 Domain Name System (DNS) servers at LBL, LVMR-NOC, AOFA-NOC . These servers are authoritative for the *es.net* domain and serve as secondary servers for several Sites. Anycast addressing is used so that requests are routed to the nearest server. The servers also contain pointers to Site managed DNS servers that resolve ESnet allocated addresses assigned to the site.

2.2.3.2 Service Level

ESnet maintains the IPv4 address registration in the ARIN whois database for all the ESnet address space. DNS servers and the domain names are registered by Network Solutions. All of the ESnet address space is registered with the International Routing Registry (IRR). ESnet also provides added support to sites by maintaining route registration in the RADB for site owned address space.

2.2.4 IPv4 Route Management

2.2.4.1 Service

ESnet manages routing tables in its routers to insure packets destined to a given location are correctly routed. OSPF and i-BGP are used as interior routing protocols and BGP is used as the exterior routing protocol.

2.2.4.2 Service Level

ESnet employs inbound and outbound Access Control Lists (ACLs), from both peers and Sites, to insure the quality of the ESnet routing tables. These filters are applied at both the packet level and the routing update level.

Packet filters block packets with invalid source and destination address from admission to the network. Additionally, Loose Reverse Path Forwarding is used to discard packets with invalid source addresses.

Route update filters are used to verify the BGP updates from each external BGP neighbor. These filters are used to prevent invalid routes advertised by a neighbor from entering the routing tables. In the cases where the neighbor announces too many routes to filter, a default filter, blocking well-know invalid routes, is used.

Site route and packet filters are relatively static and are changed manually at the request of the ESnet Site Coordinator. The filters for other peers are update routinely, by scripts, using information from the Internet Routing Registry, the peer and data from the ESnet border router.

The ESnet routers are periodically polled for routing information and any anomalies are announced, via e-mail, to the engineering staff.

BGP sessions are configured to send state transition traps to the ESnet Network Management System. The traps signal that an external peer of ESnet has either gone down or has become once again responsive and the operations staff is notified.

2.2.5 IPv4 Multicast

2.2.5.1 Service

Although the IPv4 multicast protocols are designated “experimental” by the Internet Engineering Task Force, ESnet provides full IPv4 multicast service for Any Source Multicast (ASM) and Single Source Multicast (SSM).

ESnet enables Sparse-Mode Protocol Independent Multicast (PIM-SM) on all production internal links. Many ESnet sites are multicast enabled and connect to the ESnet Multicast Service via a combination of MBGP (Multi-protocol Border Gateway Protocol), PIM-SM, and MSDP (Multicast Source Discovery Protocol). ESnet connects to a number of external multicast enabled networks at peering points via the same protocol set.

ESnet participates in commercial multicast testbeds at the Equinix peering facilities.

More information is available at www.es.net/multicast.

2.2.5.2 Service Level

ESnet has architected a redundant or Logical RP (Rendezvous Point) infrastructure for enhanced reliability. A multicast beacon is available to help in the diagnosis of connectivity problems when they arise.

2.3 IPv6 Services

ESnet provides all IPv6 services necessary for ESnet sites to fully participate in the Global Internet.

ESnet runs IPv6 native in the core backbone. ISIS and i-BGP are used as interior routing protocols and BGP is used as the exterior routing protocol. The ESnet's Email, DNS, and Web services are IPv6 capable. IPv6 service is provided to any ESnet site requesting IPv6 connectivity.

An ESnet Site Coordinator can request an IPv6 prefix (block of addresses) from ESnet's allocation of IPv6 address space. The IPv6 offering is run as a production service and ESnet accepts and pursues trouble calls related to IPv6.

More information is available at the ESnet web site.

2.3.1 Routing and Peering With Commercial Networks

2.3.1.1 Service

ESnet is a Tier 1 network that provides DOE scientists access to all Internet sites by managing a full complement of Global Internet routes (currently about 1120 IPv6 prefixes) filtered from 44 peers at 11 general peering points. ESnet is present at the commercial peering points referenced in Table 2-7 ESnet IPv6 Commercial Peering Points.

2.3.1.2 Service Level

ESnet's connection to the global Internet is resilient and provides optimal routes by decentralizing commercial peering both by location and multiplicity of peers. ESnet's routing infrastructure selects optimal routes and filters sub-optimal routes, keeping them in reserve for backup connectivity. ESnet personnel monitor the routes and, in case of problems, work with all tiers of ISPs to resolve issues.

The vast majority of ESnet's commercial peerings are settlement free. ESnet does not pay an ISP for default. The reason for this is two-fold:

1) ESnet's has been peering with commercial networks since before the breakup of the original NSFnet and is a well-known and respected ISP.

2) ESnet carries many Site networks that can only be reached by peering with ESnet.

EQUINIX	Ashburn VA
EQUINIX	San Jose CA
MAE-WEST ATM	Sunnyvale CA
MANLAN	New York City, NY
PAIX	Palo Alto CA

Table 2-7 ESnet IPv6 Commercial Peering Points

2.3.2 Routing and Peering with Research and Education Networks

2.3.2.1 Service

ESnet provides sites with high-speed IPv6 connectivity to the US, European, and Pacific Rim research and education (R&E) communities through multiple high-speed (2.5-10 Gb/s) peering points (see Table 2-9 Research and Education IPv6 Peering Points). ESnet currently peers with the Research and Education Networks noted in Table 2-11 Research and Education Ipv6 Peering Partners.)

2.3.2.2 Service Level

The high-speed R&E Ipv6 peerings are continuously monitored for connectivity. ESnet's Network Monitoring System is being upgraded to properly parse IPv6 address in traps, in order to generate an alarm if an IPv6 BGP peering goes down. To enhance redundancy and optimize connectivity, ESnet may peer with the same R&E network in multiple places.

Atlantic Wave	College Park MD
MANLAN	New York NY
MAX	College Park MD
NGIX-East	College Park, MD
NGIX-West	Sunnyvale, CA
Pacific Wave	Seattle WA
Pacific Wave	Sunnyvale CA
StarLight	Chicago IL
SDSC	San Diego CA

Table 2-9 Research and Education IPv6 Peering Points

AARNET	Australian Academic and Research Network (AARNet)

AMPATH	AMericasPATH
ASCC	Academia Sinica Computing Center
CANARIE-CA*NET	Canarie/CA*net
CENIC	Corporation for Education Network Initiatives
CERN	CERN European Laboratory for Particle Physics
CLARA	Cooperación Latino Americana de Redes Avanzadas
CUDI	Corporación Universitaria para el Desarrollo de Internet A. C.
DREN	High Performance Computing Modernization Program
GEANT/DANTE	GEANT IP Service
IIJ	Internet Initiative Japan
INTERNET2	Internet 2
KREONET2	KREONET2
Los Nettos	Los Angeles Regional Network
MAX-GIGAPOP	Mid-Atlantic Crossroads
MREN	Metropolitan Research and Education Network
NASA-AMES	NASA Ames
NREN	NREN
PacketNet	National LambdaRail
REANNZ	Research and Education Advanced Network New Zealan
SDSC - SDNAP	San Diego Supercomputer Center
TANET-I2	Taiwan Network
TransPac2	TransPac2
WIDE	WIDE
WISCNET	Wisconsin's education, research and public service network

Table 2-11 Research and Education Ipv6 Peering Partners

2.3.3 ESnet IPv6 Address and Space Management

2.3.3.1 Service

ESnet was allocated the first production IPv6 address space (2001:400::/32) assigned by the American Registry for Internet Numbers (ARIN). ARIN manages Internet address space allocation in North America. ESnet manages its address space and uses the space for assigning new address space to sites, the ESnet backbone and associated devices. Sites may also obtain address space from ARIN (see below).

ESnet maintains redundant IPv6 Domain Name System (DNS) servers at LBL, LVMR-NOC, AOFA-NOC. These servers are authoritative for the *es.net* domain and serve as secondary servers for several Sites. These servers also contain pointers to Site managed DNS servers that resolve ESnet allocated addresses assigned to the site.

2.3.3.2 Service Level

2.3.3.2.1 ESnet assigned IPv6 addresses

ESnet maintains the IPv6 address registration in the ARIN whois database for all the ESnet address space. DNS servers and the domain names are registered by Network Solutions. All of the ESnet address space is registered with the International Routing Registry (IRR).

ESnet also provides added support to sites by maintaining route registration in the RADB for site owned address space.

ESnet assigns IPv6 prefixes based on the ESnet IPv6 Production Addressing Plan.

2.3.3.2.2 ARIN assigned IPv6 addresses

At this time, ARIN makes two general classes of IPv6 addresses available: ISP Assigned, also known as Provider Aggregate-able (PA), and direct assignments also known as Provider Independent (PI). PA addresses are intended to meet the requirements of all end users except those with connections to multiple service providers. ESnet assigns PA address from its allocation.

ESnet Sites with more than a single service provider are referred to as "multihomed" and require either PI address space or special arrangements with one provider to announce that providers address space to a different provider. Many providers will not allow announcement of any of their address space to another provider.

To support the requirement for multihoming, current ARIN policy allows for an organization which qualifies for a direct IPv4 assignment to receive a PI IPv6 assignment. (See http://www.arin.net/index.shtml). These assignments are currently being made from 2001:500::/30. Any organization that qualifies for IPv4 address space may also qualify for an IPv6 assignment.

ESnet will accept the directly assigned prefixes from Sites (either ESnet supplied PA or a third party supplied PI prefixes) and announce them at all of its peerings. Although the acceptance of the announcements is a policy decision for each peer, ESnet does not anticipate a problem with the announcements being accepted.

2.3.4 IPv6 Route Management

2.3.4.1 Service

ESnet manages routing tables in its routers to insure packets destined to a given location are correctly routed. ISIS and i-BGP are used as interior routing protocols and BGP is used as the exterior routing protocol.

2.3.4.2 Service Level

ESnet employs inbound and outbound Access Control Lists (ACLs), from both peers and Sites, to insure the quality of the ESnet routing tables. These filters operate at both the packet level and the routing update level.

Packet filters block packets with invalid source and destination address from admission to the network. Additionally, Loose Reverse Path Forwarding discards packets with invalid source addresses.

Route filters verify the BGP updates from each external BGP neighbor. These filters are used to prevent invalid routes advertised by a neighbor from entering the routing tables. In the cases

where the neighbor announces too many routes to filter, a default filter, blocking well-know invalid routes, is used.

