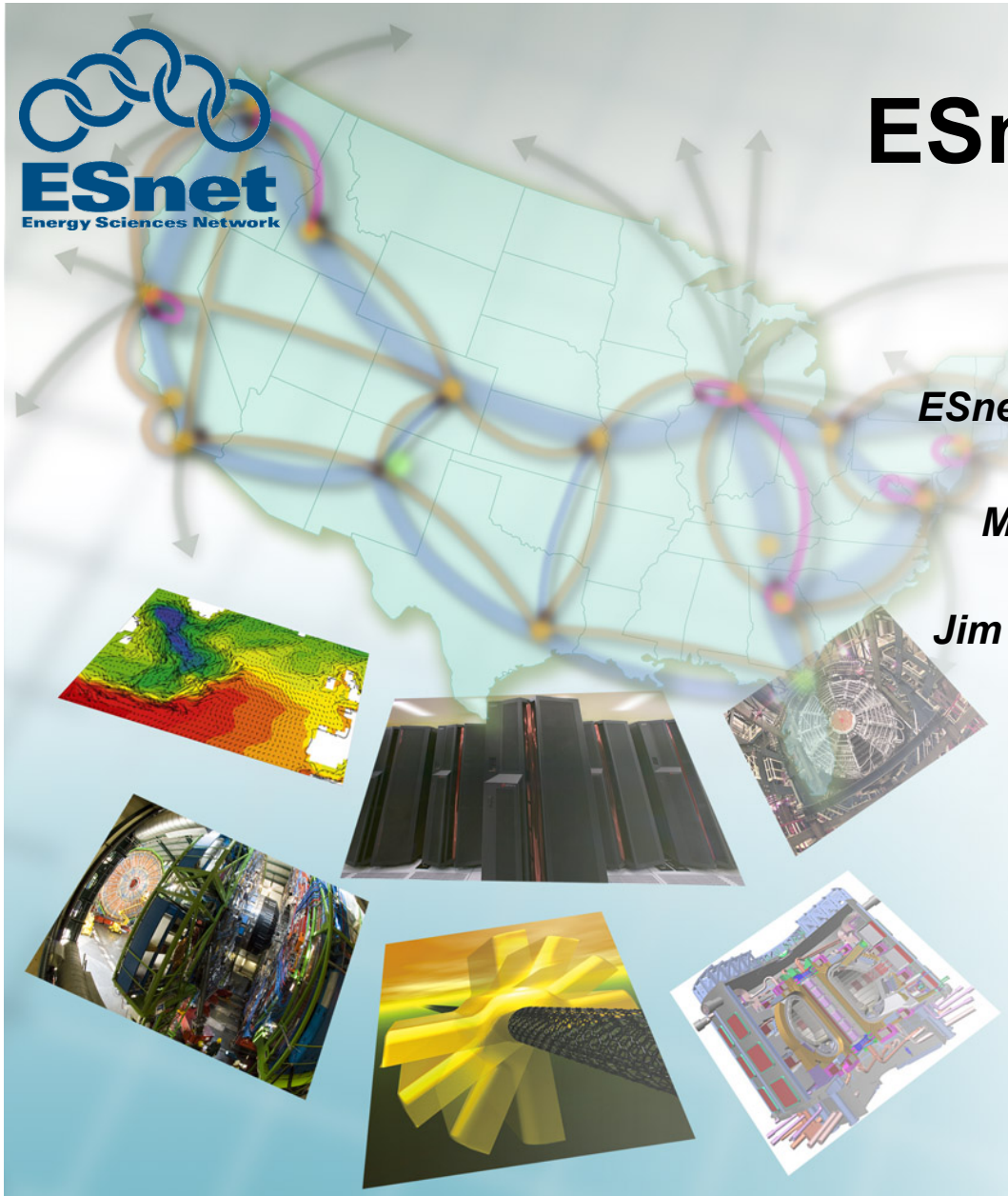


ESnet Status Update

ESCC, July 2008

William E. Johnston,
ESnet Department Head and Senior Scientist

Joe Burrescia, General Manager

Mike Collins, Chin Guok, and Eli Dart
Engineering

Jim Gagliardi, Operations and Deployment

Stan Kluz, Infrastructure and ECS

Mike Helm, Federated Trust

Dan Peterson, Security Officer

Gizella Kapus, Business Manager

and the rest of the ESnet Team

Energy Sciences Network
Lawrence Berkeley National Laboratory

wej@es.net, www.es.net

This talk is available at www.es.net/ESnet4

Networking for the Future of Science

DOE Office of Science and ESnet – the ESnet Mission

- **ESnet's primary mission is to enable the large-scale science that is the mission of the Office of Science (SC) and that depends on:**
  - Sharing of massive amounts of data
  - Supporting thousands of collaborators world-wide
  - Distributed data processing
  - Distributed data management
  - Distributed simulation, visualization, and computational steering
  - Collaboration with the US and International Research and Education community
- ESnet provides network and collaboration services to Office of Science laboratories and many other DOE programs in order to accomplish its mission

ESnet Stakeholders and their Role in ESnet

- DOE Office of Science Oversight ("SC") of ESnet
  - The SC provides high-level oversight through the budgeting process
  - Near term input is provided by weekly teleconferences between SC and ESnet
  - Indirect long term input is through the process of ESnet observing and projecting network utilization of its large-scale users
  - Direct long term input is through the SC Program Offices Requirements Workshops (more later)
- SC Labs input to ESnet
  - Short term input through many daily (mostly) email interactions
  - Long term input through ESCC

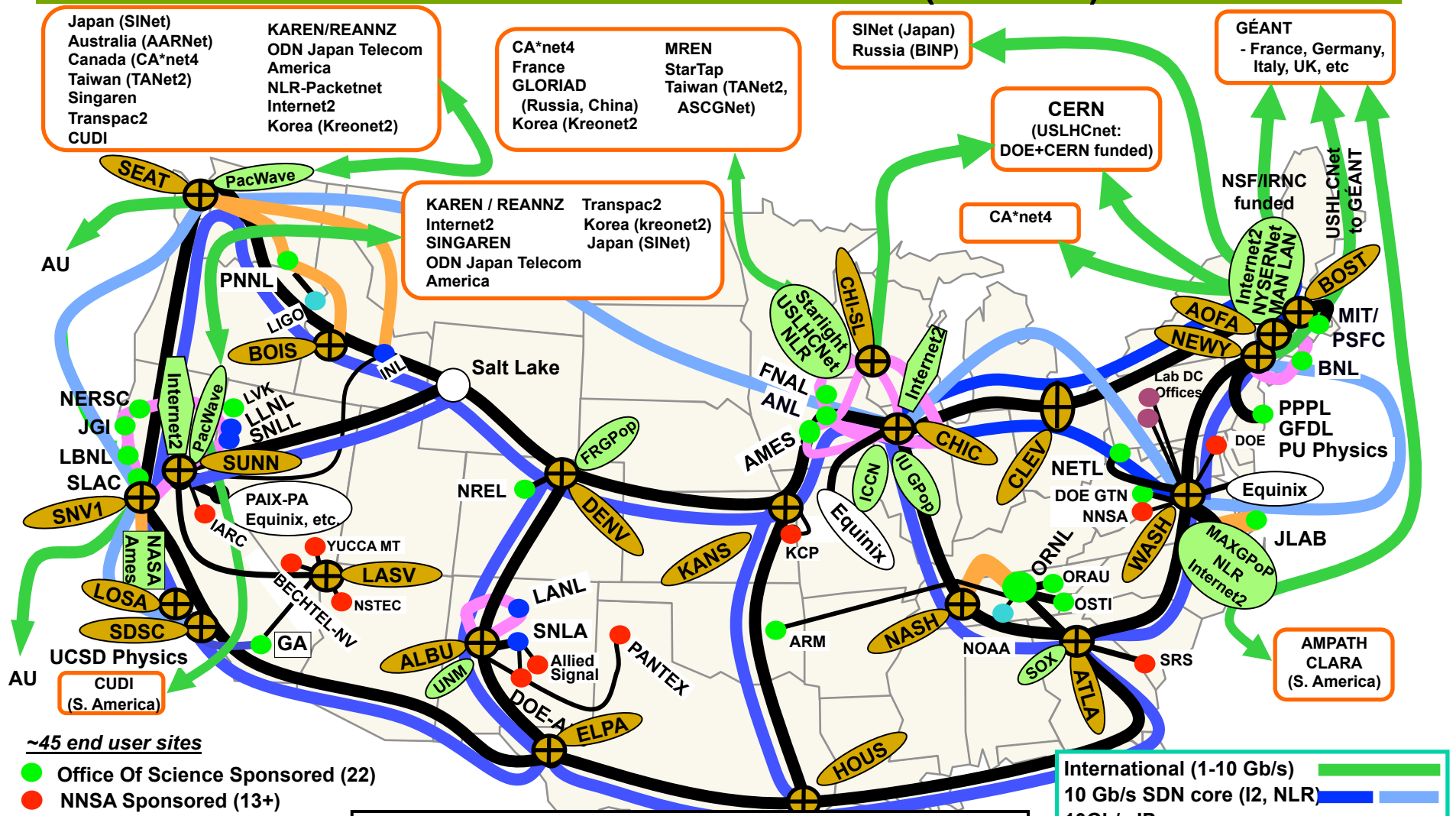
ESnet Stakeholders and the Role in ESnet

- SC science collaborators input
  - Through numerous meeting, primarily with the networks that serve the science collaborators

New in ESnet – Advanced Technologies Group / Coordinator

- Up to this point individual ESnet engineers have worked in their "spare" time to do the R&D, or to evaluate R&D done by others, and coordinate the implementation and/or introduction of the new services into the production network environment – and they will continue to do so
- In addition to this – looking to the future – ESnet has implemented a more formal approach to investigating and coordinating the R&D for the new services needed by science
  - An ESnet Advanced Technologies Group / Coordinator has been established with a twofold purpose:
    1) To provide a unified view to the world of the several engineering development projects that are on-going in ESnet in order to publicize a coherent catalogue of advanced development work going on in ESnet.
    2) To develop a portfolio of exploratory new projects, some involving technology developed by others, and some of which will be developed within the context of ESnet.
- A highly qualified Advanced Technologies lead – Brian Tierney – has been hired and funded from current ESnet operational funding, and by next year a second staff person will be added. Beyond this, growth of the effort will be driven by new funding obtained specifically for that purpose.

ESnet Provides Global High-Speed Internet Connectivity for DOE Facilities and Collaborators (12/2008)



Talk Outline

I. ESnet4

- » **Ia.** Building ESnet4
- **Ib.** Network Services – Virtual Circuits
- **Ic.** Network Services – Network Monitoring
- **Id.** Network Services – IPv6

II. SC Program Requirements and ESnet Response

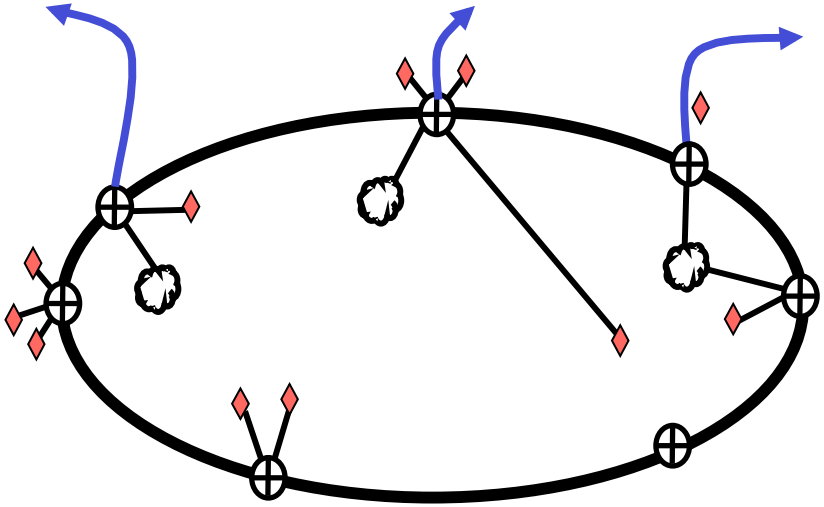
- » **IIa.** Re-evaluating the Strategy

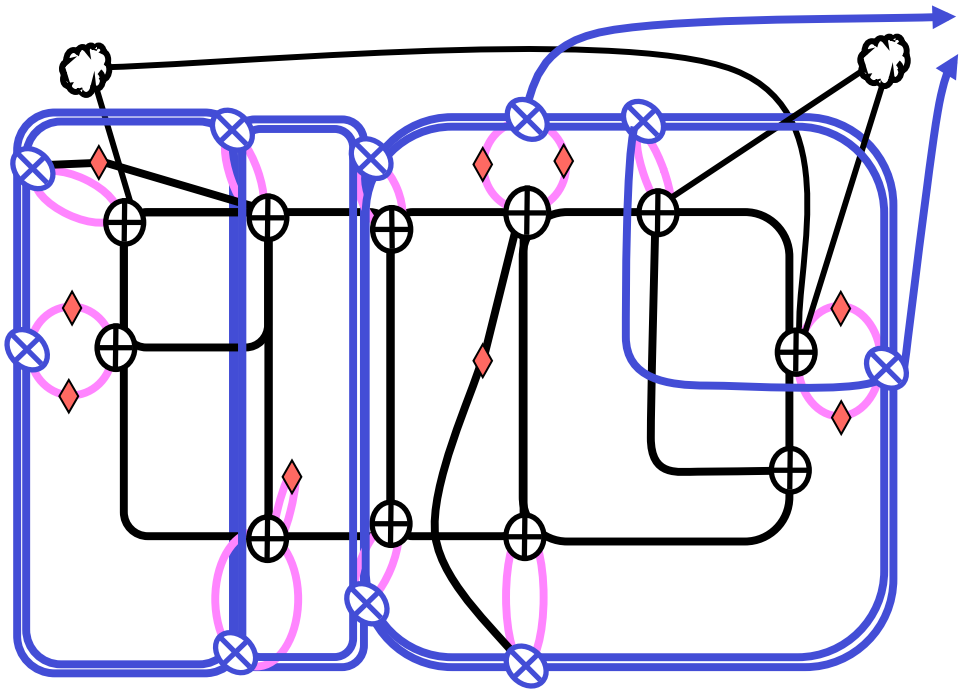
III. Science Collaboration Services

- **IIIa.** Federated Trust
- **IIIb.** Audio, Video, Data Teleconferencing
- **IIIc.** Enhanced Collaboration Services

I. ESnet4

- ESnet4 was ***built to address specific Office of Science program requirements. The result is a much more complex and much higher capacity network.***





ESnet to 2005:

- A routed IP network with sites singly attached to a national core ring
- Very little peering redundancy

ESnet4 in 2008:

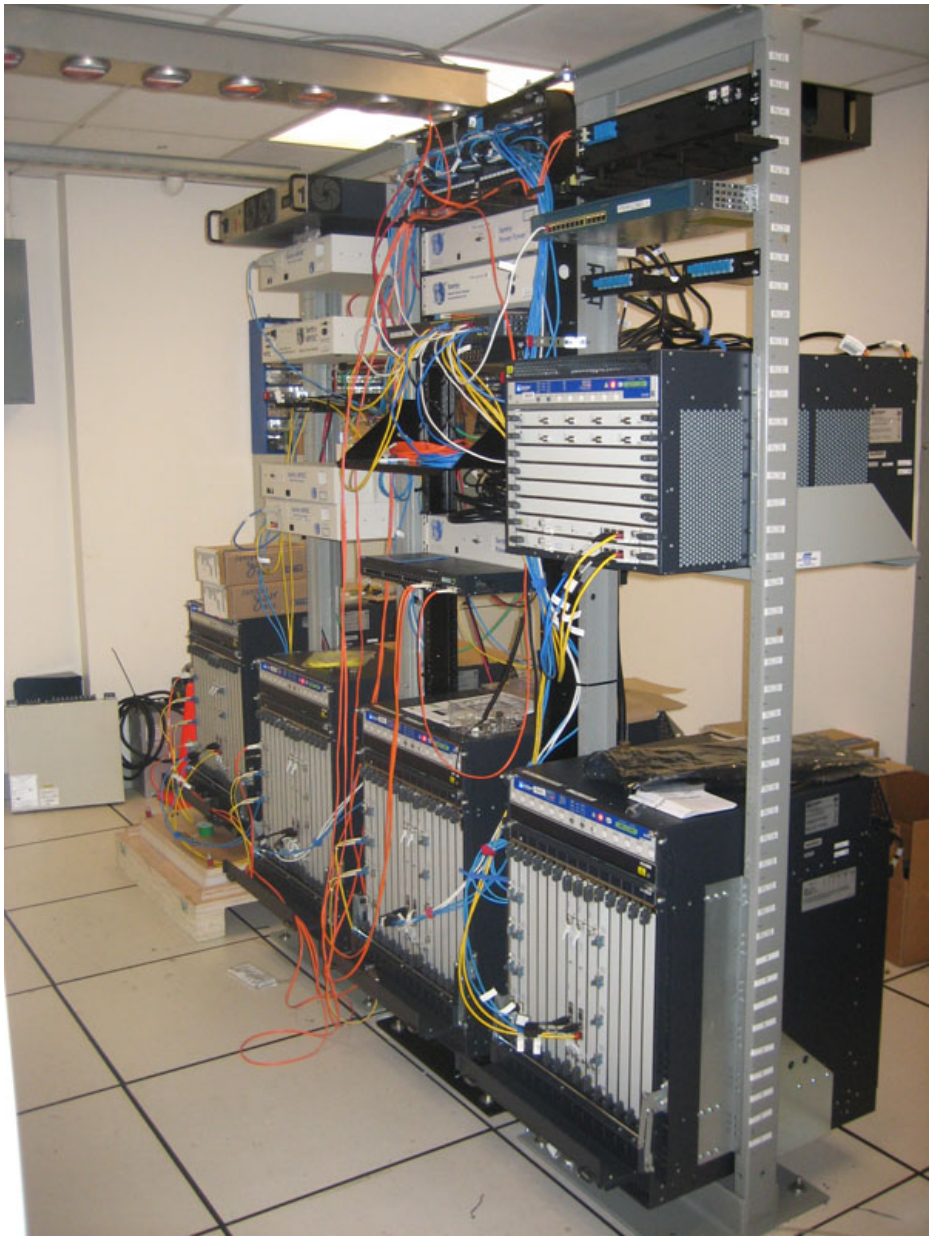
- All large science sites are dually connected on metro area rings or dually connected directly to core ring for reliability
- A switched network providing virtual circuit services for traffic engineering and guaranteed bandwidth
- Rich topology increases the reliability of the network

Ia. Building ESnet4 - SDN

State of SDN as of mid-June

(Actually, not quite, as Jim's crew had already deployed Chicago and maybe one other hub, and we were still waiting on a few Juniper deliveries.)