Site route and packet filters are relatively static and are changed manually at the request of the ESnet Site Coordinator. The filters for other peers are update routinely, by scripts, using information from the Internet Routing Registry, the peer and data from the ESnet border router.

The ESnet routers are periodically polled for routing information and any anomalies are announced, via e-mail, to the engineering staff.

The routers are configured to send traps to the ESnet Network Management System in the event of a BGP state transition indicating that an external peer of ESnet has either gone down or has become once again responsive. The Network Management system receives the traps and generates alarms for the operations staff.

2.3.4.3 Service Level

IPv6 unicast is treated as any other production network service including 24x7 support.

2.4 Network Monitoring

The ESnet Network Operations Center (NOC) is operated 24 hours a day every day. The NOC staff monitor the network continuously in real time using the Spectrum $_{tm}$, a Simple Network Management Protocol (SNMP) based network management system (NMS) from CA $_{tm.}$. Spectrum provides the NOC with network hardware and protocol alarm notifications in real time displayed all in one place on a single web browser screen with audible messages to alert operators when alarms are raised and cleared. ESnet On Call Staff (OCS) and Person Of the Week (POW) also have access to the same Spectrum tool set used at the NOC.

The NOC also monitors system log (syslog) messages from routers and switches displayed as they occur from a centralized syslog server. Email messages to trouble@es.net and phone calls are also handled promptly by the 24X7 NOC.

2.4.1 Layers 1, 2, 3 Monitoring

2.4.1.1 Service

The ESnet NMS continuously monitors 74 routers, 302 Border Gateway Protocol (BGP) adjacencies with external networks and 748 point to point circuits and other directly connected network address ranges. Alarms are generated for link failures, interior protocols, threshold exceeded for interface error, system loading, environmental parameters and authentication failures. Both of the fault tolerant Spectrum NMS servers use polling and asynchronous device trap messages to generate network alarms on the web based operations console.

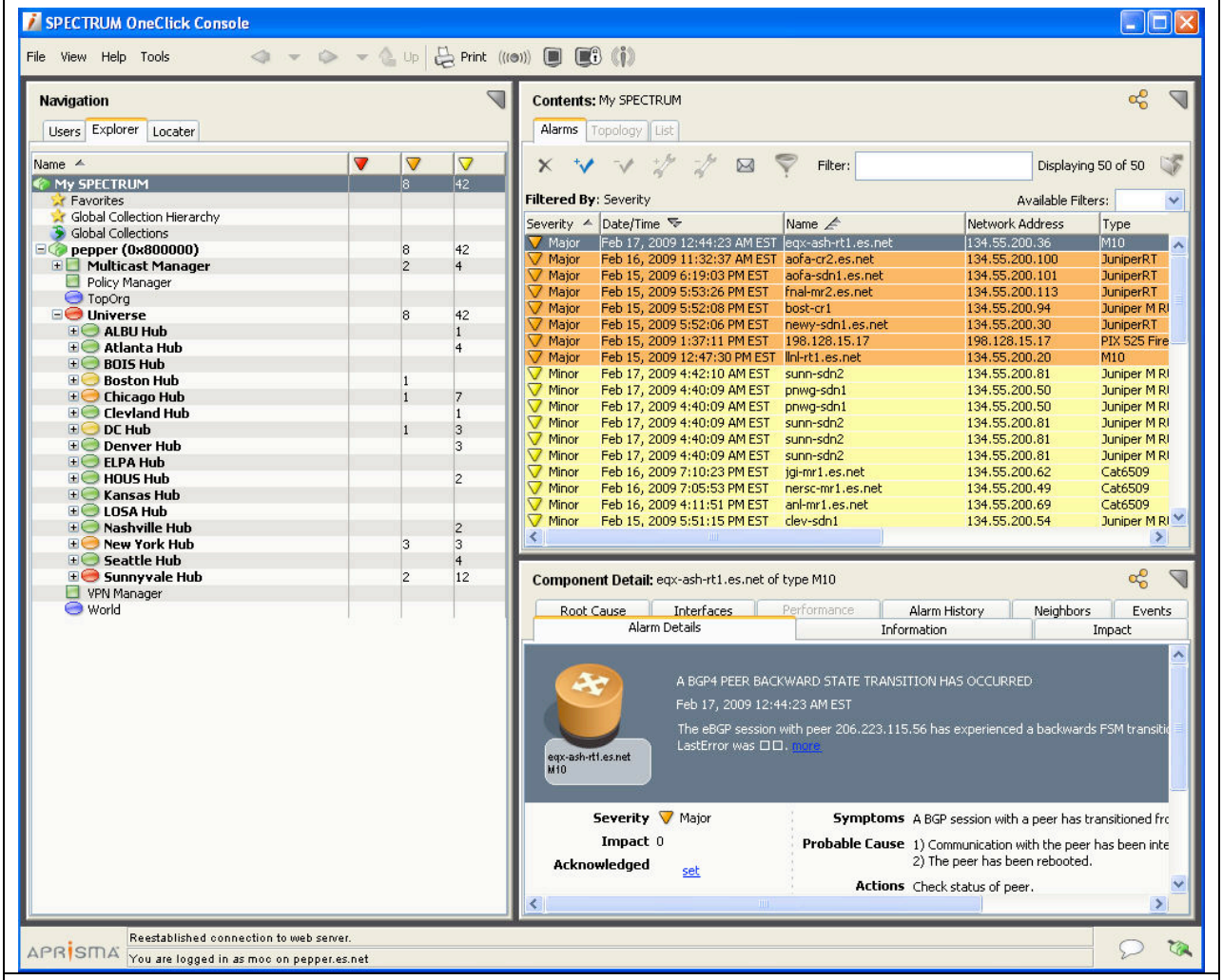

Figure 2-3 Spectrum Alarm View

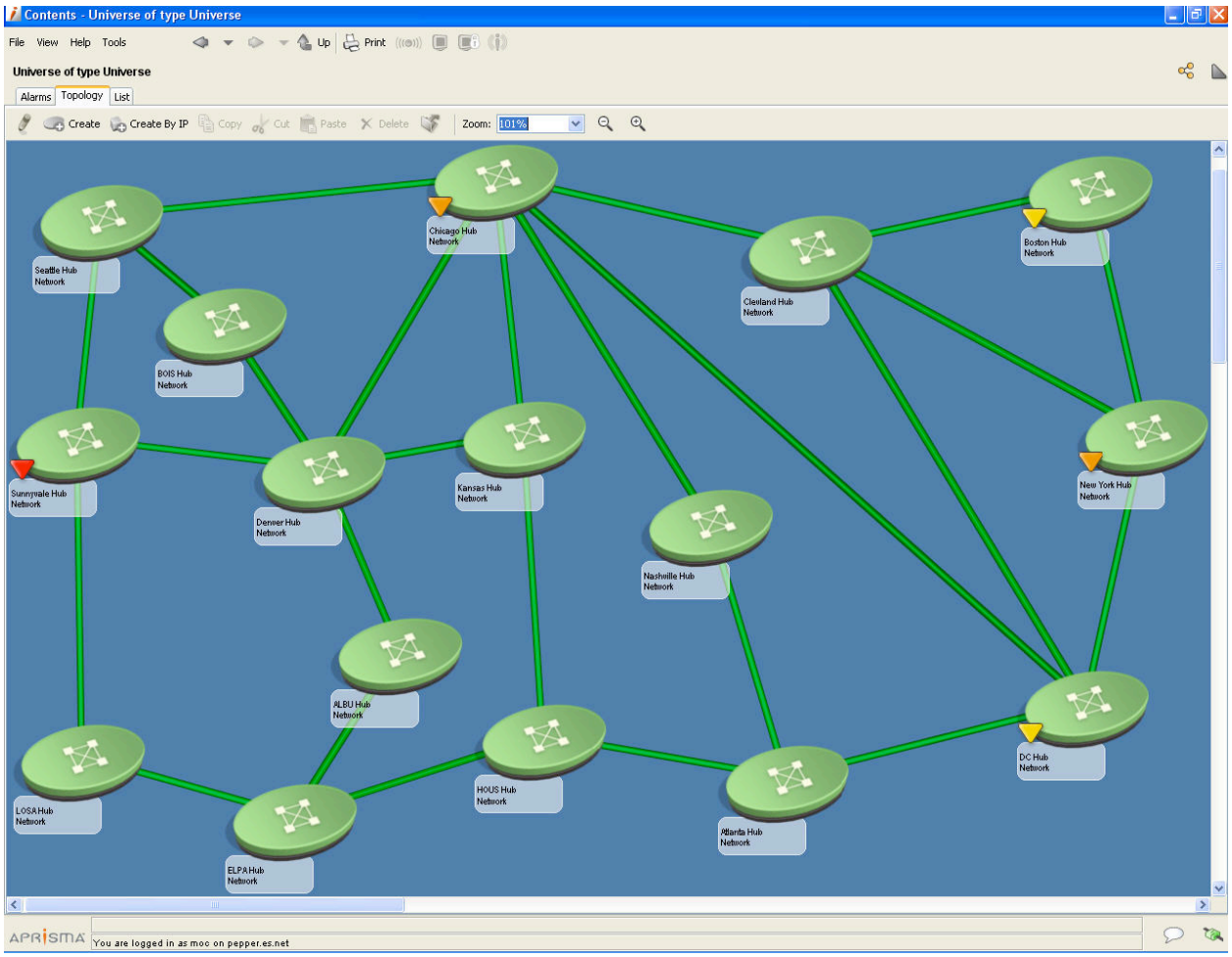

Figure 2-4 Spectrum network topology view

ESnet participated in the Beta testing of the Spectrum MPLS Transport Manager product. This new product enables ESnet to monitor OSCARS virtual circuits using MPLS protocol state. With this tool ESnet have the ability auto-discover the OSCARS topology and generate real-time service alarms the same NOC display already in use at ESnet.