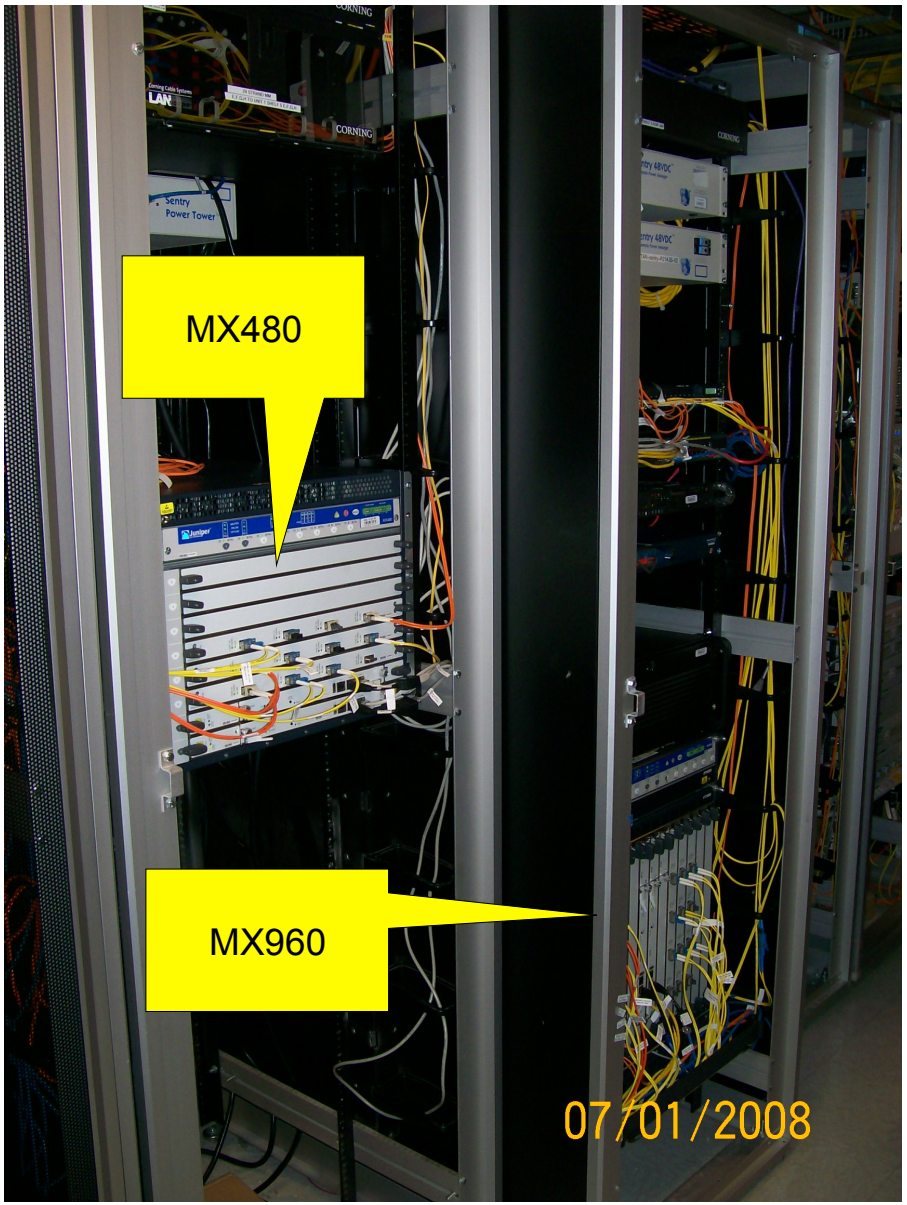
Building ESnet4 - State of SDN as of mid-July



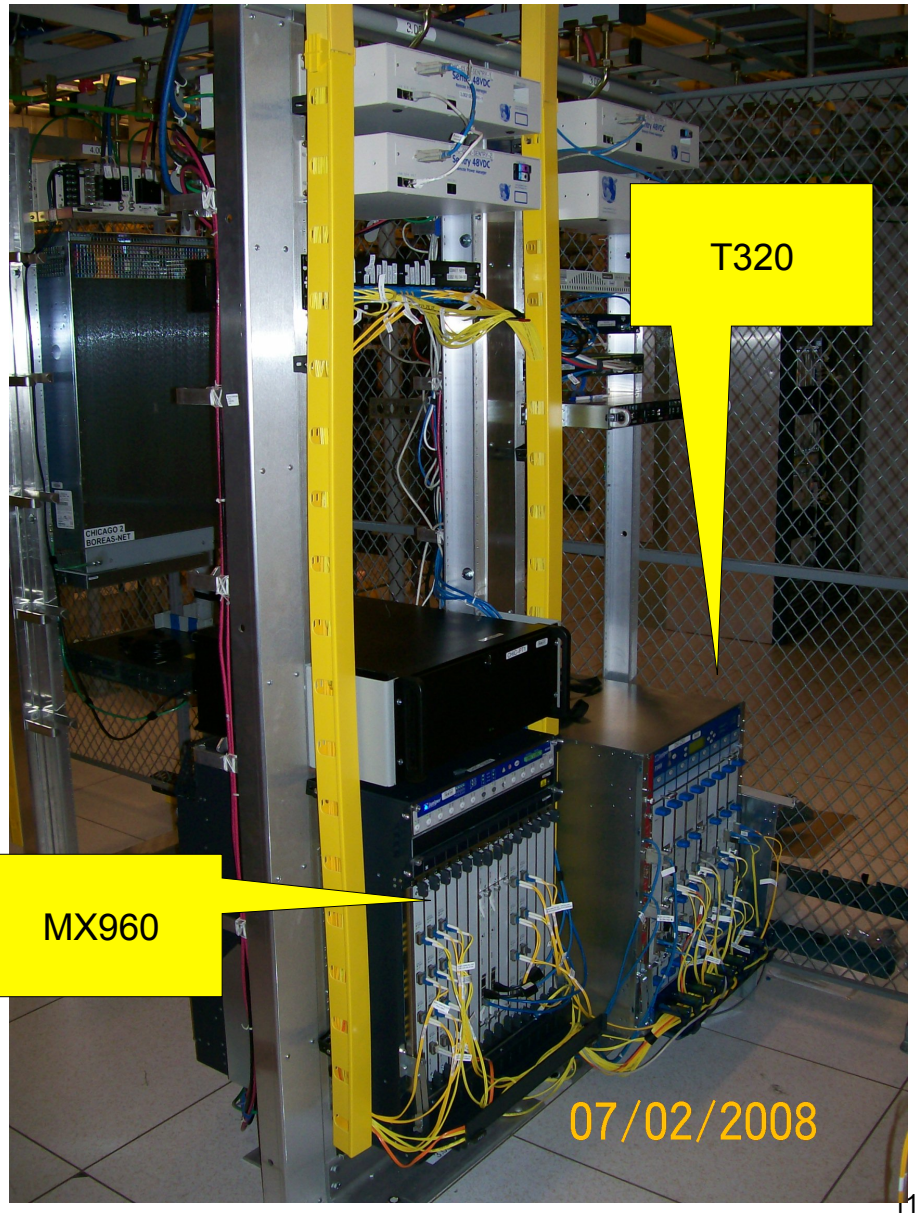
- Router/switches undergoing configuration and "burn-in" testing prior to deployment in SDN. These devices are in the ESnet configuration lab connected to dev net (ESnet in-house development network).
- The larger devices are Juniper MX960s and are for Pacific Northwest GigaPoP (Seattle), Denver, Atlanta, and Nashville.
- The smaller unit is an MX480 and is the IP core router for Kansas City
  - This device is primarily a three-way node that implements the cross-country loop to Houston – through there will probably also be a connection to NNSA's KC Plant.

ESnet4 SDN Chicago Hubs, Complete!

Starlight

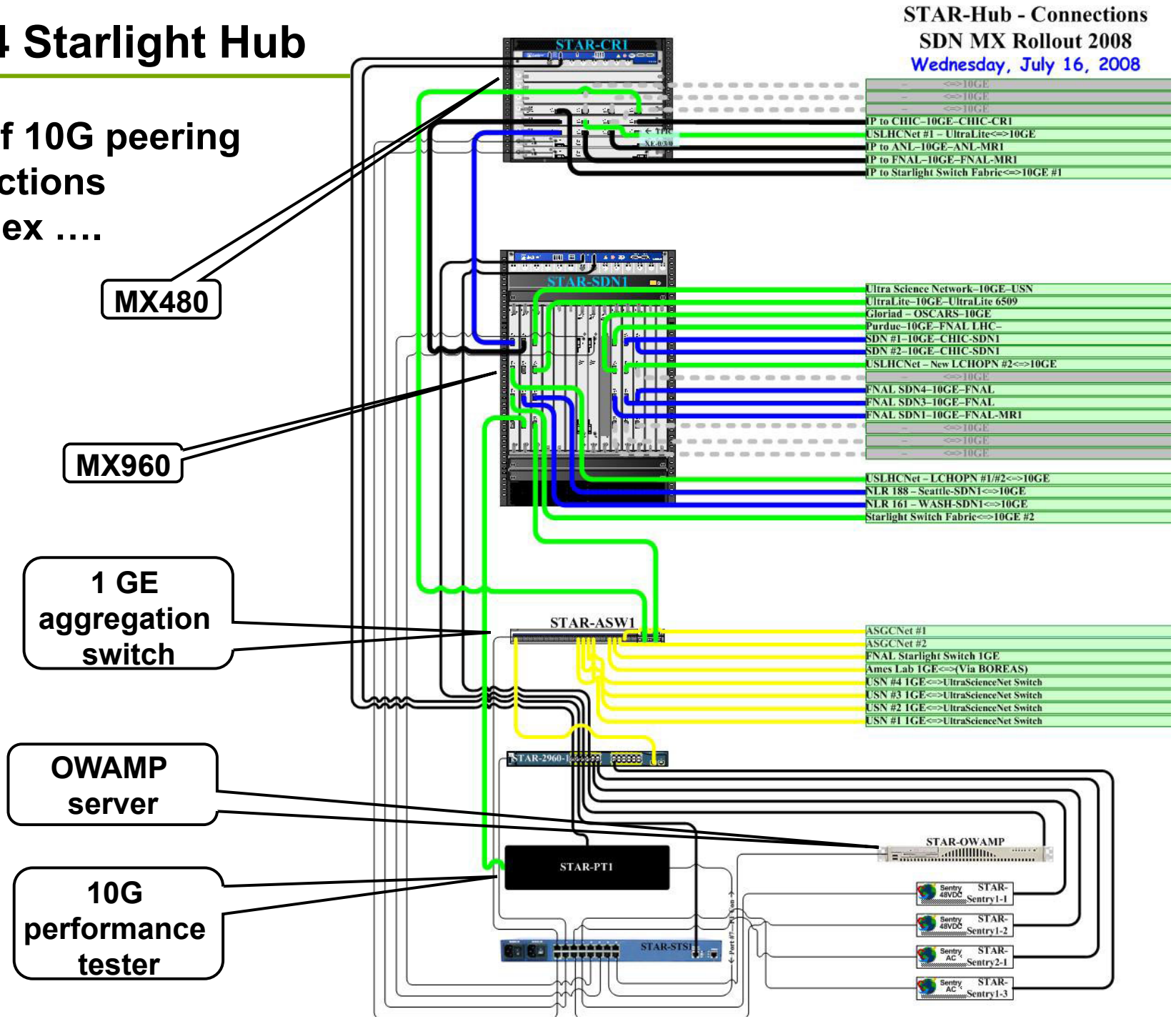


600 West Chicago (MondoCondo)



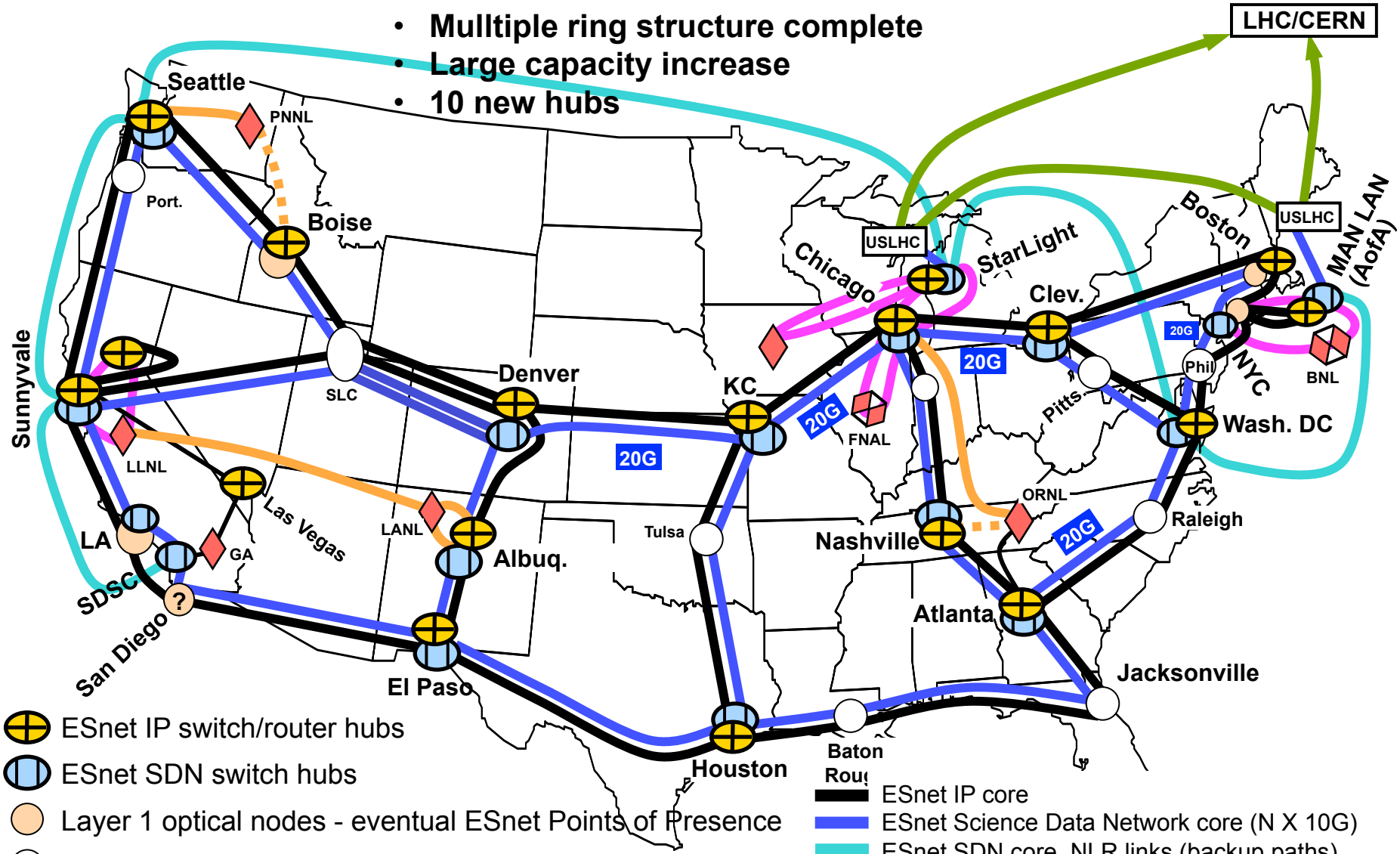
ESnet4 Starlight Hub

- A lot of 10G peering connections
- Complex ….