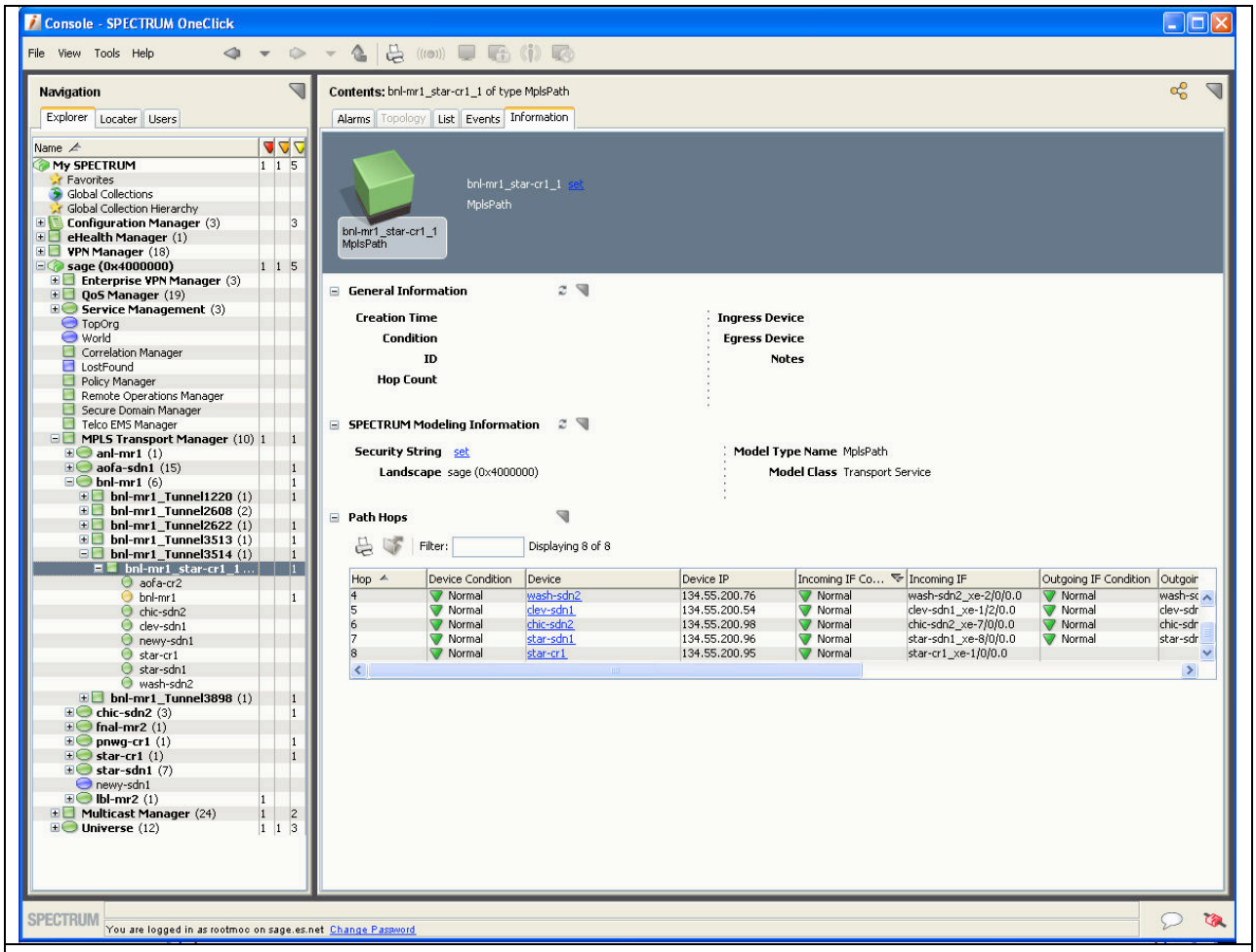

Figure 2-5 Spectrum MPLS monitoring of OSCARS virtual circuits

2.4.1.2 Service Level

Based on NMS alarms and syslog message information the ESnet NOC responds in real-time to network problems at the physical device layer and circuit link layer.

Physical layer response – The ESnet NOC responds to physical device alarms raised by the NMS and messages displayed on the syslog server by first level troubleshooting. Escalation is to the On Call Support (OCS) person. The OCS may then decide to escalate the issue by contacting the engineering Person of the Week (POW) or an external organization such as a customer site or contracted carrier.

Link layer circuit outage – The ESnet NOC responds to link layer circuit & interface alarms raised by the NMS or displayed on the syslog server by first level troubleshooting. Escalation is

to the OCS. The OCS may then decide to escalate the issue by contacting the POW or an external organization such as a customer site or contracted carrier.

2.4.2 Protocol Monitoring

2.4.2.1 Service

Based on NMS alarms and syslog message information the ESnet NOC responds in real-time to network protocol layer problems.

2.4.2.2 Service Level

Network Protocol response - When a BGP protocol peering relationship transitions out of the "established" state, the associated ESnet router immediately sends a SNMP trap message to both of the Spectrum servers which raise a BGP protocol alarm on the Spectrum alarm view (Figure 2-3). The NOC staff responds to these BGP protocol alarms by contacting the peer network if outage is unplanned and ESnet has not already received some form of communication acknowledging this protocol outage.

2.5 Network Measurements

2.5.1 Bandwidth Measurement Infrastructure

2.5.1.1 Service

The ESnet Measurement Infrastructure is a collection of services deployed on hosts across the ESnet backbone used for active performance monitoring between measurement points between ESnet hubs, sites, and our external partners. It is based on perfSONAR tools and protocols, which are being developed in collaboration with Internet2, GEANT2 and numerous others.

The achievable bandwidth portion of this infrastructure supports performing active measurements to determine the peak bandwidth that a test can reach between 2 points using BWCTL, a tool developed by Internet2. It supports both ad-hoc measurements and regularly scheduled tests.

This provides an end-user with an inside view of the achievable capacity in real-time within the ESnet backbone, and between external locations to different points in the ESnet backbone. The capability is extremely useful in investigating bandwidth throughput issues, confirming functionality after maintenance or upgrades, and establishing baseline expectations.

2.5.1.2 Service Level

Due to the intrusive nature of throughput tests, test parameters are controlled based on the end point IP addresses, and different restrictions are imposed based on endpoint classes. For example, higher bandwidth and longer duration tests are allowed to ESnet sites, while tests that transit the ESnet commodity Internet connections are severely restricted.

Results from regularly scheduled tests are stored in perfSONAR Measurement Archives and historical trends can be retrieved and graphed using perfSONAR visualization tools.

Ad hoc tests are run using BWCTL. The list of ESnet measurement points that support ad hoc BWCTL tests for our community can be found at the Fasterdata (fasterdata.es.net) site on the "Active perfSONAR Services" page.

ESnet will work with any ESnet site, or other network entity to setup regularly scheduled measurements. ESnet provides measurement servers in the major ESnet hubs. The partner is responsible for providing, configuring and maintaining the server on their end. ESnet will schedule the regular measurements, archive the results, and provide a web interface to support visualizing and analyzing the results.

2.5.2 Latency Measurement Infrastructure

2.5.2.1 Service

The latency measurement portion of the ESnet Measurement Infrastructure supports active measurement tests to determine the one-way delay between 2 points using OWAMP, a tool developed by Internet2.

This provides an end-user with a detailed accurate view of the delay and jitter across the ESnet backbone, and between external locations to different points in the ESnet backbone. The capability is extremely useful in investigating loss, queuing or congestion issues, confirming functionality after maintenance or upgrades, and establishing baseline expectations.

2.5.2.2 Service Level

ESnet will work with any ESnet site, or other network entity to setup regularly scheduled one way delay measurements. ESnet provides measurement servers in the major ESnet hubs. The partner is responsible for providing, configuring and maintaining the server on their end. ESnet will schedule the regular measurements, archive the results, and provide a web interface to support visualizing and analyzing the results.

Results from regularly scheduled tests are stored in perfSONAR Measurement Archives and historical trends can be retrieved and graphed using perfSONAR visualization tools.

Ad hoc tests are run using OWAMP. The list of ESnet measurement points that support ad hoc OWAMP tests for our community can be found on the web site. at the Fasterdata (fasterdata.es.net) site on the "Active perfSONAR Services" page.

2.5.3 Uptime Measurements

2.5.3.1 Service

Availability is one of the principle metrics used to evaluate a network service. Service availability is of primary interest to ESnet and ESnet Sites. In order to maintain high service levels ESnet develops tools to effectively measure and manage the wide ranging and varied types of service interruptions.

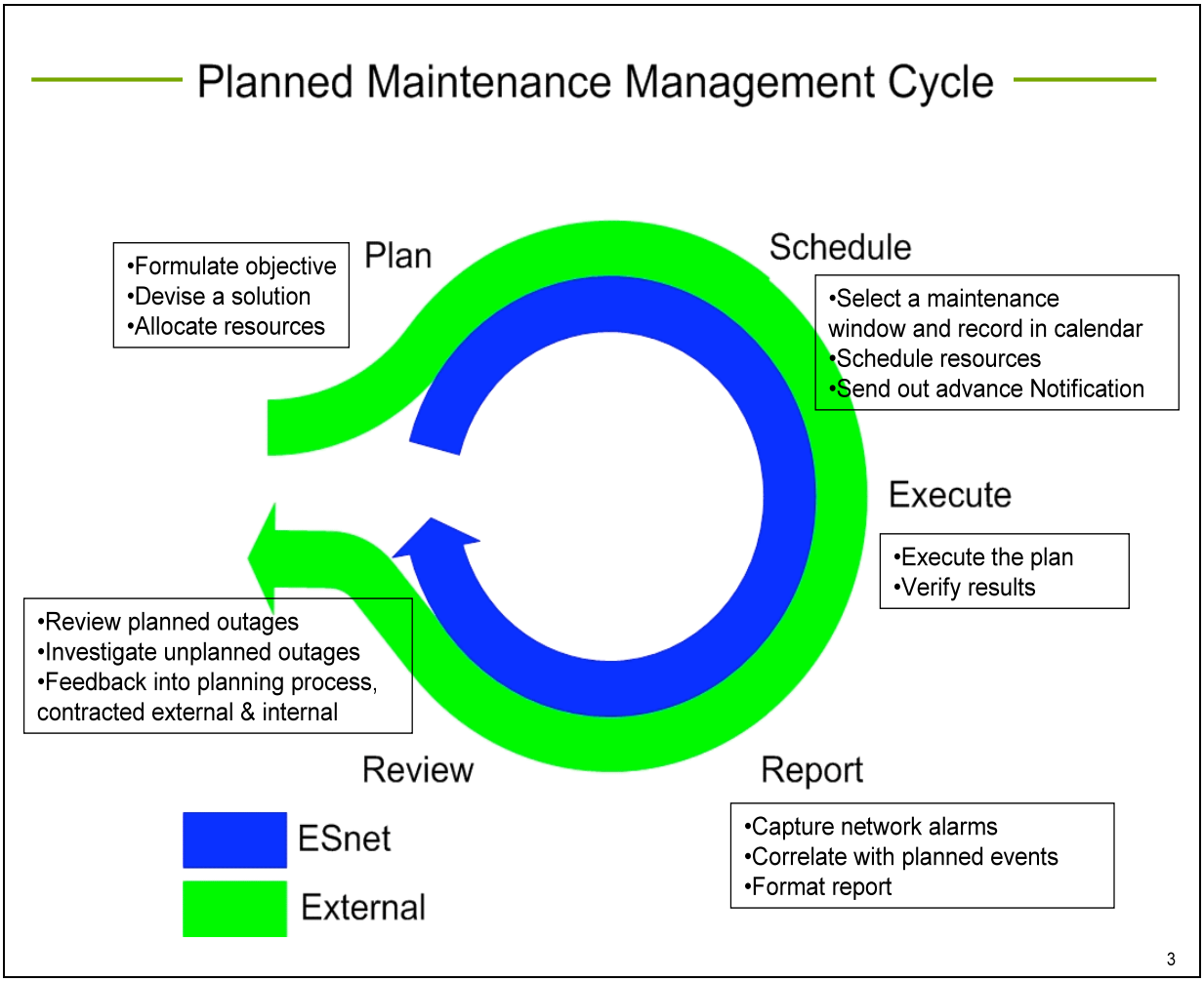

Figure 2-6 Planned Maintenance Cycle

In order to continuously improve performance ESnet makes every effort to understand and learn from past experience. The closed feedback loop in the above figure (Figure 2-6) depicts the ESnet process for systematic improvement through feeding lessons learned back into the planning cycle.

A benefit of accurately measuring service availability is that these metrics can be used to differentiate ESnet from other Internet Service Providers. Ease of demonstrability further differentiates ESnet from other networks.

2.5.3.2 Service Level

Router and router interface outage alarm reports are run against the Spectrum NMS on an hourly schedule. ESnet uses this data to compute customer site and network core availability reports. ESnet measures each site's availability along the paths originating at the customer border interface back to this site's ESnet core hub facility. The core network availability is reported

separately from all customer availability. Router unreachable alarms on a core router will appear in both the backbone metric and any associated site metrics.

ESnet customer site availability reports are publicly available at http://calendar.es.net/

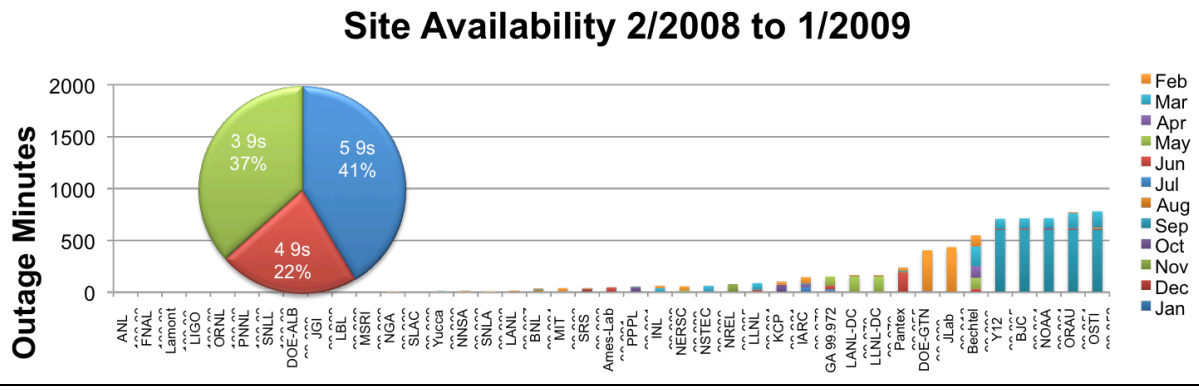

Figure 4

2.5.4 Usage Measurements

2.5.4.1 Service

The capability to collect and analyze network statistics data is one of the most important tools for identifying both long term and real time network issues. In addition to long term and real-time utilization reports, ESnet has developed the capability to analyze traffic usage according to flows. Flow data is analyzed to generate short term detailed reports, as well as long term trending information. Flow analysis is an elemental component in making traffic engineering decisions and in trouble shooting the network.

2.5.4.2 Service Level

Each of the 2400+ (physical and logical) interfaces in ESnet is tracked by two completely distinct network statistics systems.

The new system, ESnet eXtensible SNMP system (ESxSNMP), was developed in house to specifically meet the needs of ESnet. It polls for data every 30 seconds, archives the raw data indefinitely, detects changes in the network automatically and supports distributed collection and querying of the data. ESxSNMP is used to generate monthly statistics reports and provides the data to ESnet's perfSONAR Measurement Archive (MA). Work is underway to present a highly interactive view of the data for both real time and long term data visualization. To handle the long term data storage a simple database called Time Series Data Base (TSDB) was developed in conjunction with but distinct from ESxSNMP. ESxSNMP and TSDB are available as open source software at http://code.google.com/p/esxsnmp/ and http://code.google.com/p/tsdb/ respectively.

The legacy system is based on an in-house modified version of the widely used Multi Router Traffic Grapher (MRTG) network traffic monitoring system in association with the RRD (Round Robin Database) logging and graphing software. MRTG/RRD is configured to poll every 60 seconds, and the data is archived in the round robin database for up to two years. This system is used to plot interface traffic on a given interface over a specified time range. MRTG is not suitable for continued use due to it's high level of hands on daily maintenance, inadequate data retention and scaling issues; as a result the MRTG system will be phased out in favor of ESxSNMP the near term.

Each of the collection systems is fully redundant to reduce data loss. Data from both collection systems are interchangeable via a transformation process. Before MRTG is phased out a second, independent instance of ESxSNMP will be deployed on the east coast to both minimize data loss and provide disaster recovery. All data from either system are available to ESnet staff and site personnel on demand

Netinfo is a web portal that allows the ESnet community to access much of the network performance data that is collected on a regular basis. This tool allows sites to view current data from ESnet devices relating to traffic, latency, and routing information. It also provides access to the Flow collections, Maintenance Calendar, and Outage Reports.

A PerfSONAR Measurement Archive (MA) is used to publish statistics from one of the ESnet statistics systems. This MA exports data in the Global Grid Forum Network Measurement Working Group Schema version 2.0 via web services. The data exported by this service consists of the link utilization data for all ESnet site access, backbone and R&E peering links. The utilization data, with metadata including interface names, descriptions and addresses is exported to all users with no access restrictions.

2.6 Other Network Services

2.6.1 Virtual Circuits

2.6.1.1 Service

ESnet can provide user driven virtual circuits (VCs) to cater to special needs as required by the science requirements of ESnet end-sites. The virtual circuit service was designed around the following functional requirements:

- Configurable: The instantiation of the VC is dynamic and is driven by the user's requirements.
- Predictable: The instantiated VC should have predictable properties that the user can leverage.
- Schedulable: Guaranteed bandwidth is a scarce resource and as such should be obtained through a resource allocation process that is schedulable.
- Usable: The service must be easy and intuitive to use.
- Reliable: Resiliency strategies should be largely transparent to the user.
- Informative: Users should be provided useful information about reserved resources and VC status so that they can make intelligent decisions.

- Scalable: The underlying network should be able to manage its resources to provide the appearance of scalability to the user.
- Secure: The user must have confidence that both ends of the VC is connected to the intended termination points, and that the VC cannot by "hijacked" by a third party while in use.
- Provide Traffic Isolation: Allow the user to use non-standard or aggressive protocols without affecting other users.

2.6.1.2 Service Level

ESnet has developed the On-demand Secure Circuits and Advanced Reservation System (OSCARS) to address the need for VC services. This service supports user reservable layer 3 (IP), and layer 2 (Ethernet VLAN) VCs. To address the functional requirements of the VC service (see 2.6.1.1), OSCARS was implemented using the following approach.

- Configurability: Each VC is tailored to meet the user's request for bandwidth demand, time usage, and end-point terminations. This is done using a combination of tools, mainly a scheduler to manage the reservations, and standard protocols in the network such as OSPF-TE, RSVP-TE, MPLS-TE, and LDP to manage Label Switched Paths (LSPs) which constitute the VCs.
- Predictability: Each VC is traffic engineered with multiple constraints such as bandwidth, latency, hop-count, and congestion, to reduce unexpected service behaviors.
- Scheduling: The RSVP-TE protocol lacks the capability to support future reservations. To compensate for this deficiency, provision planning of each VC is done on a centralized server that keeps track of link-state reservations for the entire network. This is to prevent over-subscription of guaranteed bandwidth on a link-by-link basis.
- Usability: Users can initiate a VC reservation request either though a web based user interface, or an API. The former allows not only users to manage their individual VC requests, but also permits designated network site coordinators to manage all VC reservations that terminate at their site (regardless of the requesting user).
- Reliability: The current resiliency plan calls for redundant (or backup) VCs to be configured. This may appear to be inefficient as twice the amount of bandwidth is reserved (i.e. 2 VCs instead of 1). However the remarking of over-subscribed VC packets into scavenger-service allows for backup VCs to be reserved with a nominal guaranteed bandwidth and still burst to line rate if capacity is available. This scheme allows sites to control the level of redundancy as well as bandwidth guarantees based on service criticality and service level expectations. Future resiliency plans will rely on scheduling "hot" standby paths.
- Informative: Users can receive VC status updates via a notification broker. When a user makes a reservation request, the user will be registered with the notification broker to receive updates for the corresponding VC.
- Scalability: To scale the VC service effectively within the wide area network, the VC path computation element takes into consideration all paths between the termination end-points, and not just the shortest path as defined by default routing.