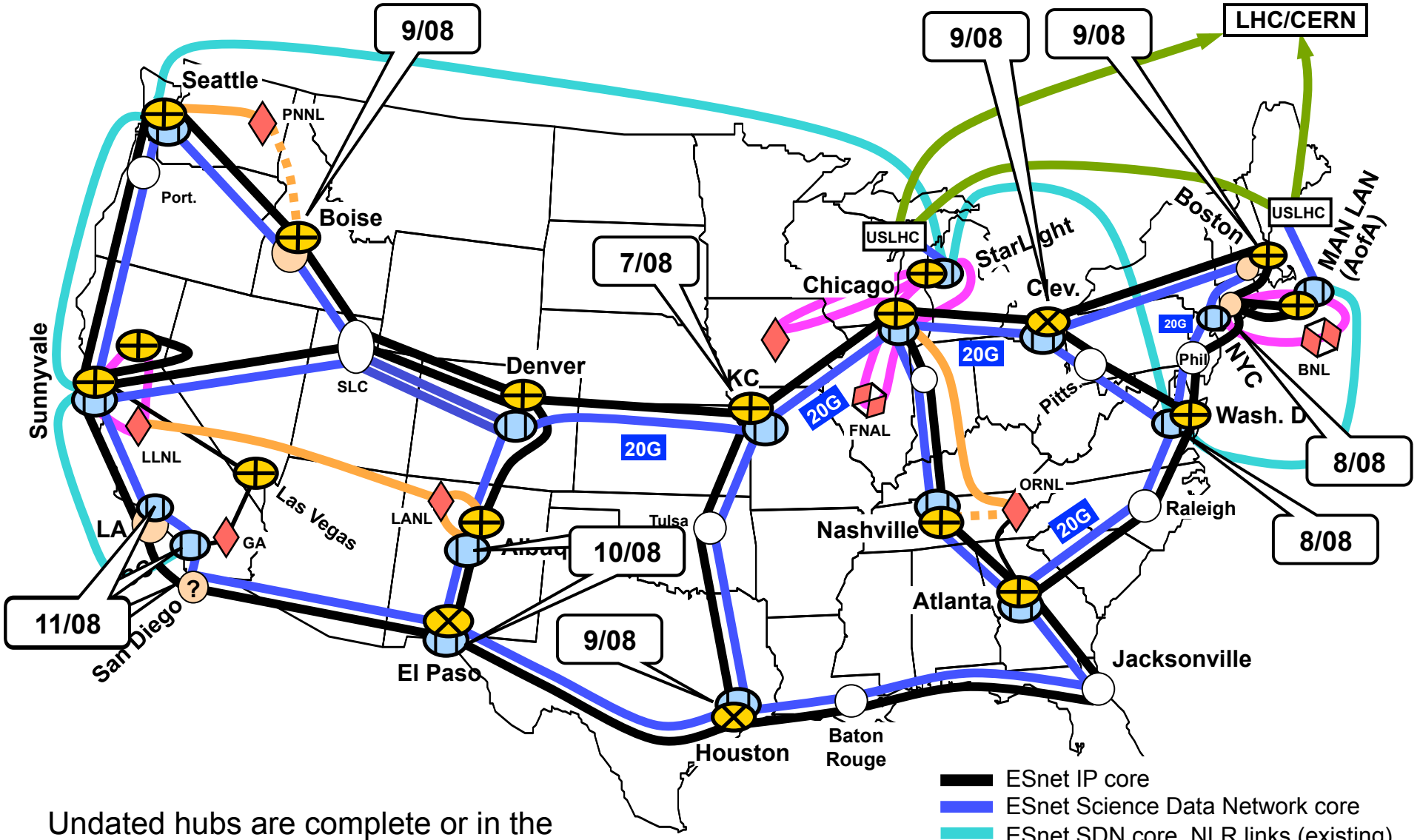


ESnet 4 Core Network – December 2008

- Mulltiple ring structure complete
- Large capacity increase
- 10 new hubs

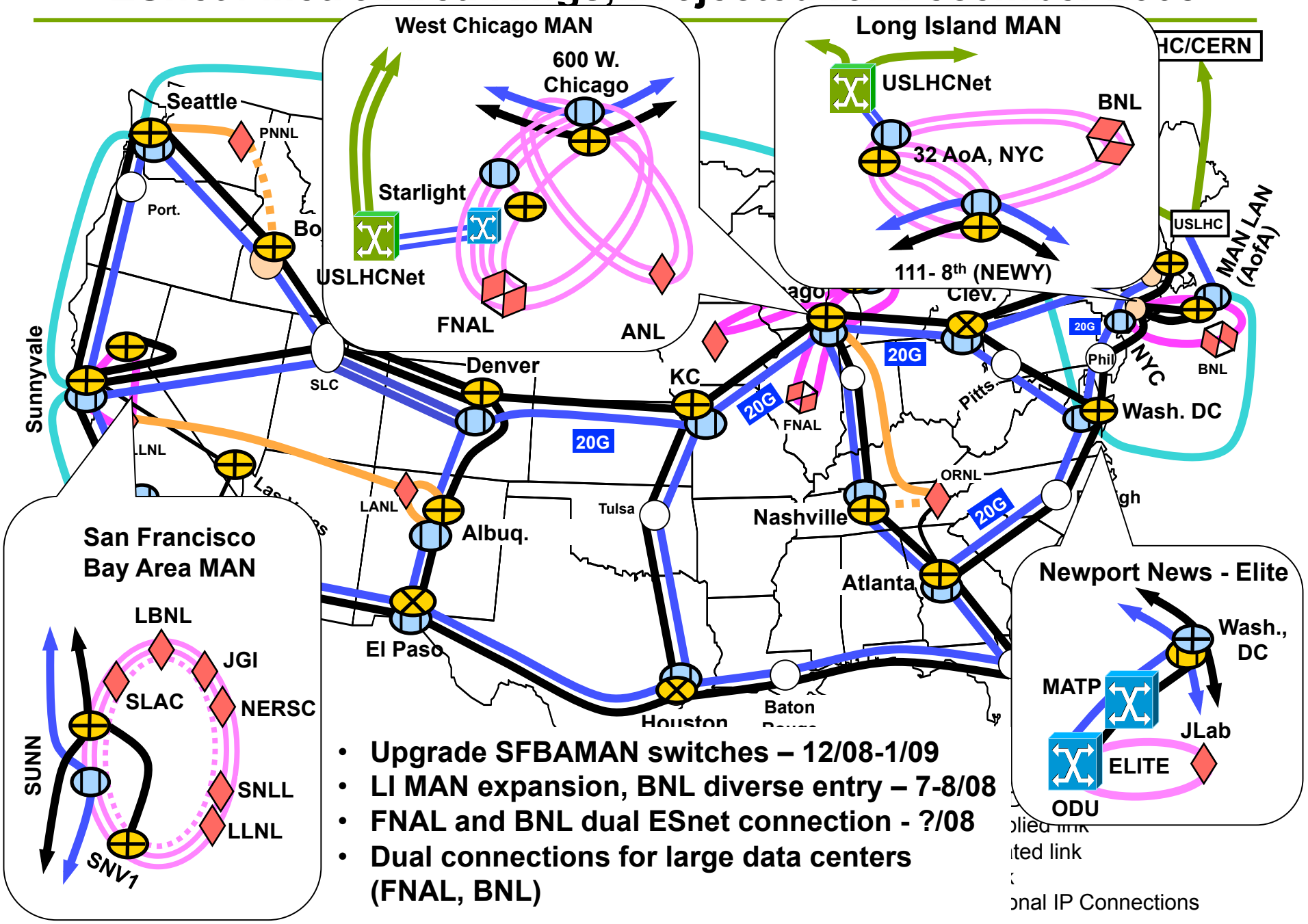


Deployment Schedule Through December 2008



Undated hubs are complete or in the process of being installed now

ESnet4 Metro Area Rings, Projected for December 2008



- Upgrade SFBAMAN switches – 12/08-1/09
- LI MAN expansion, BNL diverse entry – 7-8/08
- FNAL and BNL dual ESnet connection - ?/08
- Dual connections for large data centers (FNAL, BNL)

ted link

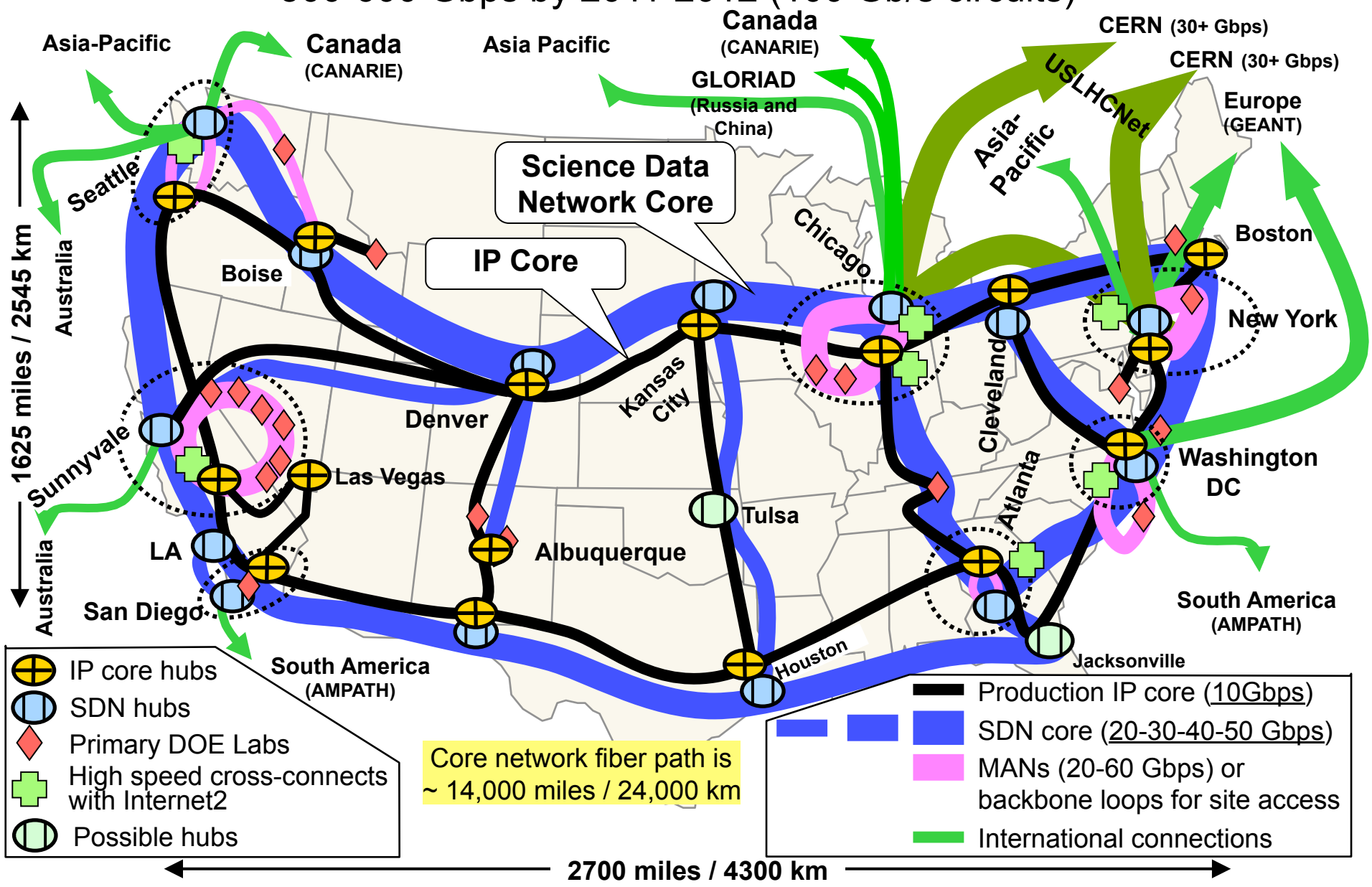
onal IP Connections

ESnet 4 SDN Factoids as of July 22, 2008

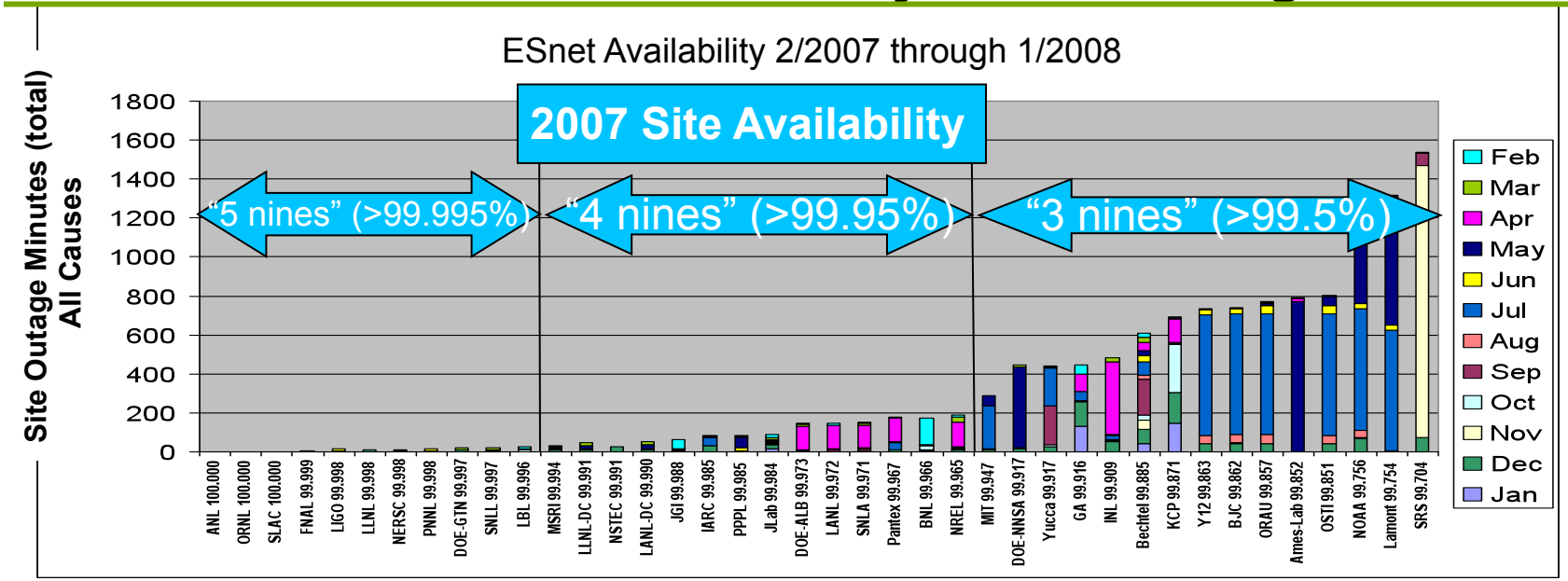
- ESnet4 SDN installation to date:
  - 10 10Gb/s backbone segments in production
  - 32 new MX series switches received
    - 4 installed so far…
    - ATLA, NASH, DENV &PNWG shipped this week
  - Enhanced hubs
    - Sunnyvale
      - 17 total connections moved including 12 10G connections
      - 6509 Removed
      - MX960 added
    - Starlight
      - 32 total connections moved including 23 10G connections
      - T320 & 6509 removed
      - MX960&480 added
    - CHIC (600 W Chicago)
      - 13 total connections moved including 12 10G connections
      - 7609 removed
      - MX960 added

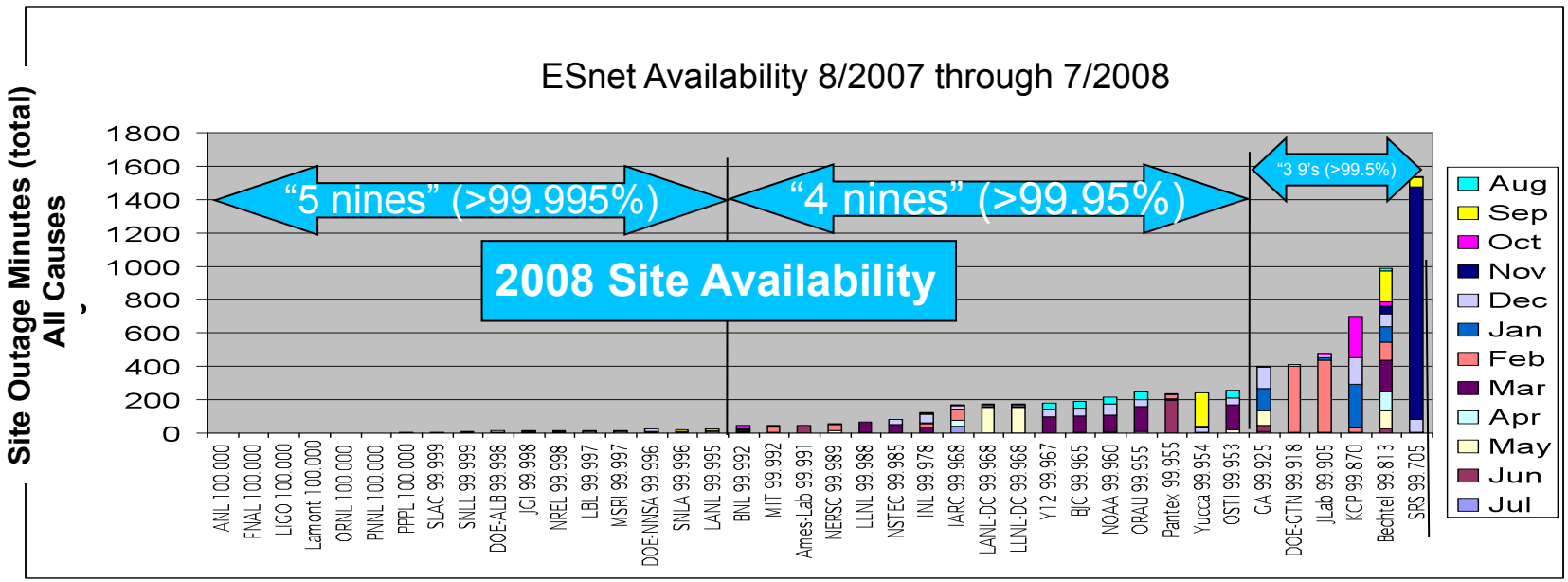
ESnet4 End-Game – 2012

Core networks 50-60 Gbps by 2009-2010 (10Gb/s circuits),
500-600 Gbps by 2011-2012 (100 Gb/s circuits)

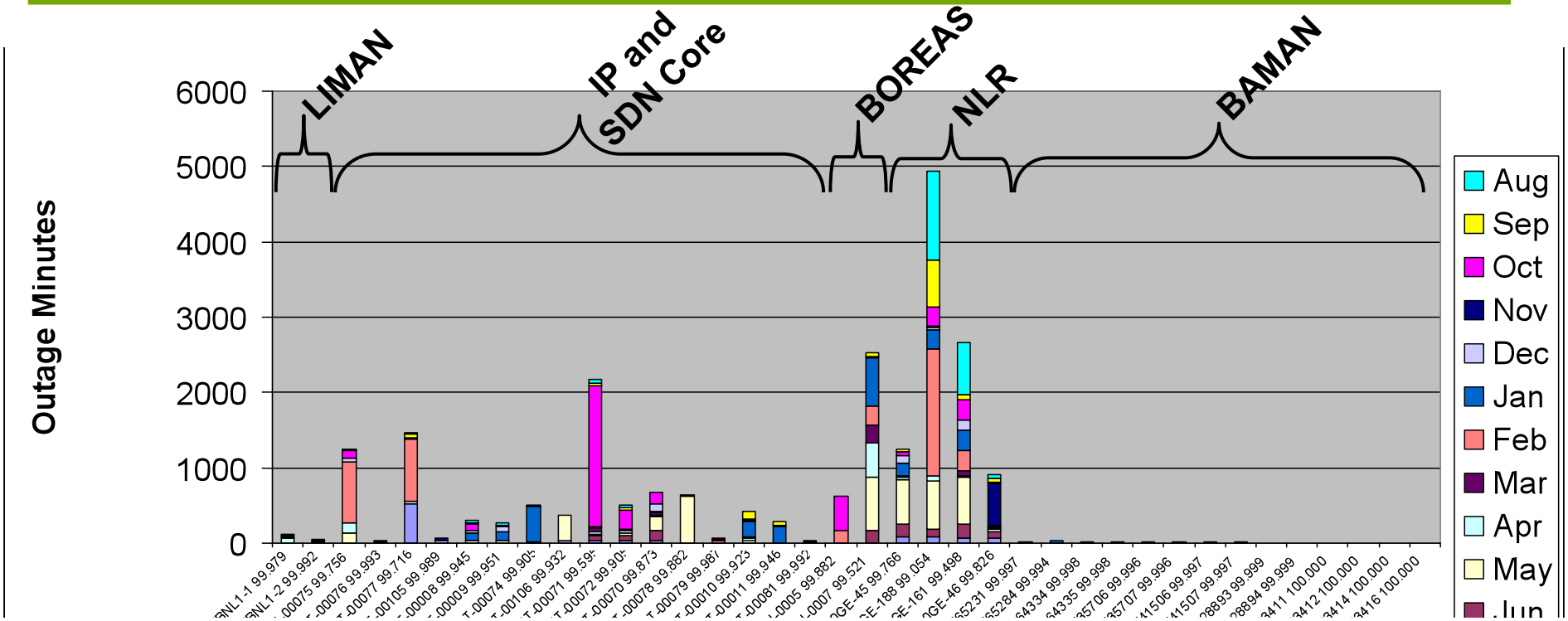


ESnet's Availability is Increasing





ESnet Carrier Circuit Outages, 8/2007-7/2008



- **IP and SDN core (Internet2 – Infinera – Level3 optical network):**
  - **The outages are understood and cause addressed – not expected to be chronic**
  - **All outages were on rings that provided redundancy so *no site impact***
- **NLR:**
  - **These outages follow a several year pattern and appear to be chronic**
  - **NLR circuit is linear (no redundancy) and mostly used in various backup strategies – no production use – *no site impact***
- **BOREAS:**
  - **Highway construction is impacting one side of the ring**
  - **Ring provides redundancy (Ames is the only ESnet site on this ring) so *no site impact***

Ib. Network Services – Virtual Circuits

Fairly consistent requirements are found across the large-scale sciences

- ***Large-scale science uses distributed systems*** in order to:
  - Couple existing pockets of code, data, and expertise into "systems of systems"
  - Break up the task of massive data analysis into elements that are physically located where the data, compute, and storage resources are located - these elements are combined into a system using a "Service Oriented Architecture" approach
- Such ***systems***
  - ***are data intensive and high-performance***, typically moving terabytes a day for months at a time
  - ***are high duty-cycle***, operating most of the day for months at a time in order to meet the requirements for data movement
  - ***are widely distributed*** – typically spread over continental or inter-continental distances
  - ***depend on network performance and availability***, but these characteristics cannot be taken for granted, even in well run networks, when the multi-domain network path is considered

➢ The system elements ***must be able to get guarantees from the network*** that there is adequate bandwidth to accomplish the task at hand

➢ The systems ***must be able to get information from the network*** that allows graceful failure and auto-recovery and adaptation to unexpected network conditions that are short of outright failure

See, e.g., [ICFA SCIC]

To Support Large-Scale Science Networks Must Provide Communication Capability that is Service-Oriented

- Configurable
  - Must be able to provide multiple, specific “paths” (specified by the user as end points) with specific characteristics
- Schedulable
  - Premium service such as guaranteed bandwidth will be a scarce resource that is not always freely available, therefore time slots obtained through a resource allocation process must be schedulable
- Predictable
  - A committed time slot should be provided by a network service that is not brittle - reroute in the face of network failures is important
- Reliable
  - Reroutes should be largely transparent to the user
- Informative
  - When users do system planning they should be able to see average path characteristics, including capacity
  - When things do go wrong, the network should report back to the user in ways that are meaningful to the user so that informed decisions can about alternative approaches
- Scalable
  - The underlying network should be able to manage its resources to provide the appearance of scalability to the user
- Geographically comprehensive
  - The R&E network community must act in a coordinated fashion to provide this environment end-to-end

The ESnet Approach for Required Capabilities

- Provide ***configurability, schedulability, predictability***, and reliability with a flexible virtual circuit service - OSCARS
  - User* specifies end points, bandwidth, and schedule
  - OSCARS can do fast reroute of the underlying MPLS paths
- Provide useful, comprehensive, and meaningful ***information on the state of the paths***, or potential paths, to the user
  - perfSONAR, and associated tools, provide real time information in a form that is useful to the user (via appropriate abstractions) and that is delivered through standard interfaces that can be incorporated in to SOA type applications **R&D**
  - Techniques need to be developed to monitor virtual circuits based on the approaches of the various R&E nets - e.g. MPLS in ESnet, VLANs, TDM/grooming devices (e.g. Ciena Core Directors), etc., and then integrate this into a perfSONAR framework **R&D**

* User = human or system component (process)

The ESnet Approach for Required Capabilities

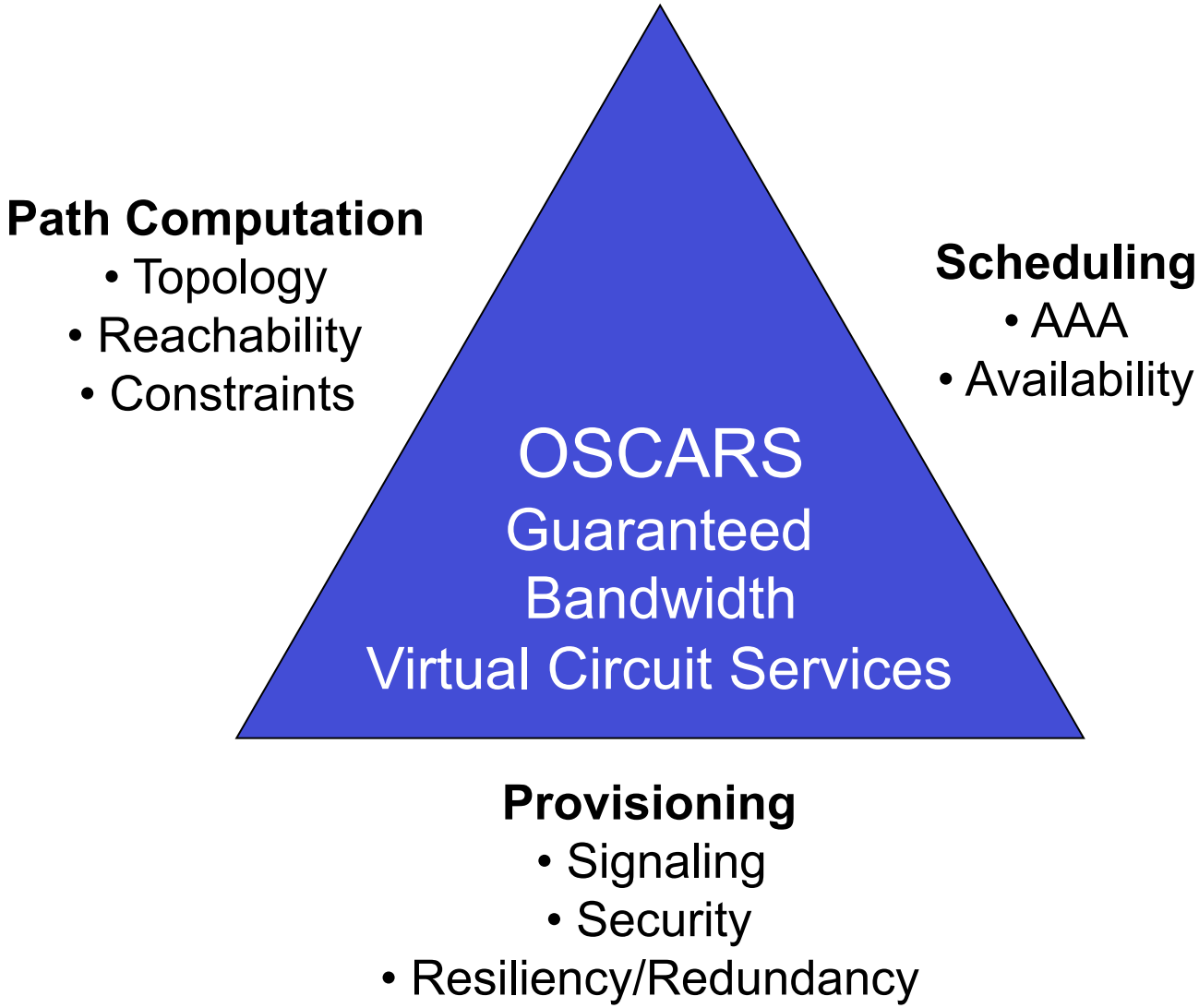
- ***Reliability*** approaches for Virtual Circuits are currently under investigation and are topics for R&D **R&D**
- ***Scalability*** will be provided by new network services that, e.g., provide dynamic wave allocation at the optical layer of the network **R&D**
- ***Geographic ubiquity*** of the services can only be accomplished through active collaborations in the global R&E network community so that all sites of interest to the science community can provide compatible services for forming end-to-end virtual circuits
  - Active and productive collaborations exist among numerous R&E networks: ESnet, Internet2, Caltech, DANTE/GÉANT, some European NRENs, some US regionals, etc.

The ESnet Approach for Required Capabilities

- User experience in the first year of OSCARS operation has revealed several new capabilities that are required
  - The usefulness of permitting over subscribing a path is needed to
    - accommodate backup circuits
    - allow for site managed load balancing
  - It is becoming apparent that there is a need to direct routed IP traffic onto SDN in a way transparent to the user **R&D**
    - Many issues here
  - More on these in the OSCARS section and talk

OSCARS Overview

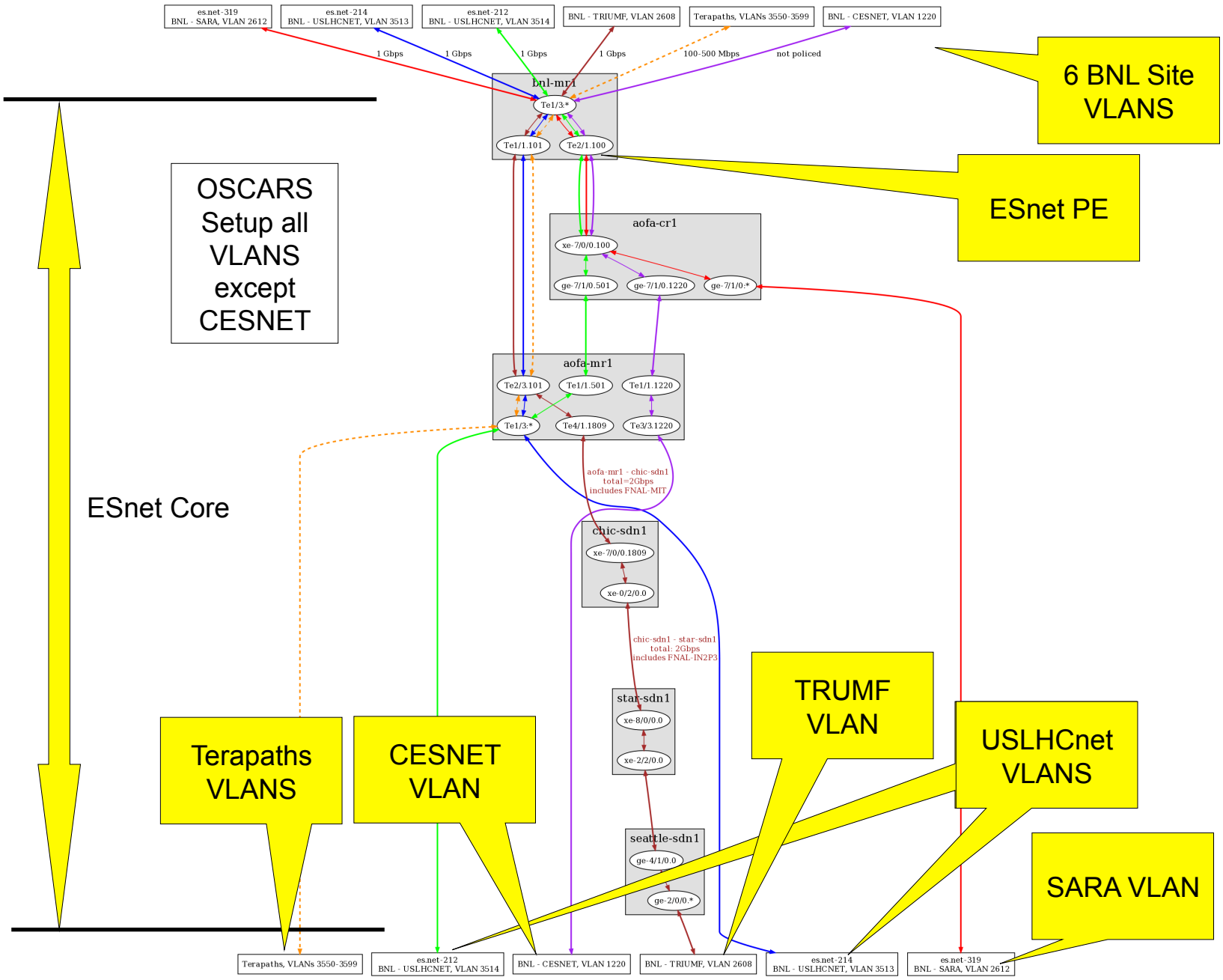
On-demand Secure Circuits and Advance Reservation System



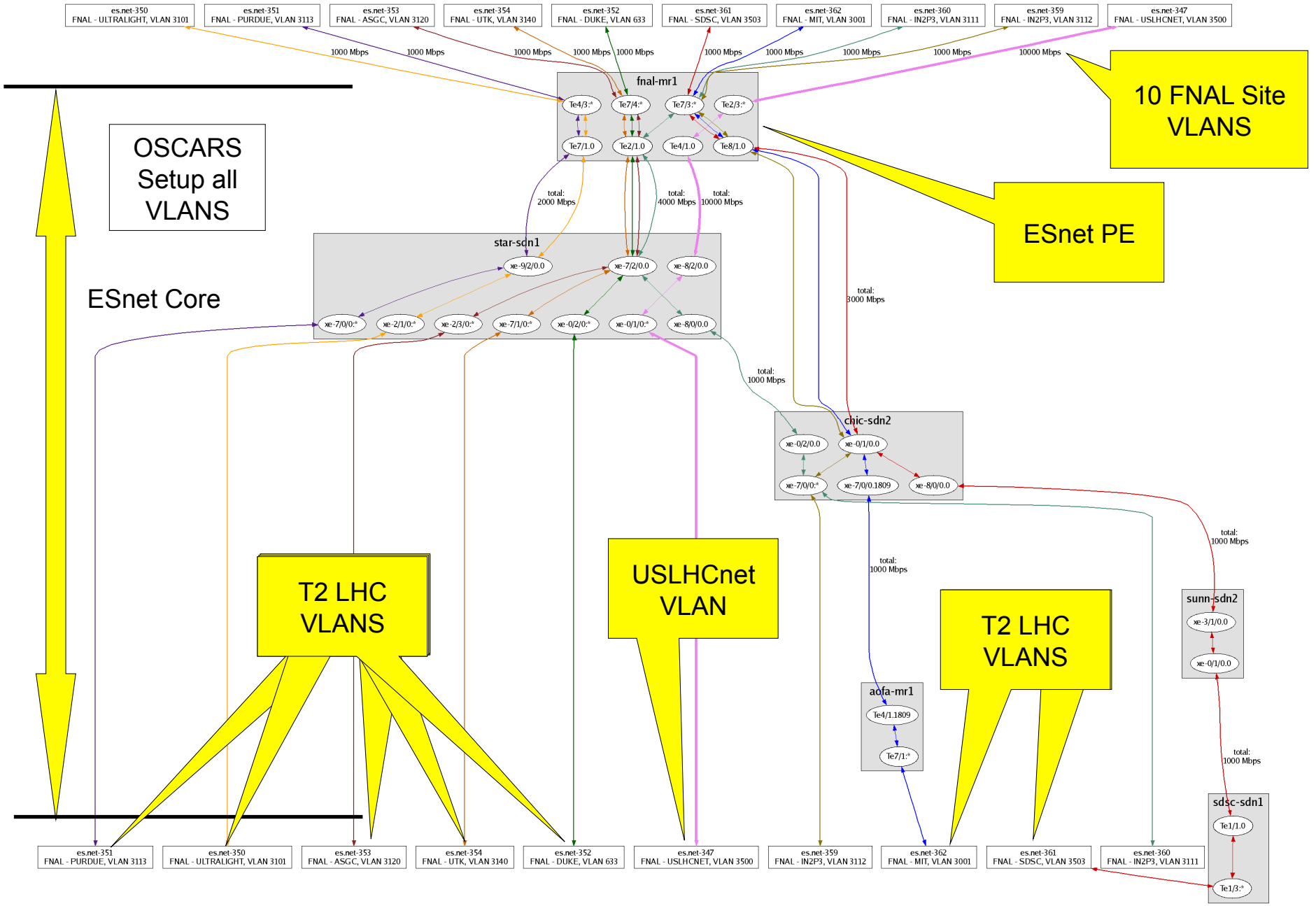
OSCARS Status Update

- ESnet Centric Deployment
  - Prototype layer 3 (IP) guaranteed bandwidth virtual circuit service deployed in ESnet (1Q05)
  - Prototype layer 2 (Ethernet VLAN) virtual circuit service deployed in ESnet (3Q07)
  - Support soft reservations (2Q08)
  - Automatic graph generation of VCs (2Q08)
  - Support site administrator role (2Q08)
- Inter-Domain Collaborative Efforts
  - Terapaths
    - Inter-domain interoperability for layer 3 virtual circuits demonstrated (3Q06)
    - Inter-domain interoperability for layer 2 virtual circuits demonstrated at SC07 (4Q07)
  - LambdaStation
    - Inter-domain interoperability for layer 2 virtual circuits demonstrated at SC07 (4Q07)
  - I2 DCN/DRAGON
    - Inter-domain exchange of control messages demonstrated (1Q07)
    - Integration of OSCARS and DRAGON has been successful (1Q07)
  - GEANT2 AutoBAHN
    - Inter-domain reservation demonstrated at SC07 (4Q07)
  - DICE
    - First draft of topology exchange schema has been formalized (in collaboration with NMWG) (2Q07), interoperability test demonstrated 3Q07
    - Initial implementation of reservation and signaling messages demonstrated at SC07 (4Q07)
  - Nortel
    - Topology exchange demonstrated successfully 3Q07
    - Inter-domain interoperability for layer 2 virtual circuits demonstrated at SC07 (4Q07)
  - UVA
    - Demonstrated token based authorization concept with OSCARS at SC07 (4Q07)
  - OGF NML-WG
    - Actively working to combine work from NMWG and NDL
    - Documents and UML diagram for base concepts have been drafted (2Q08)
  - GLIF GNI-API WG
    - In process of designing common API and reference middleware implementation

OSCARS Managed External Circuit Topology at BNL



OSCARS Managed External Circuit Topology at FNAL



OSCARS Adapting to User Experience

- Original design capabilities
  - Guaranteed bandwidth VCs
    - ***Over-provisioning of the overall SDN path is prevented at reservation request time***
      - i.e. each new reservation request is vetted against available capacity for the entire duration of the reservations ⟹ dynamic updates of not just the current reserved bandwithd loaded topology, but into the future as well (this is key to ensuring sufficient bandwidth for all VC guarantees)
    - ***Over-subscription (once the VC is in use) is prevented by policing (hard drop) at time of use***
    - All reserved VCs configured to transit ESnet as Expedited Forwarding Class traffic

OSCARS Adapting to User Experience

- Current updated capabilities
  - Guaranteed Bandwidth VC with Path Over-Subscription
    - ***Over-provisioning of overall path is still prevented at reservation request time***
    - ***Over-subscription is allowed during VC use***
      - Traffic below policed rate will transit ESnet as Expedited Forwarding Class
      - Traffic above policed rate is not dropped, but remarked as Scavenger Class (so this traffic only moves of there is unutilized bandwidth on the path)
        - » This allows sites to provision multiple VCs along the same path and manage the use of these locally
  - Considerations
    - Implementation of above enhancements are technology specific
      - not all network implementations have multiple forwarding classes (multiple traffic priorities
    - End-to-end inter-domain dynamic VCs may not support over-subscription
    - Multi-lab coordination may be required to effective utilize bandwidth available in Scavenger Class

Ic. Network Services – Network Measurement

- ESnet
  - Goal is to have ***10G testers and latency measuring*** capabilities at all hubs
    - About 1/3 of the 10GE bandwidth test platforms & 1/2 of the latency test platforms for ESnet 4 have been deployed.
      - 10GE test systems are being used extensively for acceptance testing and debugging
      - Structured & ad-hoc external testing capabilities will be enabled soon.
  - Work is progressing on ***revamping the ESnet statistics collection***, management & publication systems
    - ESxSNMP & TSDB & PerfSONAR Measurement Archive (MA)
    - PerfSONAR TS & OSCARS Topology DB
    - NetInfo being restructured to be PerfSONAR based
- LHC and PerfSONAR
  - PerfSONAR based network measurement pilot for the Tier 1/Tier 2 community is ready for deployment.
  - A proposal from DANTE to deploy a perfSONAR based network measurement service across the LHCOPN at all Tier1 sites is still being evaluated by the Tier 1 centers

Network Services – IPv6

- ESnet provides production IPv6 service
  - IPv6 fully supported by ESnet NOC and engineering
  - IPv6 supported natively by ESnet routers
  - http://www.es.net/hypertext/IPv6/index.html
- Network-level IPv6 services include:
  - Address allocation for sites
    - Some sites have already been assigned IPv6 space
    - More are welcome!
  - Full IPv6 connectivity (default-free IPv6 routing table)
    - High-speed R&E peerings with Americas, Europe, Canada, Asia
    - Numerous commodity Internet IPv6 peerings as well
  - Diverse IPv6 peering with root name servers

Routine Use of IPv6 by ESnet

- IPv6 support services
  - ESnet web, mail and DNS servers are fully IPv6 capable
  - ESnet has a Stratum 1 IPv6 time (NTP) server per coast
  - Open source software mirrors – FreeBSD and Linux
    - Open IPv6 access
    - See http://www.es.net/hypertext/IPv6/ipv6-mirror-servers.html
  - ESnet staff use IPv6 to access these services on a routine basis
- Future plans for IPv6 enabled services
  - perfSONAR
  - Performance testers

ESnet - a production IPv6 network

In anticipation of the scalability problems with IPv4 (the current Internet Protocol), the Internet Engineering Task Force (IETF) has produced a comprehensive set of specifications that define the next generation Internet Protocol known as IPv6.