- Security: The VCs, which are created using MPLS LSPs, are "secure" within the edges of the ESnet WAN. Packets cannot be injected into or withdrawn from the VC except at the termination end-points. Packets entering the VCs are filtered based on the information provided by the user's reservation requests (i.e. IP flow specification for layer 3 VC, and Ethernet VLAN ID for layer 2 VCs).
- Provide Traffic Isolation: By creating distinct MPLS LSPs for each VC, policing each VC individually on ingress, and using QoS to isolate VC packets (that are not over-subscribed) into a separate class-of-service queue, normal IP routed best-effort traffic is not affected by VC traffic even when unfriendly protocols are used.

Currently this VC service supports 17 production VCs for the LHC project.

2.6.2 Traffic Engineering

2.6.2.1 Service

ESnet employs a variety of techniques to make the best use of the resources deployed, these include: scavenger service, and site specific traffic engineering to support programmatic needs such as LHC Tier1 to Tier2 support at FNAL

2.6.2.2 Service Level: Traffic Engineering

The scavenger service deployed in ESnet enables a user to utilize unused network capacity without impacting the default best-effort class of service. The bit marking (DSCP) selected for this service and the service itself is consistent with Internet2's QBone Scavenger Service (QBSS). This service provides a mechanism for sites to transfer huge amounts of data in the "background" without affecting day-to-day traffic. This is done by allocating a separate queue within each router in ESnet and configuring it with an aggressive drop profile and minimal service quota.

Site specific traffic engineering range from simple static routing, to more complex policy based routing involving QoS and admission controlled virtual circuits (as outlined in 2.6.1). Policy based routing in association with virtual circuits provide a powerful tool for directing specific traffic flows over specific paths.

2.7 Engineering and User Support Services

ESnet provides comprehensive user support, including "owning" all trouble tickets involving ESnet users (including problems at the far end of an ESnet connection) until they are resolved. A 24 x 7 operations center and 24 x 7 on-call network engineers provide continuous operational support of the network: trouble@es.net, 1-800-33-ESnet (1-800-333-7638).

ESnet also provides custom engineering services to facilitate special science needs via specialized routing, traffic engineering, etc., when other solutions are not available.

2.7.1 24 x 7 Support of Network Problems

2.7.1.1 Service

ESnet is staffed to respond to network trouble calls 24x7x365.

2.7.1.2 Service Level

ESnet has an escalation procedure for handling network outages. The first line of troubleshooting is handled by the 24x7 Operations personnel of the Oakland NOC. If they determine that they can not resolve a problem, they will notify the next level of support, the On-Call Support (OCS) person. The OCS team is staffed by WAN hardware engineers who are compensated to be on-call 24 X 7 on a rotating basis. The OCS will respond within 20 minutes of being contacted to the NOC. If the scope of the problem is beyond the knowledge base of the OCS, the problem is escalated directly to a member of the Engineering Staff designated the Person of the Week (POW).

2.7.2 Trouble Ticket System

2.7.2.1 Service

ESnet operates and maintains its own trouble ticket system, based on the commercial Remedy product, which is capable of tracking problems from first report to final resolution. The trouble ticket system is used to track all requests, outages, alarms, etc., including all end-user requests.

2.7.2.2 Service Level

The trouble ticket system is backed up with redundant hardware and software at a commercially hardened backup facility in New York City.

Personnel at ESnet Sites have access to their tickets, and all related communication, via a simple "finger" interface. Web and application interfaces are used by ESnet personnel.

2.7.3 Assisting Site Network Personnel

2.7.3.1 Service

ESnet maintains a highly collaborative and interactive relationship with the DOE Labs and scientists for planning, configuration, and operation of the network

ESnet staff work with Site networking personnel to set help up connectivity. This includes assisting in configuring routing protocols with the site including multicast, traffic engineering and backup routes.

2.7.3.2 Service Level

Much of this service is provided during regular business hours or by a Site requesting off hours assistance. ESnet experts can, and often do, assist sites with internal routing issues on a best effort basis.

2.7.4 ESnet Owns Trouble Tickets Until Resolved

2.7.4.1 Service

ESnet's mission is to enable the network based aspects of SC science, and that includes troubleshooting network problems wherever they occur. To that end, ESnet takes responsibility for or "owns" all trouble tickets involving ESnet users (including problems at the far end of an ESnet connection) until they are resolved.

2.7.4.2 Service Level

Not only will ESnet accept a trouble ticket from any ESnet user trying to communicate to any host both at another ESnet site or another ISP but ESnet will accept a trouble ticket from any user, homed in any ISP, trying to communicate with an ESnet site.

"Owing" the ticket entails tracking the problem to resolution and may involve follow up with other sites or other ISPs as needed. A recent example involved a user at DESY in Germany having poor performance to an ESnet site Femi National Laboratory. The path between the two sites crossed 6 network domains. ESnet personnel helped organize the debugging effort and suggested the technique which eventually solved the problem. For details, please refer to: http://www.es.net/pub/esnet-doc/FNAL-DESY%20v1%200.pdf.

2.7.5 Scheduled Maintenance Calendar

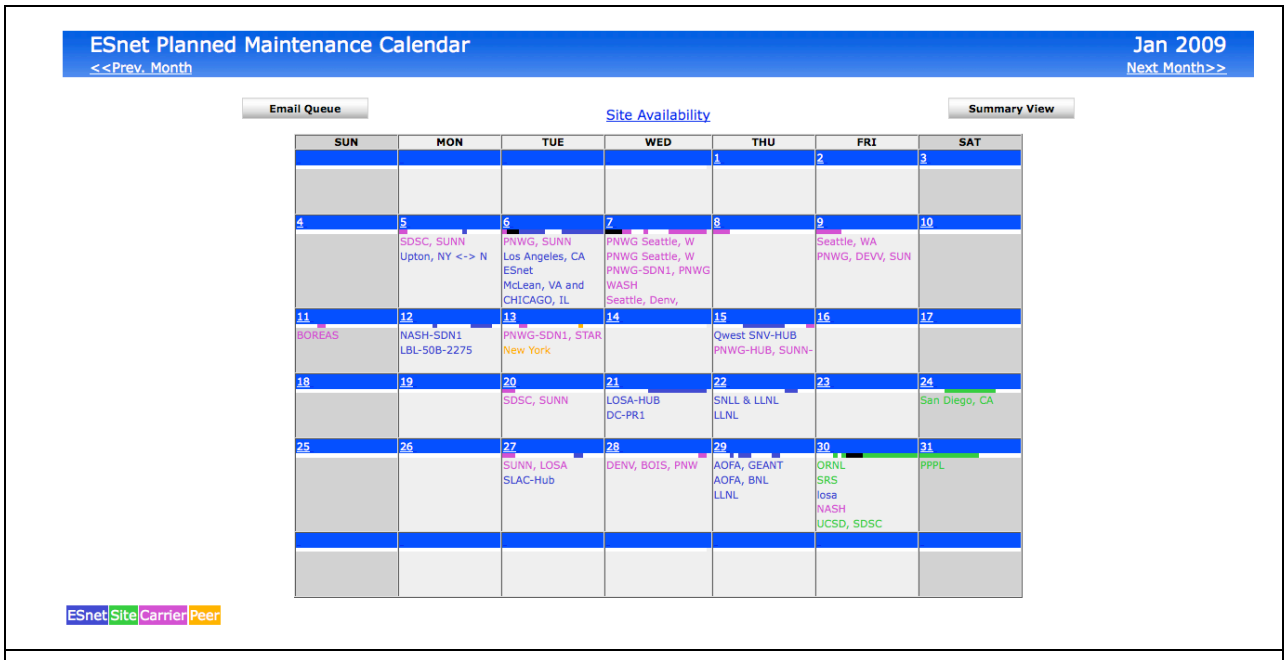

Figure 2-7 Planned Maintenance Calendar

2.7.5.1 Service

ESnet has developed an advanced network availability management system built around a web calendar user interface. The result is a publicly accessible network maintenance and

outage information clearinghouse for outage information past present and future, relevant either directly or indirectly to ESnet network availability.

This system is available in a read only version to the public on the external ESnet web site http://calendar.es.net/cgi-bin/pmcalendar.pl.

2.7.5.2 Service Level

The planned maintenance calendar application is able to correlate observed outages reported by the Spectrum NMS with ESnet trouble tickets information, providing a mechanism for quickly indexing through many years worth of trouble tickets to find important ticket information relative to a particular event.

This user-friendly system provides the NOC staff with an email notification template that helps to compose and queue for delivery, standardized network maintenance and alert notification email. Just as important, is that the structured email template makes it possible to extract important network maintenance and event information from otherwise unstructured email. Preserving important event information in this way allows ESnet to automatically correlate email notifications with the observed NMS outage alarms. Externally generated service affecting email from sites and peers is also input using the same Email template by an operator ensuring that this data is also archived in machine-readable form.

ESnet site availability reports not only display precisely when a service outage occurred but also go a long way to explaining why it occurred. The new and recent generation of ESnet Metropolitan Area Networks (MANs), are designed with a high degree of fault tolerance not found in earlier ESnet topologies.

2.7.6 Email

2.7.6.1 Service

It is critical that ESnet be able to communicate with its customers, suppliers, and collaborators. To this end, ESnet runs its own mail system. ESnet mail systems support IMAP, POP, and Webmail.

2.7.6.2 Service Level

ESnet mail service consists of geographically redundant, highly available mail routers located on the East and West coast of the USA.

The mail routers handle both in IPv4 and IPv6 traffic with the ability to route onto different protocols if needed. Meaning that a message coming in on IPv6 can route to its final hop via IPv4 and IPv4 can be routed via IPv6. All emails in and out bound to/from ESnet are handled by the four mail routers with no direct client and/or server system access. All messages are filtered, virus checked and harmful contents removed to protect the workstations or other servers and services. The mail routers deploy dynamic firewalls and auto blocking services. Additionally, the only ports advertised for service are 25, 366, 465, and 38752. All other traffic is filtered and/or blocked. The mail routers also handle all ESnet mailing lists both public and private. These are administered in either public or private fashion with all messages being scanned for virus, Trojans, and other harmful content.

The ESnet IMAP & POP servers handle only message storage and are not able to relay or send email. They can only receive email from the above approved secure mail routers. Access methods to the message storage system are only supported via an encrypted means, no clear text is allowed from either the client or the connecting mail routers. Clients must use one of the following supported methods to connect: TLS, SSL, SSL with 5MD5-Digest, or SSL w/MD5-CRAM.