ESnet has a long history with IPv6 - ESnet participated extensively in early IPv6 testing and deployment, and received the first production IPv6 address allocation in July 1999. ESnet declared IPv6 to be a fully supported production service in August 2002. ESnet's web, mail and DNS servers are fully IPv6 capable, and ESnet staff use IPv6 for routine tasks every day.

ESnet IPv6 Services

ESnet has several services available to aid in the deployment of IPv6 by ESnet sites. These include IPv6 address space, open source software mirrors available via IPv6, and IPv6 time servers. ESnet engineers maintain full global IPv6 connectivity, just as for IPv4.

- To request IPv6 address space for an ESnet site, the ESnet site coordinator should contact ESnet through established channels - IPv6 service is available to any ESnet site that requests it
- ESnet maintains open-access IPv6-enabled open source software mirrors - please see the mirror server documentation for details
- Instructions on using ESnet IPv6 time servers are available to site coordinators
- IPv6 is fully supported by the ESnet NOC and engineering staff

New
See www.es.net – "network services" tab – IPv6 link

ESnet IPv6 Peering

ESnet has an IPv6 presence at several commercial and Research and Education exchanges. Please see our peering page for more information.

ESnet IPv6 History

ESnet was an early participant in the 6bone IPv6 testbed project, and provided overall management for the 6bone until the 6bone was decommissioned. ESnet also received the first production IPv6 address allocation in July 1999. More information is available on our IPv6 history page.

II. SC Program Requirements and ESnet Response

Recall the Planning Process

- Requirements are determined by
  1) Exploring the plans of the major stakeholders:
     - 1a) Data characteristics of instruments and facilities
       - What data will be generated by instruments coming on-line over the next 5-10 years (including supercomputers)?
     - 1b) Examining the future process of science
       - How and where will the new data be analyzed and used – that is, how will the process of doing science change over 5-10 years?
  2) Observing traffic patterns
     - What do the trends in network patterns predict for future network needs?
- The assumption has been that you had to add 1a) and 1b) (future plans) to 2) (observation) in order to account for unpredictable events – e.g. the turn-on of major data generators like the LHC

(1a) Requirements from Instruments and Facilities

Network Requirements Workshops

- Collect requirements from two DOE/SC program offices per year
- ESnet requirements workshop reports: http://www.es.net/hypertext/requirements.html
- Workshop schedule
  - BES (2007 – published)
  - BER (2007 – published)
  - FES (2008 – published)
  - NP (2008 – published)
  - ASCR (Spring 2009)
  - HEP (Summer 2009)
- Future workshops - ongoing cycle
  - BES, BER – 2010
  - FES, NP – 2011
  - ASCR, HEP – 2012
  - (and so on...)

Requirements from Instruments and Facilities

- Typical DOE large-scale facilities are the Tevatron accelerator (FNAL), RHIC accelerator (BNL), SNS accelerator (ORNL), ALS accelerator (LBNL), and the supercomputer centers: NERSC, NCLF (ORNL), Blue Gene (ANL)
- These are representative of the 'hardware infrastructure' of DOE science
- ***Requirements from these can be characterized*** as
  - ***Bandwidth***: Quantity of data produced, requirements for timely movement
  - ***Connectivity***: Geographic reach – location of instruments, facilities, and users plus network infrastructure involved (e.g. ESnet, Abilene, GEANT)
  - ***Services***: Guaranteed bandwidth, traffic isolation, etc.; IP multicast

(1b) Requirements from Case Studies on Process of Science

Case studies on how science involving data is done now, and how the science community sees it as changing, were initially done for a fairly "random," but we believe them to be representative, set of facilities and collaborations.

- Advanced Scientific Computing Research (ASCR)
  - NERSC (LBNL) (supercomputer center)
  - NLCF (ORNL) (supercomputer center)
  - ACLF (ANL) (supercomputer center)
- Basic Energy Sciences
  - Advanced Light Source
    - Macromolecular Crystallography
  - Chemistry/Combustion
  - Spallation Neutron Source
- Biological and Environmental
  - Bioinformatics/Genomics
  - Climate Science
- Fusion Energy Sciences
  - Magnetic Fusion Energy/ITER
- High Energy Physics
  - LHC
- Nuclear Physics
  - RHIC (heavy ion accelerator)

Network Requirements Workshops - Findings

- Virtual circuit services (traffic isolation, bandwidth guarantees, etc) continue to be requested by scientists
  - OSCARS service directly addresses these needs
    - http://www.es.net/OSCARS/index.html
    - Successfully deployed in early production today
    - ESnet will continue to develop and deploy OSCARS
- Some user communities have significant difficulties using the network for bulk data transfer
  - fasterdata.es.net – web site devoted to bulk data transfer, host tuning, etc. established
  - NERSC and ORNL have made significant progress on improving data transfer performance between supercomputer centers

Network Requirements Workshops - Findings

- Some data rate requirements are unknown at this time
  - Drivers are instrument upgrades that are subject to review, qualification and other decisions that are 6-12 months away
  - These will be revisited in the appropriate timeframe

Science Network Requirements Aggregation Summary

Science Drivers Science Areas / Facilities	End2End Reliability	Near Term End2End Band width	5 years End2End Band width	Traffic Characteristics	Network Services
ASCR: ALCF	-	10Gbps	30Gbps	• Bulk data • Remote control • Remote file system sharing	• Guaranteed bandwidth • Deadline scheduling • PKI / Grid
ASCR: NERSC	-	10Gbps	20 to 40 Gbps	• Bulk data • Remote control • Remote file system sharing	• Guaranteed bandwidth • Deadline scheduling • PKI / Grid
ASCR: NLCF	-	Backbone Bandwidth Parity	Backbone Bandwidth Parity	•Bulk data •Remote control •Remote file system sharing	• Guaranteed bandwidth • Deadline scheduling • PKI / Grid
BER: Climate	-	3Gbps	10 to 20Gbps	• Bulk data • Rapid movement of GB sized files • Remote Visualization	• Collaboration services • Guaranteed bandwidth • PKI / Grid
BER: EMSL/Bio	-	10Gbps	50-100Gbps	• Bulk data • Real-time video • Remote control	• Collaborative services • Guaranteed bandwidth
BER: JGI/Genomics	-	1Gbps	2-5Gbps	• Bulk data	• Dedicated virtual circuits • Guaranteed bandwidth

Science Network Requirements Aggregation Summary

Science Drivers Science Areas / Facilities	End2End Reliability	Near Term End2End Band width	5 years End2End Band width	Traffic Characteristics	Network Services
BES: Chemistry and Combustion	-	5-10Gbps	30Gbps	• Bulk data • Real time data streaming	• Data movement middleware
BES: Light Sources	-	15Gbps	40-60Gbps	• Bulk data • Coupled simulation and experiment	• Collaboration services • Data transfer facilities • Grid / PKI • Guaranteed bandwidth
BES: Nanoscience Centers	-	3-5Gbps	30Gbps	•Bulk data •Real time data streaming •Remote control	• Collaboration services • Grid / PKI
FES: International Collaborations	-	100Mbps	1Gbps	• Bulk data	• Enhanced collaboration services • Grid / PKI • Monitoring / test tools
FES: Instruments and Facilities	-	3Gbps	20Gbps	• Bulk data • Coupled simulation and experiment • Remote control	• Enhanced collaboration service • Grid / PKI
FES: Simulation	-	10Gbps	88Gbps	• Bulk data • Coupled simulation and experiment • Remote control	• Easy movement of large checkpoint files • Guaranteed bandwidth • Reliable data transfer

Science Network Requirements Aggregation Summary

Science Drivers Science Areas / Facilities	End2End Reliability	Near Term End2End Band width	5 years End2End Band width	Traffic Characteristics	Network Services
Immediate Requirements and Drivers					
HEP: LHC	99.95+% (Less than 4 hours per year)	73Gbps	225-265Gbps	• Bulk data • Coupled analysis workflows	• Collaboration services • Grid / PKI • Guaranteed bandwidth • Monitoring / test tools
NP: CMS Heavy Ion	-	10Gbps (2009)	20Gbps	•Bulk data	• Collaboration services • Deadline scheduling • Grid / PKI
NP: JLAB	-	10Gbps	10Gbps	• Bulk data	• Collaboration services • Grid / PKI
NP: RHIC	Limited outage duration to avoid analysis pipeline stalls	6Gbps	20Gbps	• Bulk data	• Collaboration services • Grid / PKI • Guaranteed bandwidth • Monitoring / test tools

Aggregate Capacity Requirements Tell You How to Budget for a Network But Do Not Tell You How to Build a Network

- To actually build a network you have to look at where the traffic originates and ends up and how much traffic is expected on specific paths
- So far we have specific information for
  - LHC
  - SC Supercomputers
  - RHIC/BNL

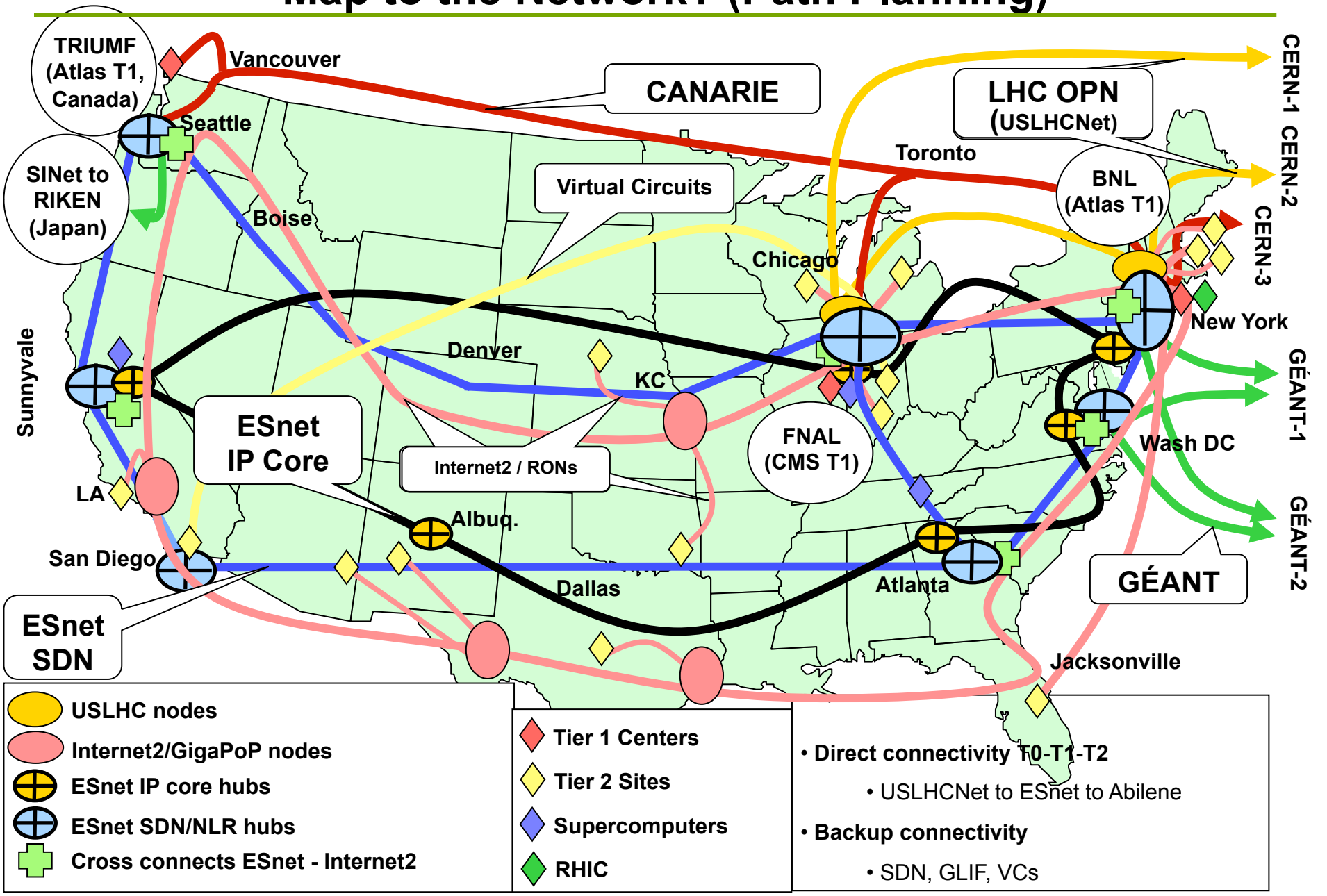
LHC ATLAS Bandwidth Matrix as of April 2008

Site A	Site Z	ESnet A	ESnet Z	A-Z 2008 Rate	A-Z 2010 Rate	
CERN	BNL	AofA (NYC)	BNL	10Gbps	20-40Gbps	
BNL	U. of Michigan (Calibration)	BNL (LIMAN)	Starlight (CHIMAN)	3Gbps	10Gbps	
BNL	Northeastern Tier2 Center	BNL (LIMAN)	Internet2 / NLR Peerings	3Gbps	10Gbps	
BNL	Great Lakes Tier2 Center	BNL (LIMAN)	Internet2 / NLR Peerings	3Gbps	10Gbps	new
BNL	Midwestern Tier2 Center	BNL (LIMAN)	Internet2 / NLR Peerings	3Gbps	10Gbps	
BNL	Southwestern Tier2 Center	BNL (LIMAN)	Internet2 / NLR Peerings	3Gbps	10Gbps	
BNL	Western Tier2 Center	BNL (LIMAN)	SLAC (BAMAN)	3Gbps	10Gbps	new
BNL	Tier3 Aggregate	BNL (LIMAN)	Internet2 / NLR Peerings	5Gbps	20Gbps	
BNL	TRIUMF (Canadian ATLAS Tier1)	BNL (LIMAN)	Seattle	1Gbps	5Gbps	

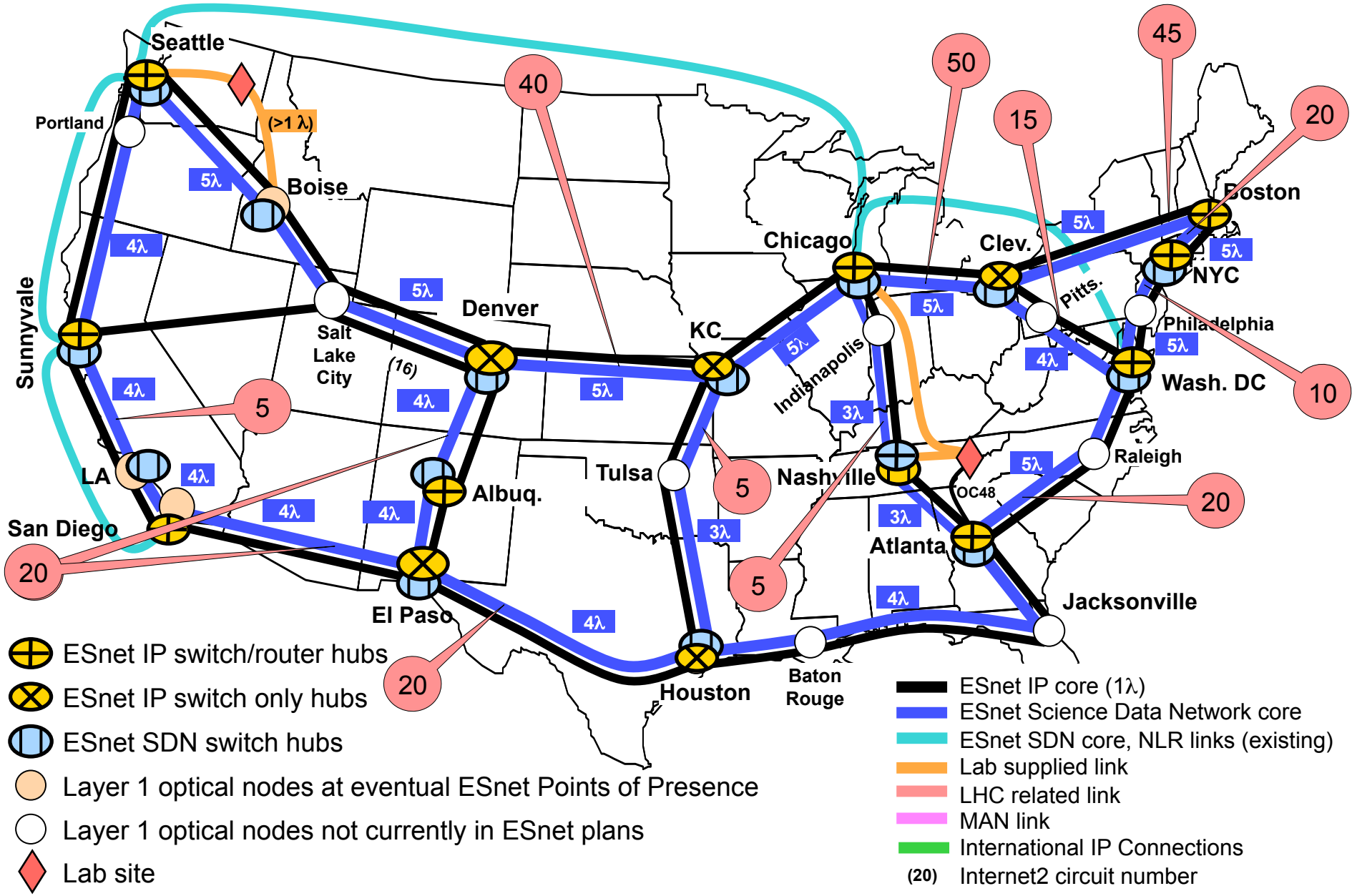
LHC CMS Bandwidth Matrix as of July 2008

Site A	Site Z	ESnet A	ESnet Z	A-Z 2008 Rate	A-Z 2010 Rate
CERN	FNAL	Starlight (CHIMAN)	FNAL (CHIMAN)	10Gbps	20-40Gbps
FNAL	U. of Michigan (Calibration)	FNAL (CHIMAN)	Starlight (CHIMAN)	3Gbps	10Gbps
FNAL	Caltech	FNAL (CHIMAN)	Starlight (CHIMAN)	3Gbps	10Gbps
FNAL	MIT	FNAL (CHIMAN)	AofA (NYC)/ Boston	3Gbps	10Gbps
FNAL	Purdue University	FNAL (CHIMAN)	Starlight (CHIMAN)	3Gbps	10Gbps
FNAL	U. of California at San Diego	FNAL (CHIMAN)	San Diego	3Gbps	10Gbps
FNAL	U. of Florida at Gainesville	FNAL (CHIMAN)	SOX	3Gbps	10Gbps
FNAL	U. of Nebraska at Lincoln	FNAL (CHIMAN)	Starlight (CHIMAN)	3Gbps	10Gbps
FNAL	U. of Wisconsin at Madison	FNAL (CHIMAN)	Starlight (CHIMAN)	3Gbps	10Gbps
FNAL	Tier3 Aggregate	FNAL (CHIMAN)	Internet2 / NLR Peerings	5Gbps	20Gbps

How do the Bandwidth – End Point Requirements Map to the Network? (Path Planning)



How do the Bandwidth – End Point Requirements Map to the Network? (Core Capacity Planning - 2010)



How do the Science Program Identified Requirements Compare to this Capacity Planning?