The ESnet webmail server provides ESnet staff with secure and safe access to their email from any system capable of using a browser with SSL connectivity. Webmail sends email by relaying mail to the smtp routers via SSL with MD5-Digest encryption. All email read by webmail is transferred from the message storage system via SSL only. Connectivity to Webmail can only be done via SSL connection, no clear access or clear text allowed.

2.8 Security

2.8.1 WAN security

2.8.1.1 Service

ESnet employs best current practices in protecting ESnet resources in the WAN.

2.8.1.2 Service Level

ESnet employs a layered security model in protecting its assets as shown in Figure 2-8 WAN Layered Security Model.

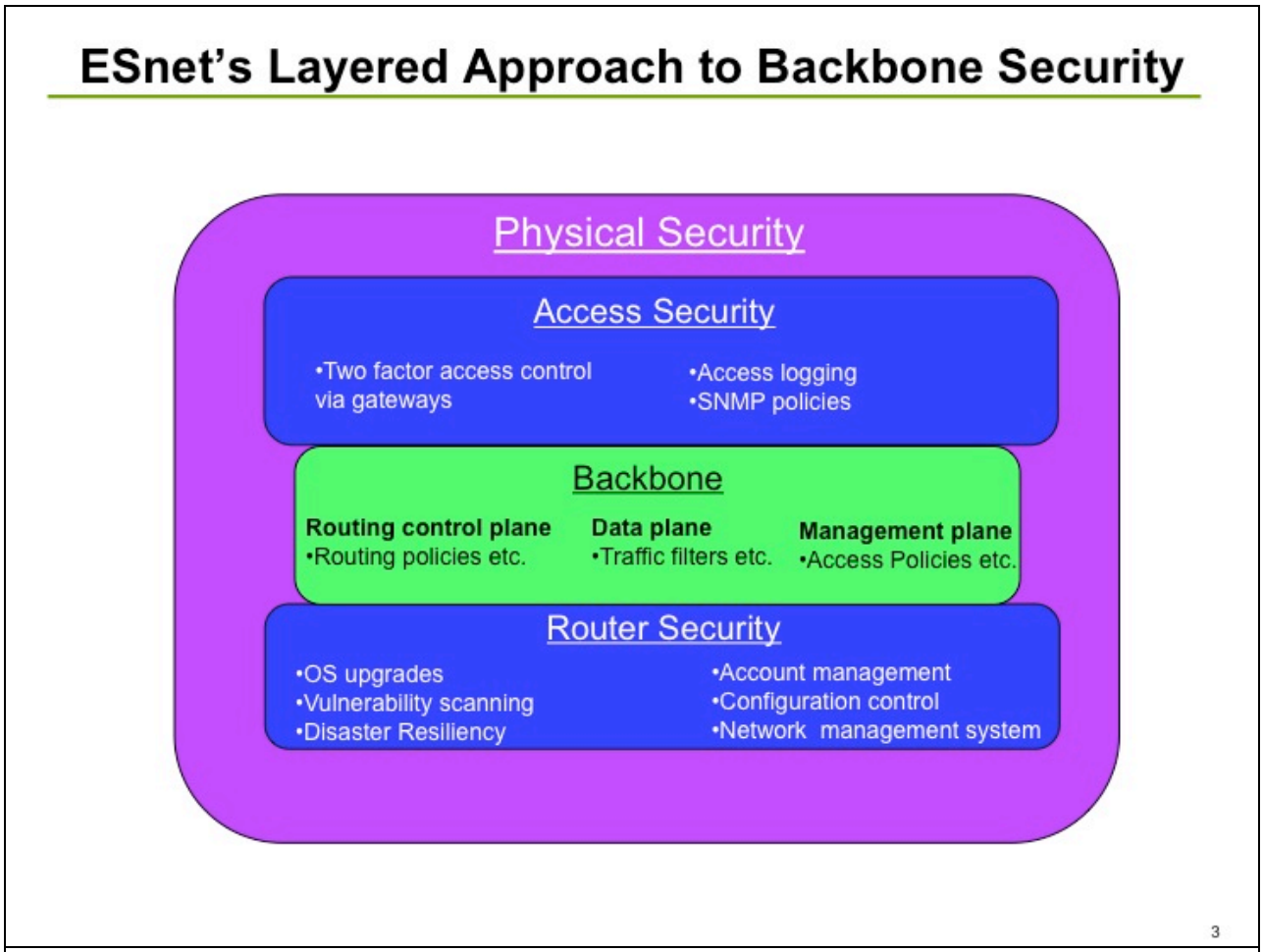

Figure 2-8 WAN Layered Security Model

ESnet locates all routing hardware in secure locations, including telecommunications carrier hotels, DOE sites and research & education exchanges. All of these sites provide significant security.

ESnet limits access to routers to SSH and SNMP protocols. SNMP supports only read access. Both SSH and SNMP access is restricted to known and trusted sources.

The control plane is protected either by using non-routable protocols (ISIS), or by limiting them to the backbone IP address space. Forging of source packets is controlled by filtering all traffic entering the network for source addresses in the backbone network.

2.8.2 Cyber Defense

2.8.2.1 Service

While Sites are primarily responsible for providing themselves with cyber protection, ESnet can and will assist sites in case of emergencies. Requests for assistance must come from the Site Coordinator. The request should be sent to trouble@es.net and, minimally, contain a request for assistance and a call back number. ESnet will take steps to defend the network resource if impacted, or comes under an attack.

2.8.2.2 Service Level

There are many different scenarios involving an attack on ESnet or its Sites. The engineering staff will analyze the situation and formulate a response depending on the nature of the incident. The tools and techniques useful in responding to intrusions include:

- **Lowering the forwarding priority of incoming traffic**. Setting the priority of a large volume of malicious traffic to Scavenger means that production traffic is queued and forwarded ahead of the malicious stream. The attack traffic will not congest normal production traffic. This technique would be most useful in the event of a large scale DOS attack originating from a 10G connected Site or peer. It could also be useful if the traffic involves virus or worm propagation. Congestion would be relieved while an appropriate signature was determined.
  This technique would also be useful in mitigating traffic coming from a slower speed external connection that is overwhelming a slower speed internal link.
- **Filtering incoming traffi**c based on destinations, sources, and/or other attack signature. Once a valid signature for an attack is determined, filters at the network edges can block it.
- **Rate limiting traffic**. This technique is potentially useful if there is a large volume of malicious traffic directed to host(s) (say a web server) that congests a site access or LAN. Traffic to the host can be rate limited to allow normal traffic access to the line.
- **Disconnectin**g, or partially disconnecting, the backbone nodes from all Peers and Sites. Disconnecting is a useful strategy if there is an overwhelming attack or infection. It allows Sites to take recovery steps in isolation before attempting to reattach to the network

2.9 Disaster Recovery and Stability

ESnet has procedures, physical resources and personnel in place to provide continuity in service in the face of a disaster.

2.9.1 Disaster Recovery and Stability

2.9.1.1 Service

The network is designed to be available even if, e.g., the NOC on the West Coast is disabled by a massive earthquake, etc.

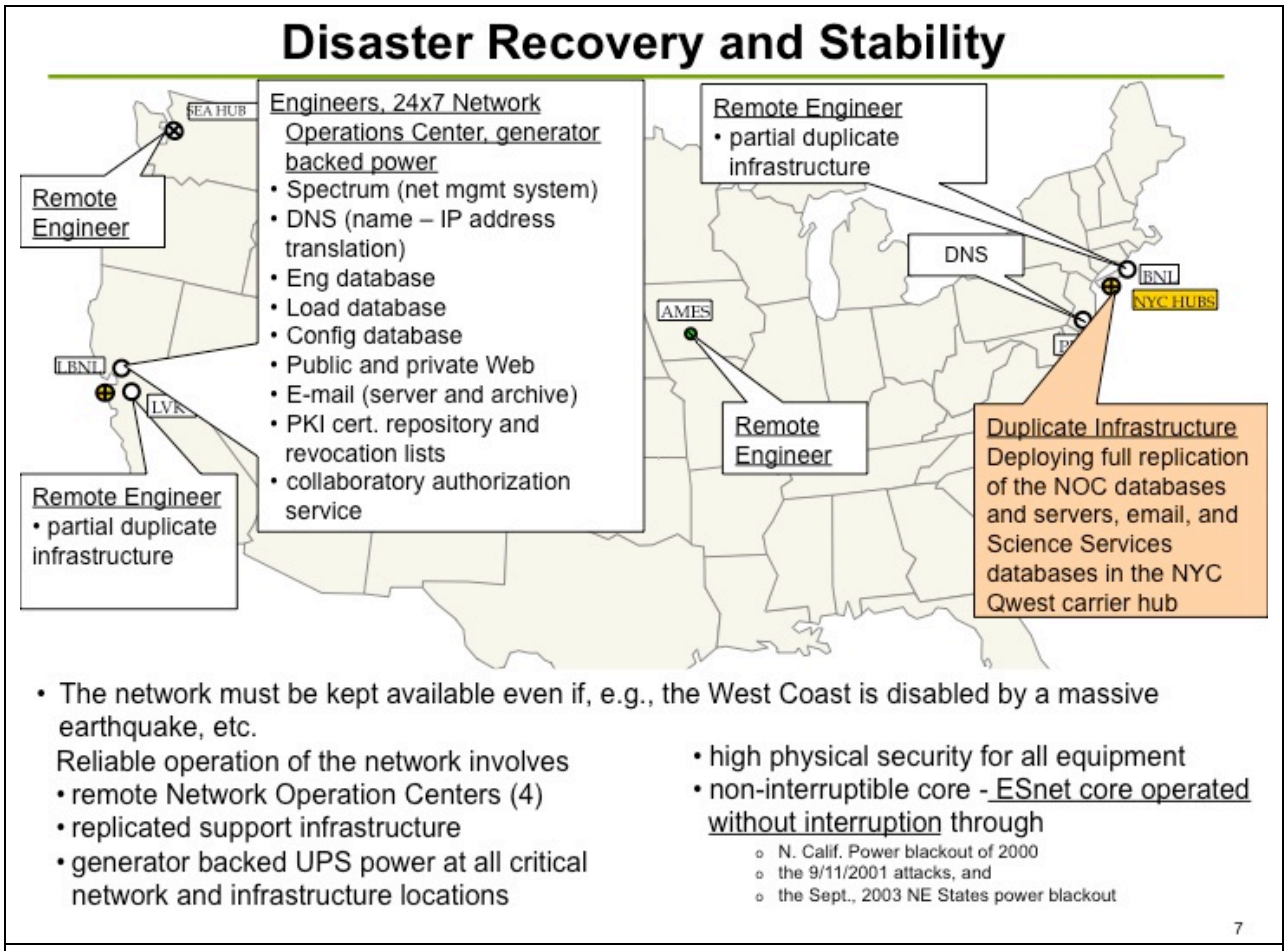

Figure 2-10 ESnet Disaster Recovery and Stability

2.9.1.2 Service Level

Stability and disaster recovery are provided through multiple Network Operations Centers (NOCs) and replicated infrastructure. An unmanned NOC was created at the Qwest AOA POP to house replicas of the NOC databases and servers necessary to monitor and manage the network. The AOA Hub was chosen since it has multiple connections to the Backbone. The servers are synchronized with the primary servers and are immediately available in the event that connectivity to the NOCs at LBNL and NERSC is lost. This, in conjunction with the distributed workforce, will facilitate network operations in the event of an emergency. Additionally, ESnet engineers have secure, out-of-band network access to all network equipment, including backup secure telephony access to all routers.