- The current network is built to accommodate the known, path-specific needs of the programs – However this is not the whole picture

Synopsis of Known Aggregate Requiremetns, 6/2008			
Science Areas / Facilities	**5 year end-to-end bandwidth requiremetns (Gb/s)**	**accounted for in current ESnet path planning**	**Unacc'ted for**
ASCR: ALCF	30	30	
ASCR: NERSC	40	40	
ASCR: NCLF	50	50	
BER: Climate	20		20
BER: EMSL/Bio	100		100
BER: JGI/Genomics	5		5
BES: Chemistry and Combustion	30		30
BES: Light Sources	60		60
BES: Nanoscience Centers	30		30
Fusion: International Collaborations	1		1
Fusion: Instruments and Facilities	20		20
Fusion: Simulation	88		88
HEP: LHC	265	265	
NP: CMS Heavy Ion	20		20
NP: JLAB	10		10
NP: RHIC	20	20	
total	**789**	**405**	**384**

Where Are We Now?

- The path-capacity map, however, so far only accounts for 405 Gb/s out of 789 Gb/s identified by the science programs

Synopsis of Known Aggregate Requiremetns, 6/2008			
	5 year end-to-end bandwidth requirements (Gb/s)	accounted for in current ESnet path planning	Unacc'ted for
total	789	405	384

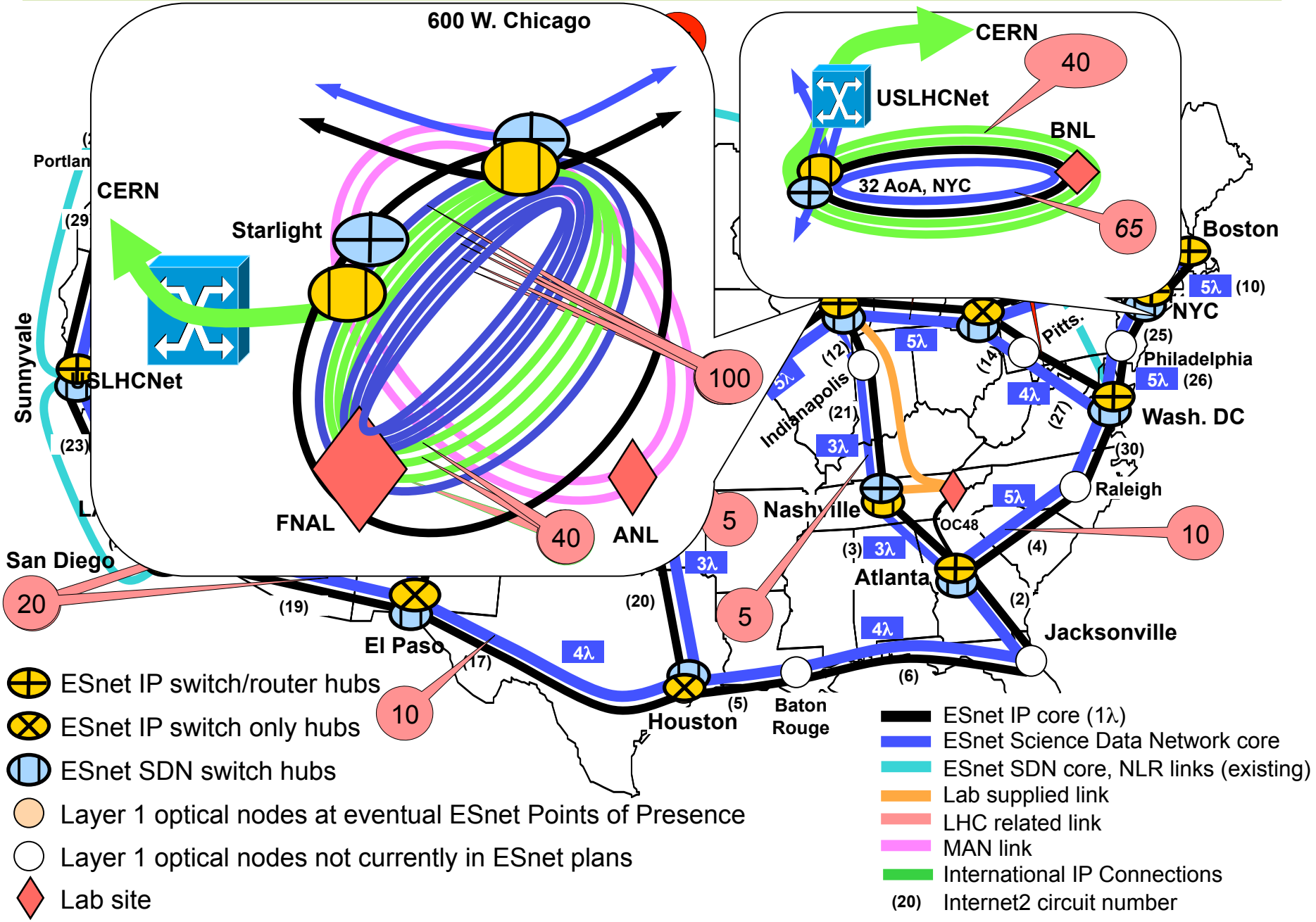
- The ESnet 5 yr budget provides for the capacity buildup of ESnet that is represented by (nominally) adding one wave per year.

ESnet Planned Aggregate Capacity (Gb/s)								
	2006	2007	2008	2009	2010	2011	2012	2013
ESnet core total inter-hub bandwidth (Gb/s)	57.50	240	530	620	740	910	920	980

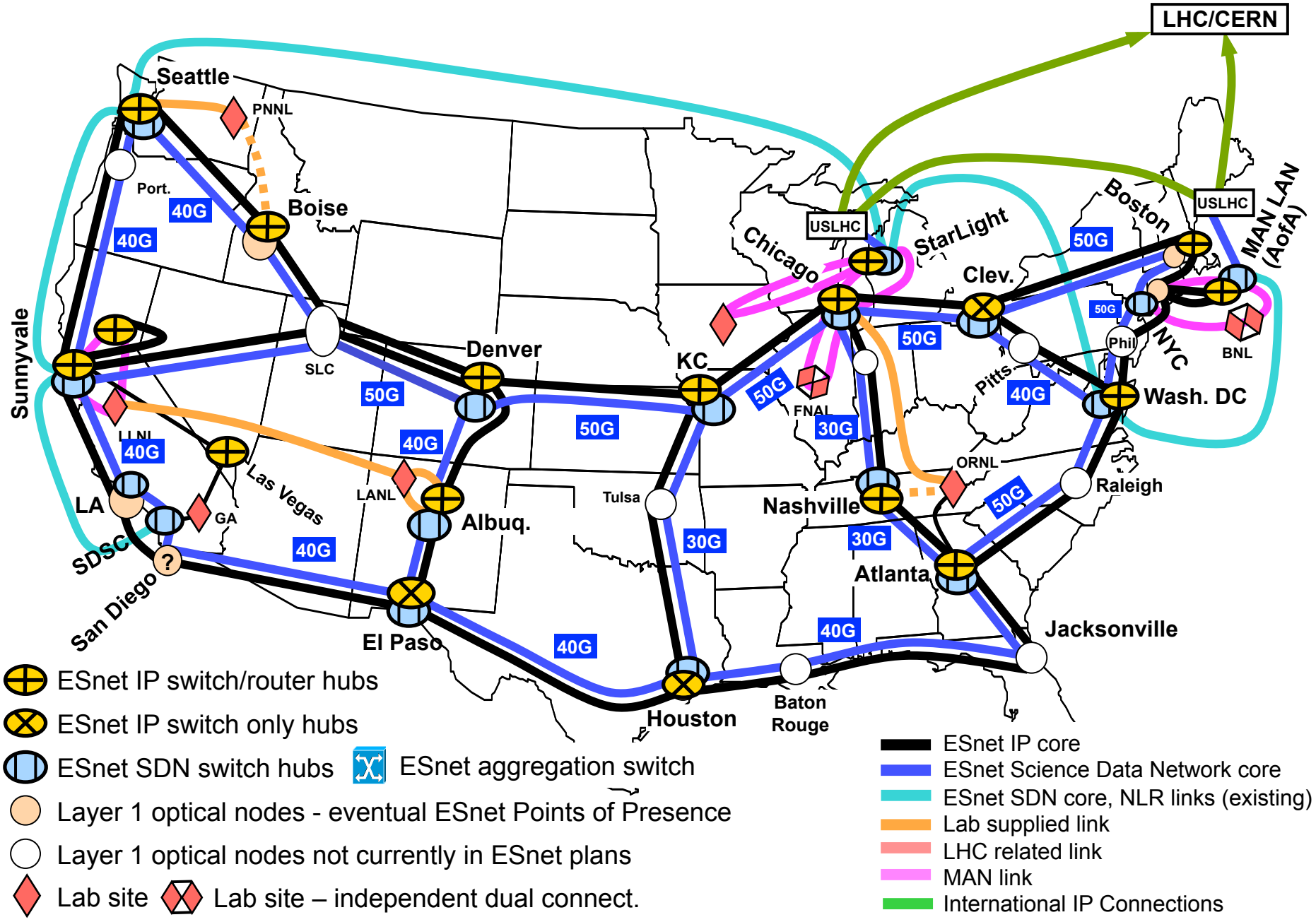
(This table is a summary of a small part of the ESnet Program reporting to OMB on plans and spending)

- The result is that the aggregate capacity growth of ESnet matches the know requirements – ***in aggregate***
- The "extra" capacity indicated above ***tries to account for the fact that there is not complete flexibility in mapping specific path requirements to the network*** infrastructure – and we have to plan the infrastructure years in advance based on incomplete science path-specific information
  - Whether this approach works is TBD, but indications are that it probably will

MAN Capacity Planning - 2010



This Sort of Analysis Leads to ESnet 4 As Planned for 2010



➢Is ESnet Planned Capacity Adequate for LHC?
(Maybe so, Maybe not)

- Several Tier2 centers (especially CMS) are capable of 10Gbps now
  - Many Tier2 sites are building their local infrastructure to handle 10Gbps
  - This means the 3Gbps estimates in the table are probably low
  - We won't know for sure what the "real" load will look like until the testing stops and the production analysis begins
  - ➢ ***Scientific productivity will follow high-bandwidth access to large data volumes ⇒incentive for others to upgrade***
- Many Tier3 sites are also building 10Gbps-capable analysis infrastructures
  - Most Tier3 sites do not yet have 10Gbps of network capacity
  - It is likely that this will cause a "second onslaught" in 2009 as the Tier3 sites all upgrade their network capacity to handle 10Gbps of LHC traffic
- ➢ ***It is possible that the USA installed base of LHC analysis hardware will consume significantly more network bandwidth than was originally estimated***
- N.B. Harvey Newman predicted this eventuality years ago

➢Observations on Reliability

- Reliability – the customers who talk about reliability are typically the ones building automated wide area workflow systems (LHC and RHIC).
  - "Transfer a data set" paradigm isn't as concerned with reliability, other than the annoyance/inconvenience of outages and their effect on a given transfer operation
  - ➢ ***However, prolonged outages can cause cascade failure in automated*** workflow systems (outage ⟹ analysis pipeline stall ⟹ data loss) since the instruments don't stop and the computing capacity is sized to analyze the data as it arrives
  - Many of our constituents are talking about moving to this model (e.g. Climate and Fusion) – this will increase demand for high reliability
  - ESnet's current strategy (ESnet4) has significantly improved reliability, and continues to do so – both theory and empirical data support this assertion

IIa. Re-evaluating the Strategy

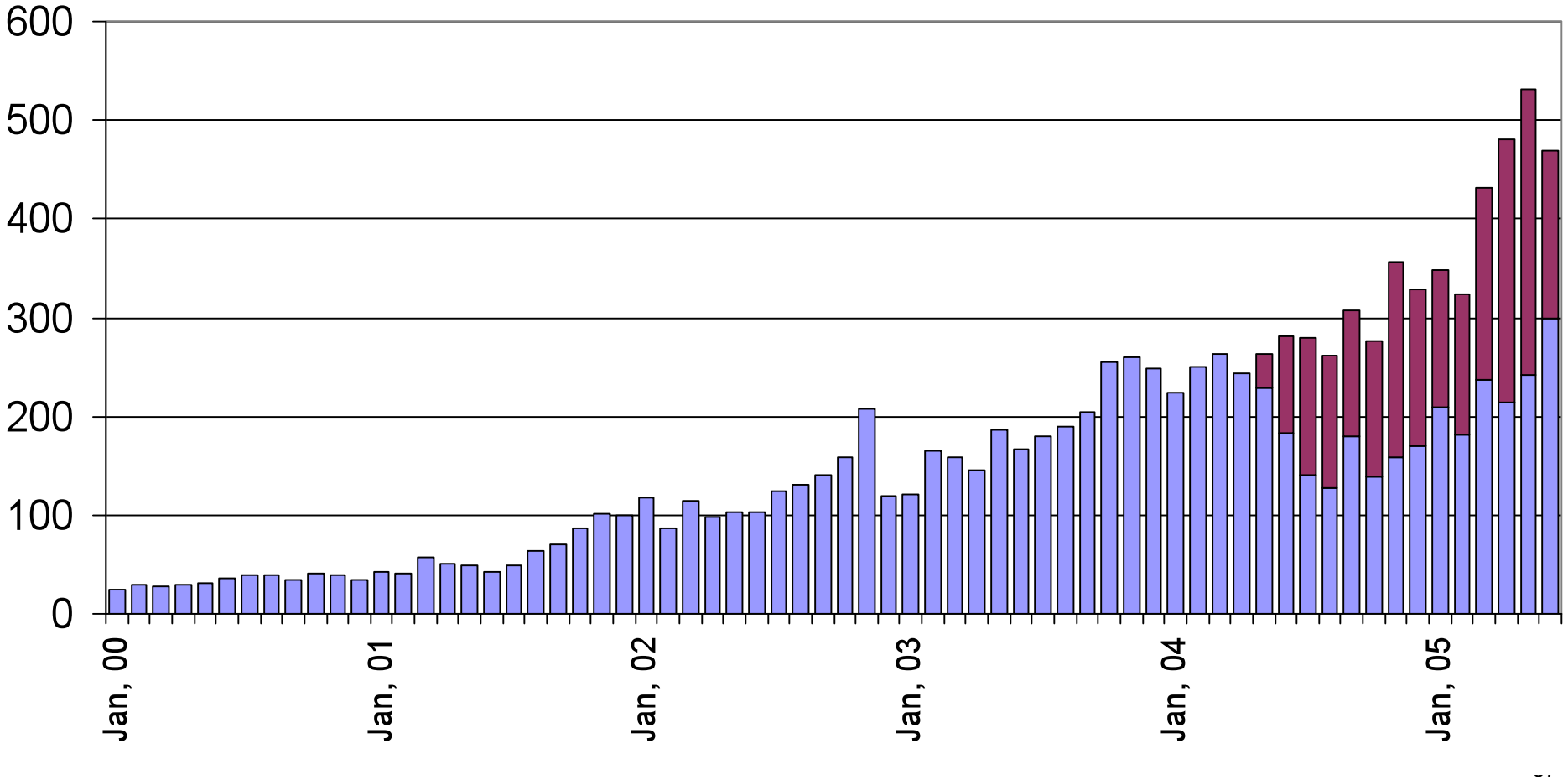
- The current strategy (that lead to the ESnet4, 2012 plans) was developed primarily as a result of the information gathered in the 2003 and 2003 network workshops, and their updates in 2005-6 (including LHC, climate, RHIC, SNS, Fusion, the supercomputers, and a few others) [workshops]
- So far the more formal requirements workshops have largely reaffirmed the ESnet4 strategy developed earlier
- ***However – is this the whole story?***

"Philosophical" Issues for the Future Network

- One can qualitatively divide the networking issues into what I will call "old era" and "new era"
- In the old era (to about mid-2005) data from scientific instruments did grow exponentially, but the actual used bandwidths involved did not really tax network technology
- In the old era there were few, if any, dominate traffic flows – all the traffic could be treated together as a "well behaved" aggregate.