3 Federated Trust

ESnet's Authentication and Trust Federation (ATF) group provides identity, authentication, and related services to various DOE Office of Science distributed computing and collaboration programs, particularly computational and data Grids.

3.1.1 DOEGrids Certification Authority

3.1.1.1 Service

The DOEGrids Certification Authority and the DOEGrids website serve the scientific programs of the DOE Office of Science, particularly in SciDAC and Grid applications, where distributed computing and PKI are required. DOEGrids has about 2200 "person" Grid certificates and about 6500 "service" Grid certificates. DOEGrids supports 12 large-scale registration authorities, each of which represents a project, an experiment, or a site (usually a DOE laboratory). Beside DOE Office of Science projects, Registration Authorities include NSF, and international participants, thanks to DOE's broad support for scientific collaboration. Some of these registration authorities have extensive structure (eg "Open Science Grid (OSG)" provides services for more than forty different projects across a broad range of scientific disciplines). About 35 agents and an equal number of GridAdmins (sub-agents) help process requests, verify identities, and authorize issuance of certificates. DOEGrids is governed by the DOEGrids Policy Management Authority, which provides guidance to CA operations and oversees DOEGrids compliance with international standards. Membership is drawn from the Registration Authorities, relying party and other liaisons, and ATF.

3.1.1.2 Service Level

Please refer to section 3.6 Federated Trust Service Level

3.1.2 ESnet Root Certification Authority

3.1.2.1 Service

ATF operates the ESnet root CA. This CA is never on a computer network, and is kept locked in a vault (with some components dispersed) when not actively engaged in a certification operation. This highly secure CA only signs subordinate CA certificates, such as the DOEGrids Certification Authority, the NERSC SLCS service, and the FusionGrid CA. The two-level CA architecture allows rapid revocation and recovery of a subordinate CA PKI should the subordinate CA be compromised.

3.1.2.2 Service Level

Please refer to section 3.6 Federated Trust Service Level

3.2 Other Certification Authorities

ATF operates a number of other certification authority servers, in partnership with Grid virtual organization or site collaborators. See "Secure Hosting" below.

3.2.1 PGP Key Server

3.2.1.1 Service

http://www.es.net/pgp/

The PGP Key Server provides storage and distribution of PGP keys. The key server allows users to search for public keys of other users, download keys for PGP use, upload signed keys, and manage key revocations. It is not a PGP certification authority but instead supports the PGP web of trust and personal certification authority mode of operation.

3.2.1.2 Service Level

Please refer to section 3.6 Federated Trust Service Level

3.2.2 Trust Federations – IGTF

3.2.2.1 Service

Trust federations bring together various stakeholders, who need to cooperate on security-related principles such as identity, roles, authorization, in order to accomplish a programmatic mission (such as a large scale science experiment, or a research program). ATF supports or is actively involved in several trust federations. ATF currently hosts the International Grid Trust Federation (IGTF), which federates regional Grid policy management authorities (PMA), and standardizes policies and practices between different regional policy management authorities. ATF is also an active participant in two of these regional PMAs: member since 2002 in the EUGridPMA (European regional); and a founding and organizing member of TAGPMA (Western Hemisphere regional). Each regional consists of each region's Grid certification authorities (sometimes extra-regional members), and each regional provides registration and policy guidance to CA operators and relying parties. ATF is also actively involved in developing standards for trust federation operation, by developing standards documents in Open Grid Federation and other venues. ATF also participates in development of trust federations in other domains, such as Shibboleth and OpenID based federations.

3.2.2.2 Service Level

Please refer to section 3.6 Federated Trust Service Level

3.3 Other Services

3.3.1 Hardware Security Modules (HSM)

3.3.1.1 Service

Hardware security modules manage the private keys (and other objects) of certification authorities and related services. The HSM reduces or eliminates the possibility of certain kinds of theft, helping to improve the security profile of a certification authority. The DOEGrids project evaluated several vendors' HSM products in 2002 and chose the nCipher HSM. HSM's are an integral part of ATF CA architecture. The nCipher product can also be used in other security-related applications and has some additional capabilities that may prove useful to collaborators (secure code, network operation, replication and reliability).

3.3.1.2 Service Level

Please refer to section 3.6 Federated Trust Service Level

3.3.2 Secure Hosting

3.3.2.1 Service

ATF in collaboration with other teams at ESnet have developed a secure logical and physical infrastructure for the DOEGrids certification authority servers. Co-hosting is available for use by other organizations. This infrastructure is oriented towards supporting security applications and their related servers, and is not a general purpose site. Grid Virtual Organizations, which

lack a physical home site, but need to support an authentication or authorization service for customers across the US or the around the world, are good candidates. Currently we are hosting certification authorities for ESnet, and FusionGrid. We host an experimental OCSP server for DOEGrids and the IGTF, and an experimental MyProxy service for OSG.

3.3.2.2 Service Level

Please refer to section 3.6 Federated Trust Service Level

3.3.3 ESnet Two Factor Authentication

3.3.3.1 Service

ATF supports ESnet's two-factor authentication internal service, based on hardware tokens, through the DOEGrids CA.

3.3.3.2 Service Level

Please refer to section 3.6 Federated Trust Service Level

3.4 Experimental Services

ATF has several non-production, pilot services available for test, collaboration, or further development to production status.

3.4.1 OCSP Server

3.4.1.1 Service

Online Certificate Status Protocol (OCSP) provides an online (client-server) method to check the validity of certificates. OCSP augments (or even replaces) the existing CRL mechanism. OCSP provides the capability of checking certificates status in near-real time, which is very difficult for the existing Grid PKI. OCSP can also outsource all the revocation checking software and data management to a server, allowing application developers to simply PKI usage and deployment. The DOEGrids experimental OCSP service demonstrated the capabilities of OCSP to the Grid community starting in 2003. The server collects CRLs from all current EUGridPMA participants and ESnet CAs, and will answer OCSP validation requests (http://amethyst.es.net/ocsp).

3.4.1.2 Service Level

Please refer to section 3.6 Federated Trust Service Level

3.4.2 MyProxy Server

3.4.2.1 Service

ATF has developed a pilot MyProxy server for PPDG (now OSG). The first phase of this project is deployment of a MyProxy server based on configurations and recommendations of NERSC and Fusion Grid MyProxy projects.

3.4.2.2 Service Level

Service remains at pilot/experimental level. Please refer to section 3.6 Federated Trust Service Level.

3.4.3 RADIUS Authentication Fabric

3.4.3.1 Service

ATF developed a RADIUS authentication fabric (ESnet RAF) to support various one-time password initiatives in DOE laboratories in 2004. A user logs in to an application, and the application forwards this "authentication query" through a network of RADIUS servers to the user's home site for verification at the home site authentication service (OTP backend, LDAP, password service, &c). RADIUS is a widely supported and widely deployed authentication and authorization service. ATF developed configuration, infrastructure, and federation techniques to support this technology securely. ATF is currently exploring potential uses of the RAF in supporting wide area wireless roaming (see Eduroam for a European version of this service), certification authority integration, VPN service, and others.

3.4.3.2 Service Level

Please refer to section 3.6 Federated Trust Service Level

3.4.4 Shibboleth (SAML2) services

3.4.4.1 Service

ATF has developed a plan and a proposal to provide X.509 certification authority services to a federation of Shibboleth identity providers (or perhaps, other SAML-2 speaking identity providing services). ATF also led the effort to make ESnet a member of the US higher education Shibboleth federation, InCommon. This is a two-pronged strategy – the CA effort provides a unique service bridging two important communities, Shibboleth-aware sites and Grid-aware resources; the InCommon effort provides a way for other ESnet services to appear as legitimate resources and Shibboleth token consumers in various federations. ESnet membership is established, and CA services are under development.

3.4.4.2 Service Level

Please refer to section 3.6 Federated Trust Service Level

3.4.5 OpenID services

3.4.5.1 Service

OpenID is a simpler and more flexible protocol that accomplishes much the same mission as Shibboleth, but with some additional features and some serious trade-offs in features and security. OpenID is also a single representative of a large ecology of identity protocols emerging from recent web research and development; we have chosen to work with this protocol because it has emerged as an important protocol in industry, and seems the most applicable to problems experienced in our community. Currently we have developed an experimental OpenID provider, capable of using our CA technology as a source of authenticators, and a few sample consumers; the next phase of this experiment will advance the OpenID provider and one or two interesting clients to production quality.

3.4.5.2 Service Level

Please refer to section 3.6 Federated Trust Service Level

3.5 Research and Development

3.5.1.1 Service

ATF has a significant research and development component, reflected in part in the Experimental Services above. ATF is involved with or has interests in other authentication and authorization related services and technologies. Currently ATF is investigating an extension of the "OCSP" idea to a full-fledged validation service, which would incorporate certificate path discovery and path validation, and more sophisticated key management protocols (XKMS, SCVP). ATF continues to develop HSM deployment and is investigating network HSM capability. Shibboleth and other federation technology is also a subject of interest. ATF is interested in studying firewall traversal and management issues as they pertain and conflict with Grid uses, and is investigating how certificates and PKI can interact with COTS firewall products.

3.5.1.2 Service Level

Please refer to section 3.6 Federated Trust Service Level

3.6 Federated Trust Service Level

All services are offered on a "best effort" basis. Support is available only during normal LBL business hours, 8 AM – 5 PM Pacific Time Monday-Friday. LBL also observes several holidays (see the LBL holiday schedule) and typically has a site-wide closure at the end of December. Absent off-hours problems, services should be available 24 hours a day, seven days a week. Problems that occur on holidays, site closures, or outside of normal business hours will be logged and ticketed, but cannot be addressed until normal business hours resume. Under some circumstances, the availability of personnel and other ESnet business may delay response to problems during normal business hours.

Problems with services are reported using standard ESnet trouble reporting (http://www.es.net/hypertext/welcome/pr/problems.html).

4 Audio, Video, Data Collaboration

ESnet Audio, Video, Data Collaboration (AVD) supports voice, video, streaming, and data collaboration technology that provide DOE Office of Science researchers and their collaborators the ability to meet and exchange information remotely as easily as if they were in the same location. At the present time, the ESnet AVD system has over 1000 registered users worldwide supporting such science initiatives as: ATLAS, D0, CDF, ILC, CMS, ZEUS, Alcator C-Mod, OSG, DOSAR, and others.

4.1 Audio and H.323 Conferencing

4.1.1.1 Service

1. Web-based registration where potential customers go to register themselves and, if required, their equipment. They can also re-visit those pages for HELP related to using the ESnet Collaboration Services. http://www.ecs.es.net

2. IP videoconferencing (H.323) is supported at ESnet with over 120 ports of video multipoint bridging capability (at 384 kbps) and a centralized video gatekeeper, a "name/address server". Customers are given valid ESnet assigned "phone" numbers for access, and are free to meet anytime in a pure "ad-hoc" manner.

3. Gateway services allow legacy H.320 ISDN video conferencing systems, and telephony users, to connect to IP H.323 meetings.

4. Audio conferencing is still telephony based and is a scheduled and reserved service. A customer schedules a meeting at this web site: http://audiobridge.es.net. Each meeting participant receives email notification of the meeting time. The audio bridge consists of 144 telephone ports.

5. Data conferencing supports thousands of users and is available via the audio bridge either with an audio meeting or by itself where it can be used with a videoconference.

6. Streaming via your browser using REAL or Quicktime is available to view H.323 meetings. These meetings can be password protected.

7. Web-based form to send email to H.323 ad-hoc meeting participants (http://www-staff.es.net/%7Emikep/adhoc/schedumail.pl)

9. Customer Testing, Assistance, and Advice: ESnet Collaboration Staff assists customers who are having trouble connecting or if they need advice on how to operate equipment. This is done on a time available basis only.

4.1.1.2 Service Level

ECS is supported 8 AM to 5 PM, Pacific Time, workdays (Monday thru Friday, except for holidays). Best effort service is provided. LBL observes several holidays (see the LBL holiday schedule) and typically has a site-wide closure at the end of December. Problems that occur on holidays, site closures, or outside of normal business hours will be logged and ticketed, but cannot be addressed until normal business hours resume. Absent off-hours problems, services should be available 24 hours a day, seven days a week.

Send email to trouble@es.net for assistance. The ECS Team will respond appropriately depending upon the emergency and workload

ECS Staff are NOT responsible for video or audio endpoints, firewalls, NATs, or LANs at the customer site. ECS Staff are only responsible for ESnet resources located in Berkeley, CA.

There is no guarantee of security or privacy in the collaboration services, though best practices are used to protect the AVD services from cyber attack.

5 Governance (Relationship and Responsiveness to Customers)

ESnet has several groups of stakeholders whose concerns have to be balanced and addressed.

ESnet is a DOE Office of Science facility whose mission and funding are established by a Federal Program Manager in the Office of Science.

ESnet sites typically have three primary constituents:

- The scientific staff concerned with a network that is architected to facilitate their science
- The general staff dependent on ESnet for Internet connectivity
- The Lab business operations staff (typically represented by the Lab CIO) dependent on ESnet for communications with the commercial world in order to conduct Laboratory business.

5.1.1 Site Requests for Network Services

5.1.1.1 Service

ESnet makes all efforts to be responsive to external Site WAN networking requirements. ESnet personnel work with Site coordinators and senior Site Network engineers to meet the Site's needs. Some examples of these requirements include enhanced maintenance response to site equipment, traffic engineering of specific flows, and additional backup connectivity.

5.1.1.2 Service Level

ESnet responds to these requests as expeditiously as possible. If a requested service incurs an additional charge to ESnet, ESnet will pass this cost to the Site. All requests of this nature must come from the designated Site Coordinator, the Alternate Site Coordinator or, in special cases, his/her delegate (see the document ESnet Site Coordinator Roles and Responsibilities).

5.1.2 Property Management of ESnet Equipment at Sites

ESnet is accountable to DOE and LBNL for all equipment it purchases, even equipment that is deployed at remote locations. This is an example of a service that is responsive to stakeholders other than the scientific and general ESnet users.

5.1.2.1 Service

ESnet maintains a property management system that ensures the tracking of ESnet property in accordance with policies and procedures approved by DOE and the requirements of the LBNL Prime Contract.

5.1.2.2 Service Level

ESnet does a full inventory of equipment once per year.

ESnet maintains its own detailed database tracking each piece of equipment in the field. Additionally, ESnet personnel comply with LBNL requirements and track the location of the equipment in the LBNL database as well.

The Site Coordinator is responsible for the traceability and the physical security of the ESnet equipment at the site, and must promptly respond to information requests by ESnet and LBNL Property Management representatives during inventories of ESnet equipment.

5.1.3 Meeting Obligations to DOE and Other Government Agencies

This is an example of a service that is responsive to stakeholders other than the scientific and general ESnet users.

5.1.3.1 Service

ESnet is required to submit periodic reports on financial and programmatic progress to DOE throughout the year. Additionally, ESnet responds to additional "one-time" requests that are not anticipated or planned for on an annual basis (e.g., the GAO Audit of ESnet's OMB-300 or the GAO's request for IPv6 readiness).

5.1.3.2 Service Level

The OMB 300 is an annual process that begins in early spring and ends in the fall (roughly a six-month process). DOE submits the report in the fall. ESnet submits a well written draft in the spring that goes through several reviews, official and unofficial, and is edited and revised after each review until a final draft is accepted by DOE in the early fall time-frame.

The Quarterly reports are a DOE-CIO requirement. These reports require reporting of the fiscal quarters data.

Appendix A: ESnet Circuit Acceptance Criteria

Background

ESnet has documents acceptance testing of circuits for a number of reasons.

1. Create Artifacts to meet management oversite & audit requirements.

   Changes to DOE project management practices are requiring more formal reporting of many aspects of the ESnet project, including network upgrades. For example, the BAMAN upgrades were milestones reported and tracked in the OMB 300 process.

2. Improve communication with our user community.

   ESnet is in a constant state of growth requiring frequent upgrades to keep up with the demands of the DOE science we support. Improved documentation of this process will help our user community understand how we are changing to meet their needs.

3. Network Engineering Support.

ESnet has always run tests on new circuits to ensure that they were performing properly before placing them in service. This process is getting more complex as our network architecture evolves include redundant site access links and backbone and the increasing use of local area network technology in metropolitian and wide area links. Rigorious documentation of the characteristics of the links before they are put into service will assist us with diagnosing problems down the line.

Acceptance Test Criteria

1. **Saturation Test**

   This test is to assure that circuit that has been delivered doesn't have any internal bottlenecks that would prevent it from running at capacity.

   The circuit shall be saturated with a s demonstrated bandwidth over 95% of the link capacity for 5 minutes.

   In situations where the integrity of the counters used to compute utilization are suspect, confirmation should be made by using a second set of counters.

2. **Loss Test**

   This test is to assure that the line meets 10 GE Standard loss rates.

   An quantity of data greater than the half duplex bandwidth of the link times 12 hours shall be transfered across the link in both directions with a frame loss rate less than or equal to the identified threashold.

Loss Thresholds

Link	Saturation Tests		Loss Test			Notes
	Line Rate[1]	95% Saturation Rate	FLR	Total Data Transfer	Acceptable Frame Loss in a 24 hour test	
10GE LanPHY	10 Gbps	9.5 Gbps	7.2E-7	100 TBytes	4000	10GE Standard & BAMAN Criteria [2]
	10 Gbps	9.5 Gbps	10E-10	100 TBytes	10	LIMAN Criteria [3]
DS3	44.736 Mbps	42.499 Mbps	7.2E-7	4.83 GBytes	38	This is the Ethernet FLR. The typical DS3 criteria allow dozens of errored seconds per day. They do not specify packet loss rates. 7.2E-7 is a lot more restrictive than the standard.

	44.736 Mbps	42.499 Mbps		4.83 GBytes	432 Errored SECONDS	One of our providers has established values of 99.5% error free seconds, or less than 432 Errored Seconds per day and 99.975 availability.

1. This is the layer2 frame rate. including Ethernet headers. It is not the IP payload rate reported by iperf.
2. This criterion is based on the 10GE standard.

   The 10GE standard specifies a BER of 10E-12 in 802.3ae clause 52. This can be converted to a frame loss rate using a formula from page 89 of "Gigabit Ethernet" by Rich Seifert, Addison Wesley 1998.

   - FLR = Frame Loss Rate
   - BER = Bit Error Rate
   - N = Bits in Frame (9K IP MTU +26 Bytes Ethernet Header)
   - FLR = $1 - (1 - BER)^{N}$
   - FLR = $1 - (1 - 10E\text{-}12)^{72,280}$
   - FLR = 7.2E-7
3. This criteria is based on the Mathis et al formula for the bounds that loss places on maximum TCP throughput as expressed in formula 4 in (see M. Mathis, J. Semke, J. Mahdavi, T. Ott, "The Macroscopic Behavior of the TCP Congestion Avoidance Algorithm",Computer Communication Review, volume 27, number 3, pp. 67-82, July 1997 using a MSS of 1460, RTT of 100 ms
4. This criterial is based on the loss rate neccessary to sustain a single 10Gb stream using Reno TCP computed by the Mathis formula. It is included in the LIMAN SOW.

Note, We will strive for 0 loss in our tests and work with the vendors to achieve this goal on as many circuits as possible.