Old Era Traffic Growth Charasteristics

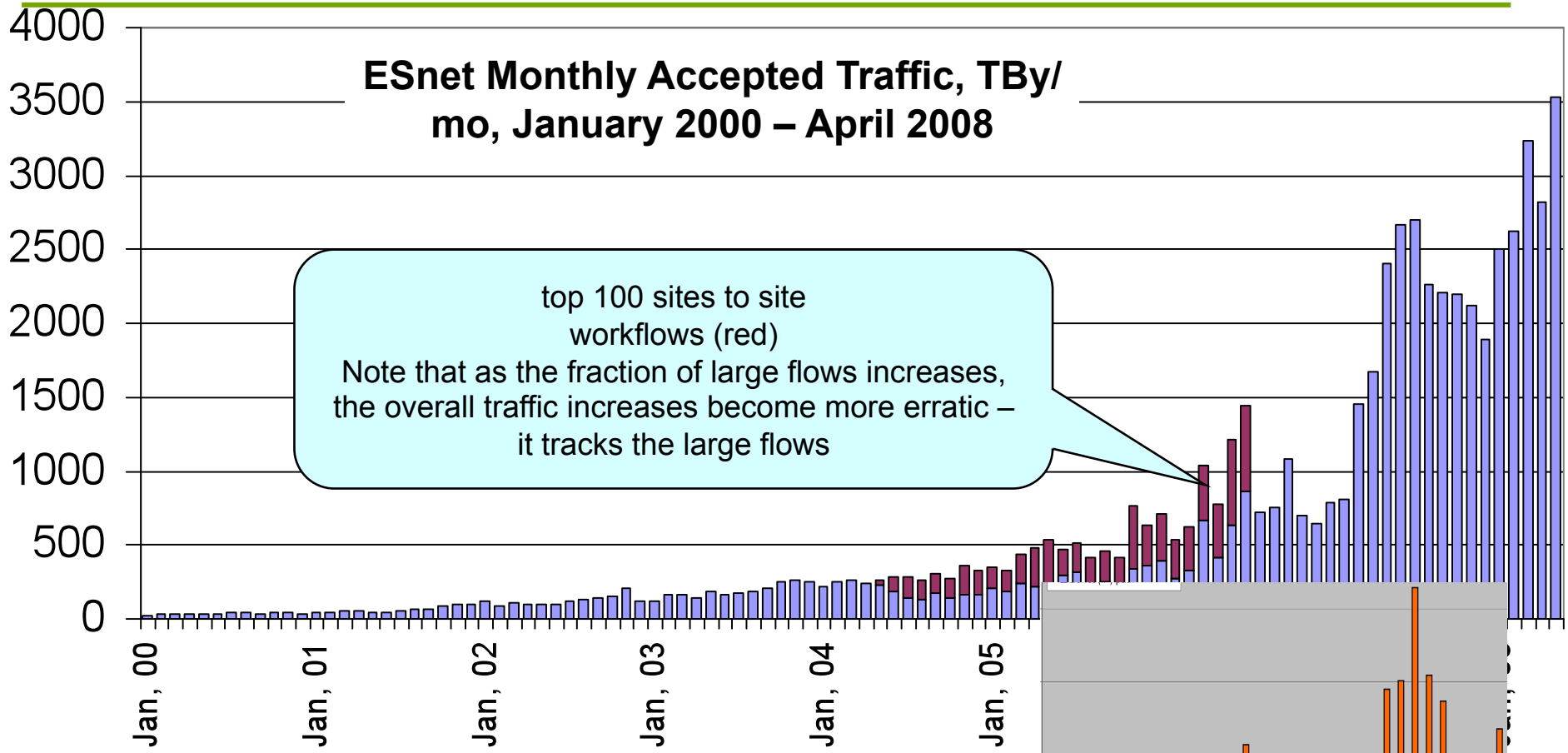
ESnet Monthly Accepted Traffic,
GBy/mo, January 2000 – June 2005



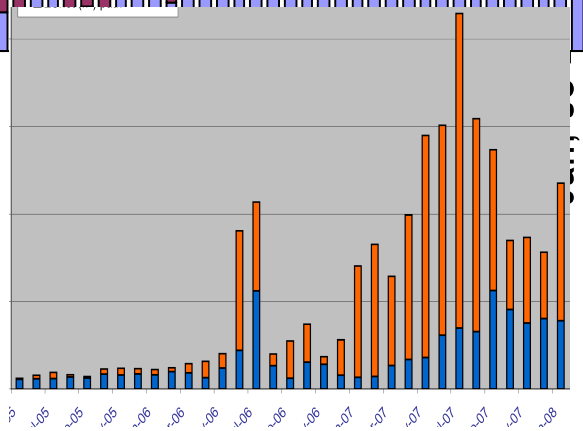
In New Era Large-Scale Science Traffic Dominates ESnet

- Large-scale science – LHC, RHIC, climate, etc. now generate a few thousand flows/month that account for about 90% of all ESnet traffic
- When a few large data sources/sinks dominate traffic then overall network usage follows the patterns of the very large users
- Managing this to provide good service to large users and not disrupt a lot of small users requires the ability to isolate these flows to a part of the network designed for them (“traffic engineering”)

Starting in mid-2005 a small number of large data flows dominate the network traffic



- ESnet is currently transporting more than 3 petabytes (3500 terabytes) per month
- Since about mid-2005 more than 50% of the traffic is now generated by the top 100 sites ⇒ large-scale science dominates all ESnet traffic



FNAL (LHC Tier 1 site) Outbound Traffic
(courtesy Phil DeMar, Fermilab)

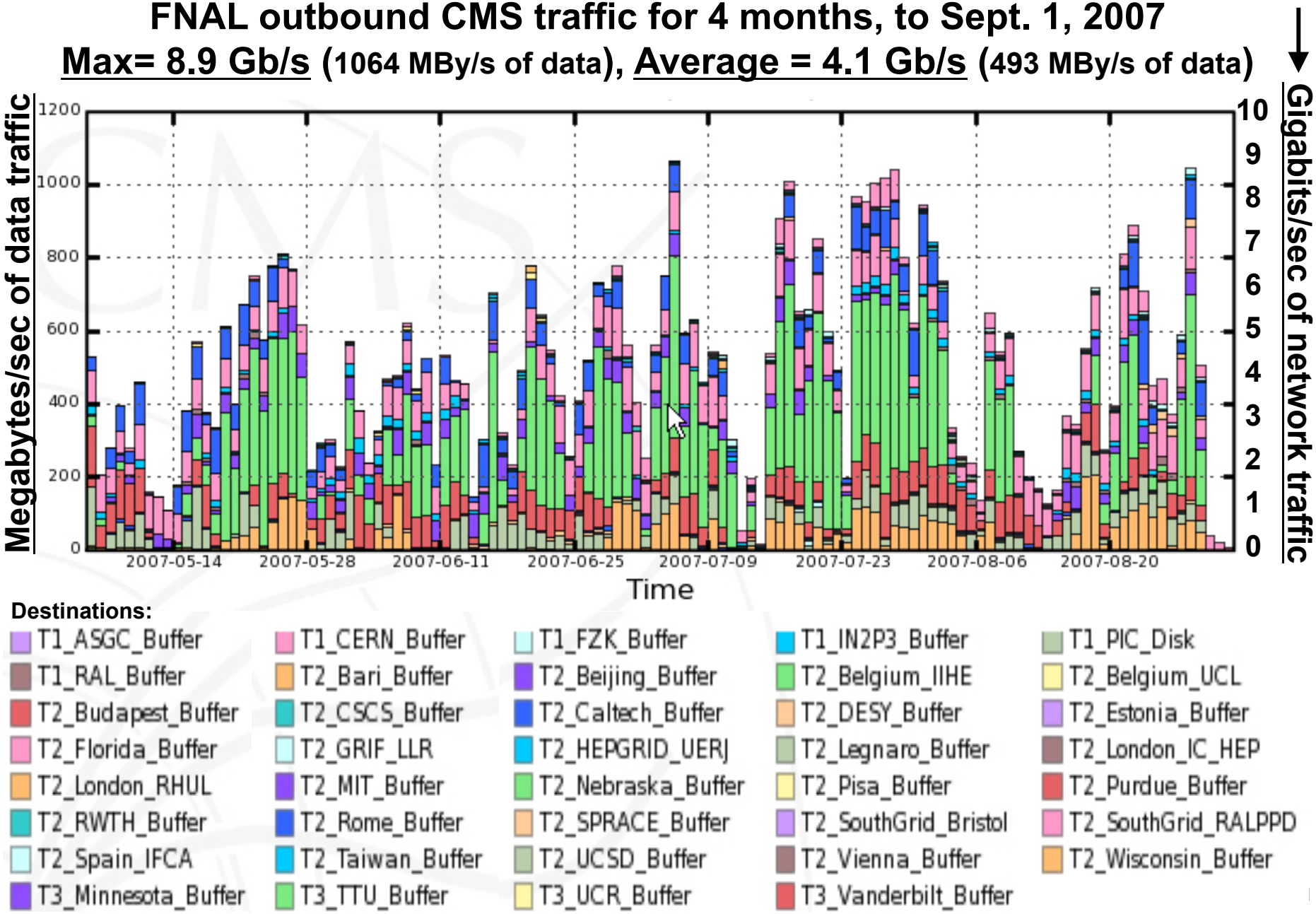
Issues for the Future Network– "New Era" Data

- Individual Labs now fill 10G links
  - Fermilab (an LHC Tier 1 Data Center) has 5 X 10Gb/s links to ESnet hubs in Chicago and can easily fill one or more of them for sustained periods of time
  - BNL has plans to host the entire LHC ATLAS dataset (up from 30%) and expects 20Gb/s sustained traffic

Individual Sites Can Now Routinely Fill 10G Circuits

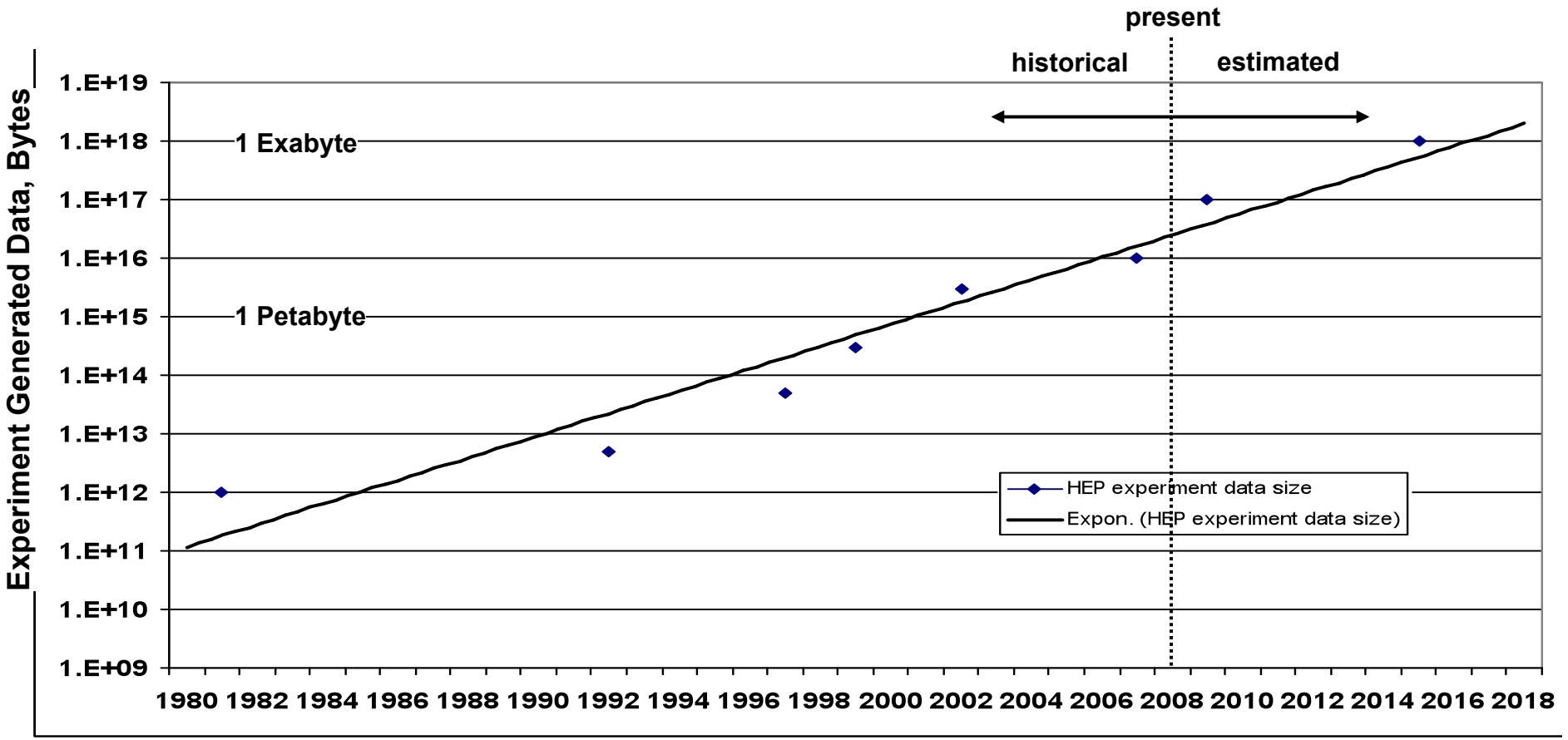
FNAL outbound CMS traffic for 4 months, to Sept. 1, 2007

Max= 8.9 Gb/s (1064 MBy/s of data), Average = 4.1 Gb/s (493 MBy/s of data)



The Exponential Growth of HEP Data is "Constant"

For a point of "ground truth" consider the historical growth of the size of HEP data sets – The trends as typified by the FNAL traffic will continue.

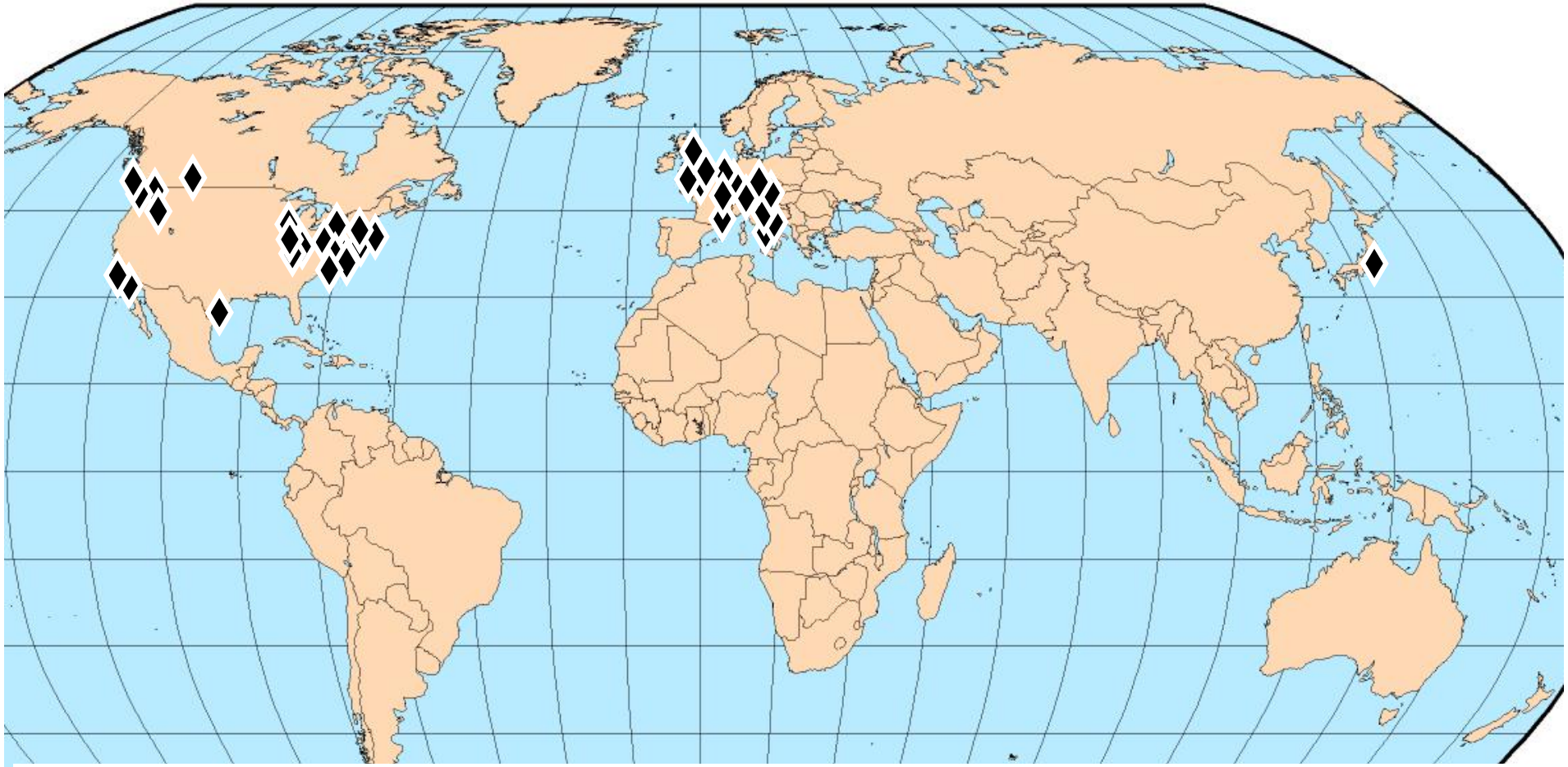


Data courtesy of Harvey Newman, Caltech,
and Richard Mount, SLAC

Issues for the Future Network

- Consider network traffic patterns – More “ground truth”
  - What do the trends in network patterns predict for future network needs?

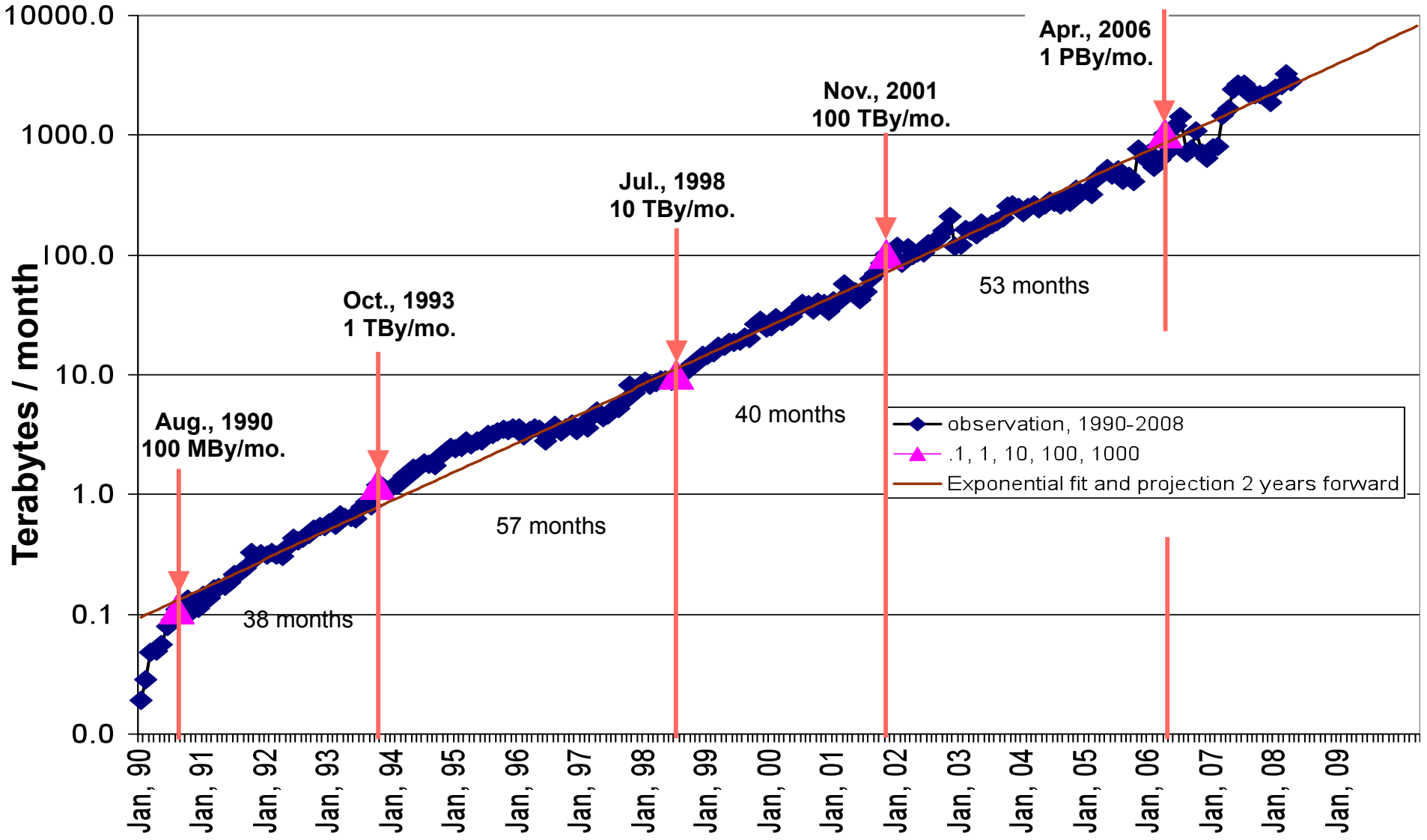
The International Collaborators of DOE's Office of Science, Drives ESnet Design for International Connectivity



Most of ESnet's traffic (>85%) goes to and comes from outside of ESnet. This reflects the highly collaborative nature of large-scale science (which is one of the main focuses of DOE's Office of Science).

⧫ = the R&E source or destination of ESnet's top 100 sites (all R&E)
(the DOE Lab destination or source of each flow is not shown)

ESnet Traffic has Increased by 10X Every 47 Months, on Average, Since 1990

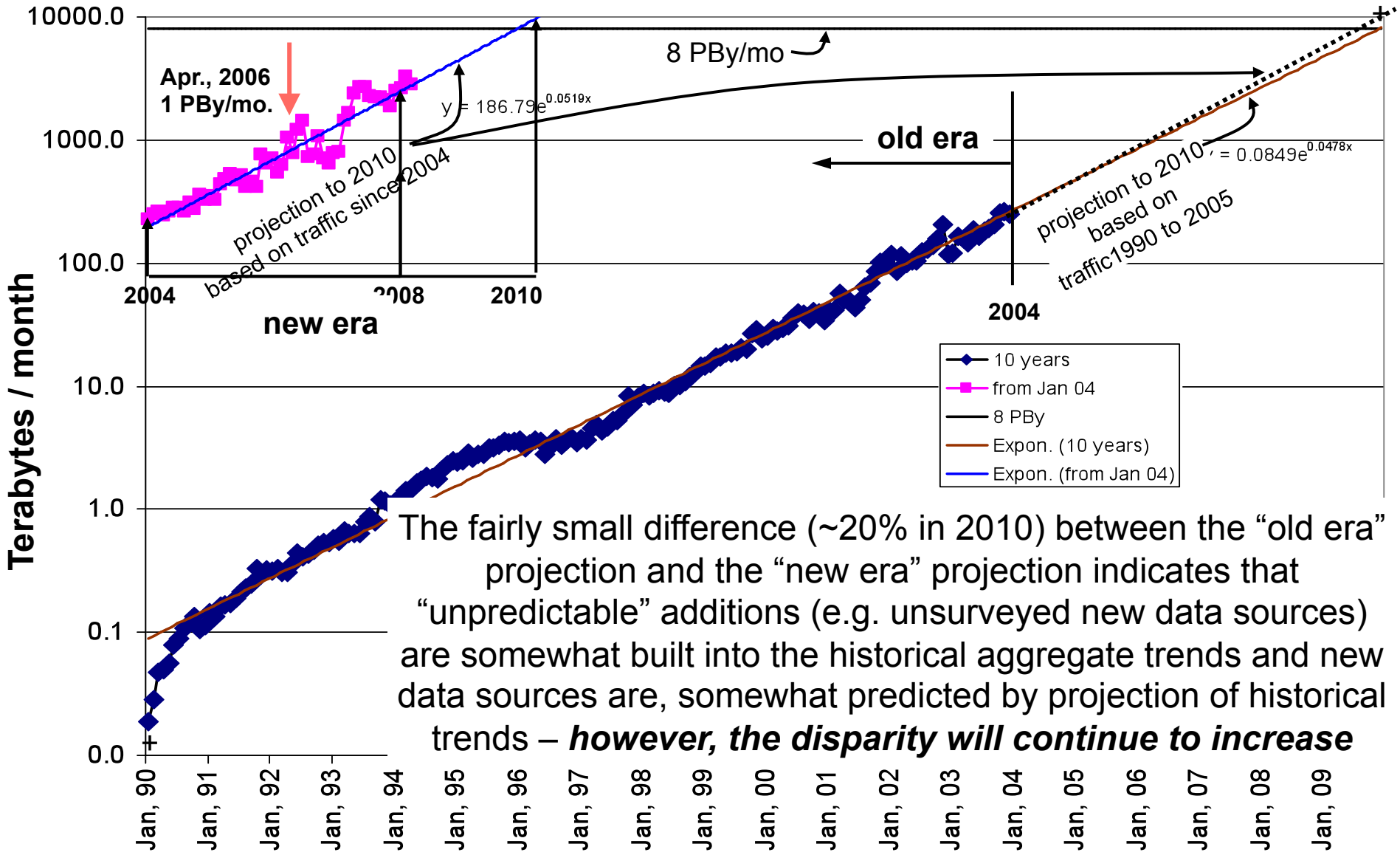


Log Plot of ESnet Monthly Accepted Traffic, January, 1990 – April 2008

Requirements from Network Utilization Observation

- Every 4 years, we can expect a 10x increase in traffic over current levels just based on historical trends
  - Nominal average load on busiest backbone paths in June 2006 was ~1.5 Gb/s
  - In 2010 average load will be ~15 Gbps based on current trends and 150 Gb/s in 2014
- Measurements of this type are science-agnostic
  - It doesn't matter who the users are, the traffic load is increasing exponentially

Projected Aggregate Network Utilization: New Era vs. Old



The fairly small difference (~20% in 2010) between the "old era" projection and the "new era" projection indicates that "unpredictable" additions (e.g. unsurveyed new data sources) are somewhat built into the historical aggregate trends and new data sources are, somewhat predicted by projection of historical trends – ***however, the disparity will continue to increase***

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

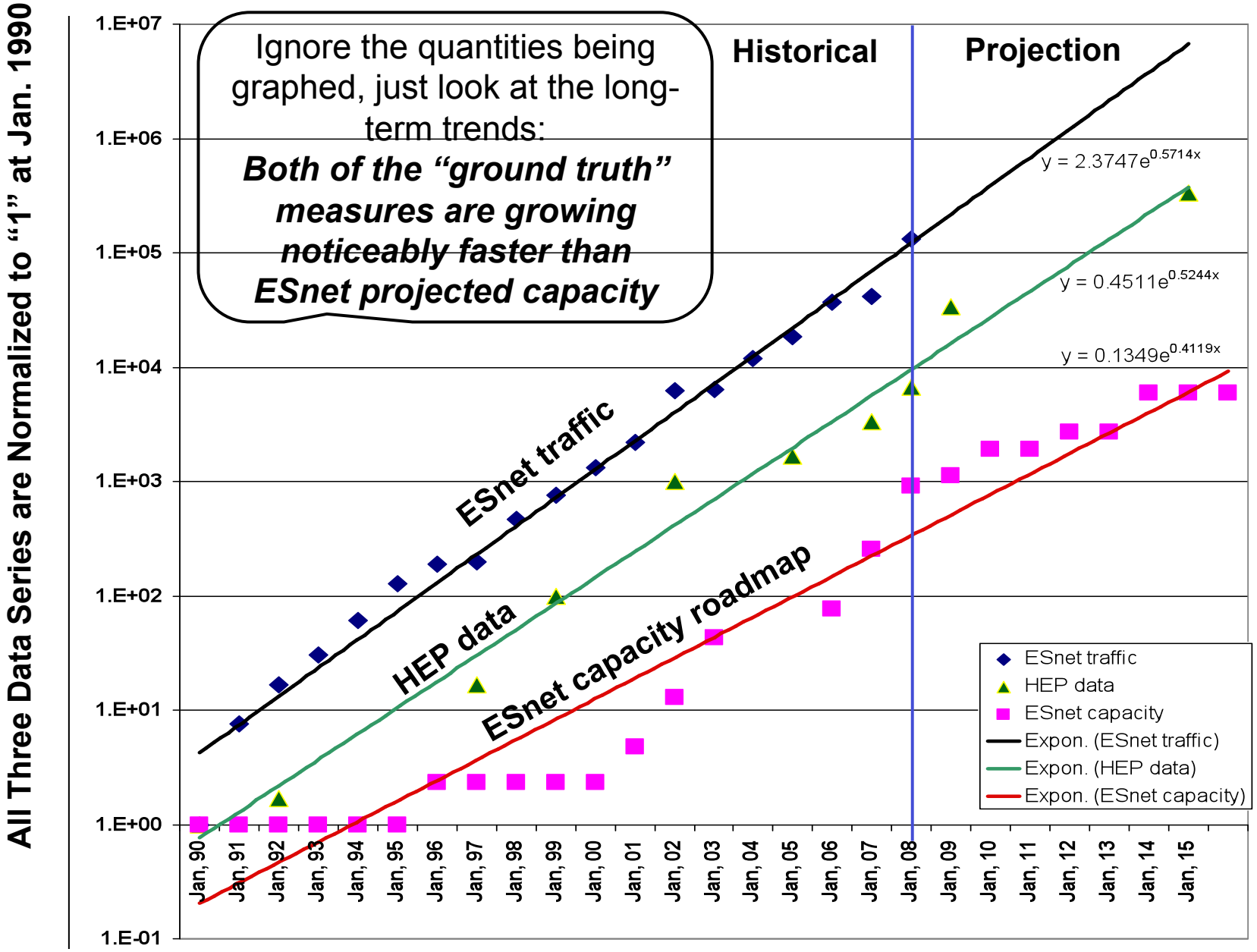
Where Will the Capacity Increases Come From?

- ESnet4 planning assumes a 5000 Gb/s core network by 2012
  - By 2012, technology trends will provide 100Gb/s optical channels in one of two ways:
    - By aggregation of lesser bandwidth waves in DWDM systems with high wave counts (e.g. several hundred)
    - By more sophisticated modulation of signals on existing waves to give 100 Gb/s per wave
- The ESnet4 SDN switching/routing platform is designed to support 100Gb/s network interfaces
- So the ESnet 2010 channel count will give some fraction of 5000 Gb/s of core network capacity by 2012 (20%? – complete conversion to 100G waves will take several years – depending on the cost of the equipment

➢ **Is this adequate to meet future needs?**

Not Necessarily!

Network Traffic, Physics Data, and Network Capacity



Re: Both of the "ground truth" measures are growing noticeably faster than ESnet projected capacity

- The lines are different units - one is rate, one is traffic volume, and one is capacity. These are all normalized to "1" at January 1990
- The only thing of interest here is the rate of growth. Since these are log plots, the significantly higher exponential growth of traffic (total accepted bytes) vs. total capacity (aggregate core bandwidth) means traffic will eventually overwhelm the capacity – "when" cannot be directly deduced from aggregate observations, but if you add this fact
    - Nominal average load on busiest backbone paths in June 2006 was ~1.5 Gb/s - In 2010 average load will be ~15 Gbps based on current trends and 150 Gb/s in 2014

Issues for the Future Network– "New Era" Data

- Just looking at the trends in current traffic growth, HEP data growth, and current (through 2010) ESnet capacity projection ….
  - The "new era" of science data (2010 and beyond) ***is*** likely to tax network technology
  - The "casual" increases in overall network capacity straightforward commercial channel capacity are less likely to easily meet future needs

Aside on Requirements Analysis and Network Planning -1

- It seems clear that the ESnet historical trends have built into them some the “unpredictables,” that is, projections from historical traffic data appear to represent some of the required total capacity, without reference to data projections from experiment and instrument usage analysis
- Does this apparent ability of the projected traffic trends to predict future network capacity requirements mean that we can plan based on aggregate traffic growth projections and dispense with detailed requirements gathering?

Aside on Requirements Analysis and Network Planning -2

- Of course not:
  1. The traffic trends provide a very high-level view of the required capacity. ***Knowing the required aggregate capacity requirement does not tell you how the network must be build*** in order to be useful. Detailed requirements analysis, such as shown for the LHC, above, tells how the network must be built.
  2. Strong coupling of the network requirements planning to the Science Program Offices and science community is absolutely essential for generating the shared sense of urgency that results in the funding required to build the network with the required capacity

Where Do We Go From Here?

- The current estimates from the LHC experiments and the supercomputer centers ***have the currently planned ESnet 2011 configuration operating at capacity*** and there are several other major instruments that will be generating significant data in that time frame
- The significantly higher exponential growth of traffic (total accepted bytes) vs. total capacity (aggregate core bandwidth) means traffic will eventually overwhelm the capacity – "when" cannot be directly deduced from aggregate observations, but if you add this fact
    - Nominal average load on busiest backbone paths in June 2006 was ~1.5 Gb/s - In 2010 average load will be ~15 Gbps based on current trends and 150 Gb/s in 2014

My (wej) guess is that problems will start to occur by 2015-16 unless new technology approaches are found

New Technology Issues

- It seems clear that we will have to have both more capacity and the ability to ***more flexibly map traffic to waves*** (traffic engineering) in order to make optimum use of the available capacity

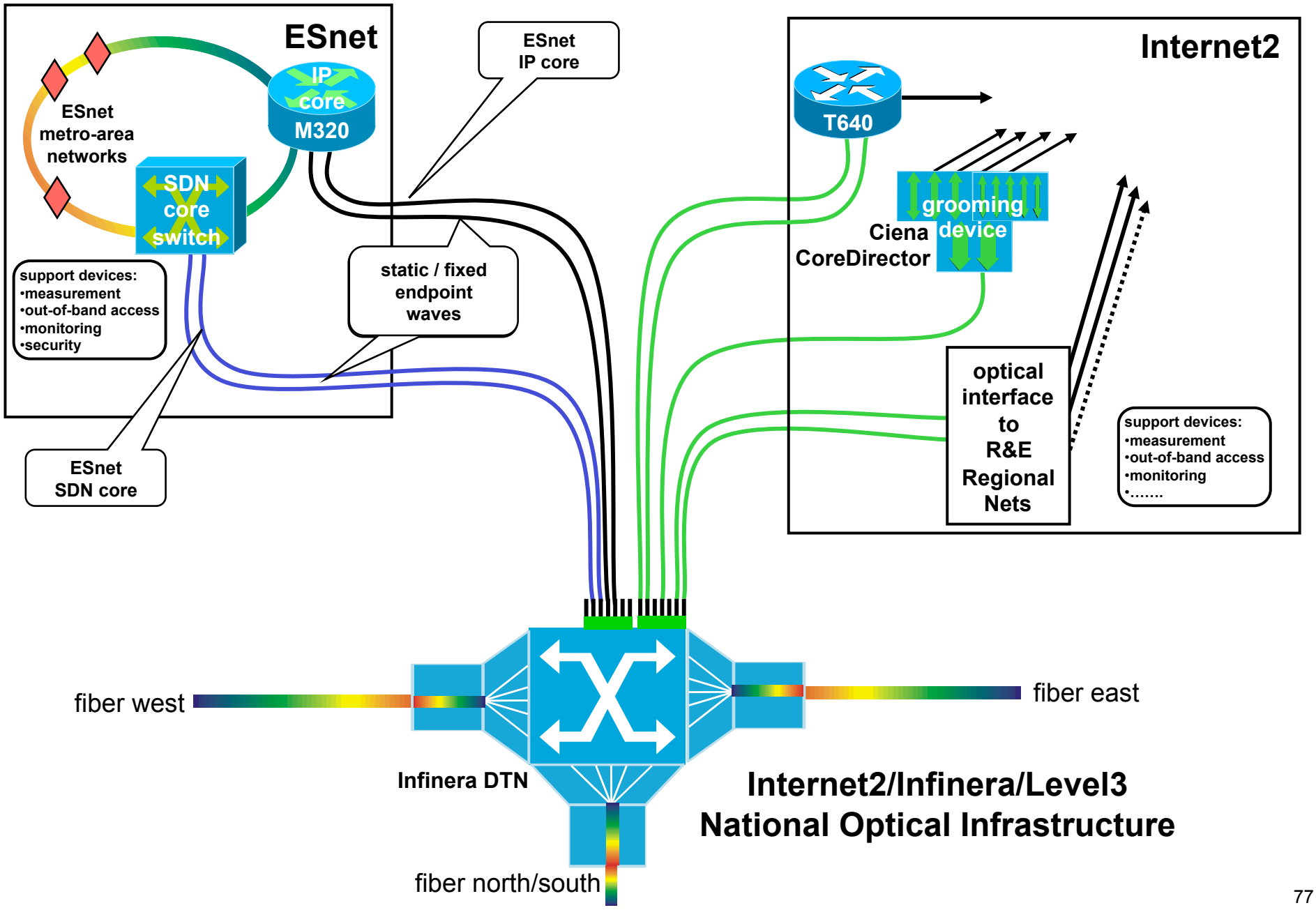
So, What Now?

- The Internet2-ESnet partnership optical network is build on dedicated fiber and optical equipment
  - The current configuration is configured with 10 ✕ 10G waves / fiber path and more waves will be added in groups of 10
- The ***current wave transport topology is essentially static*** or only manually configured - our current network infrastructure of routers and switches assumes this

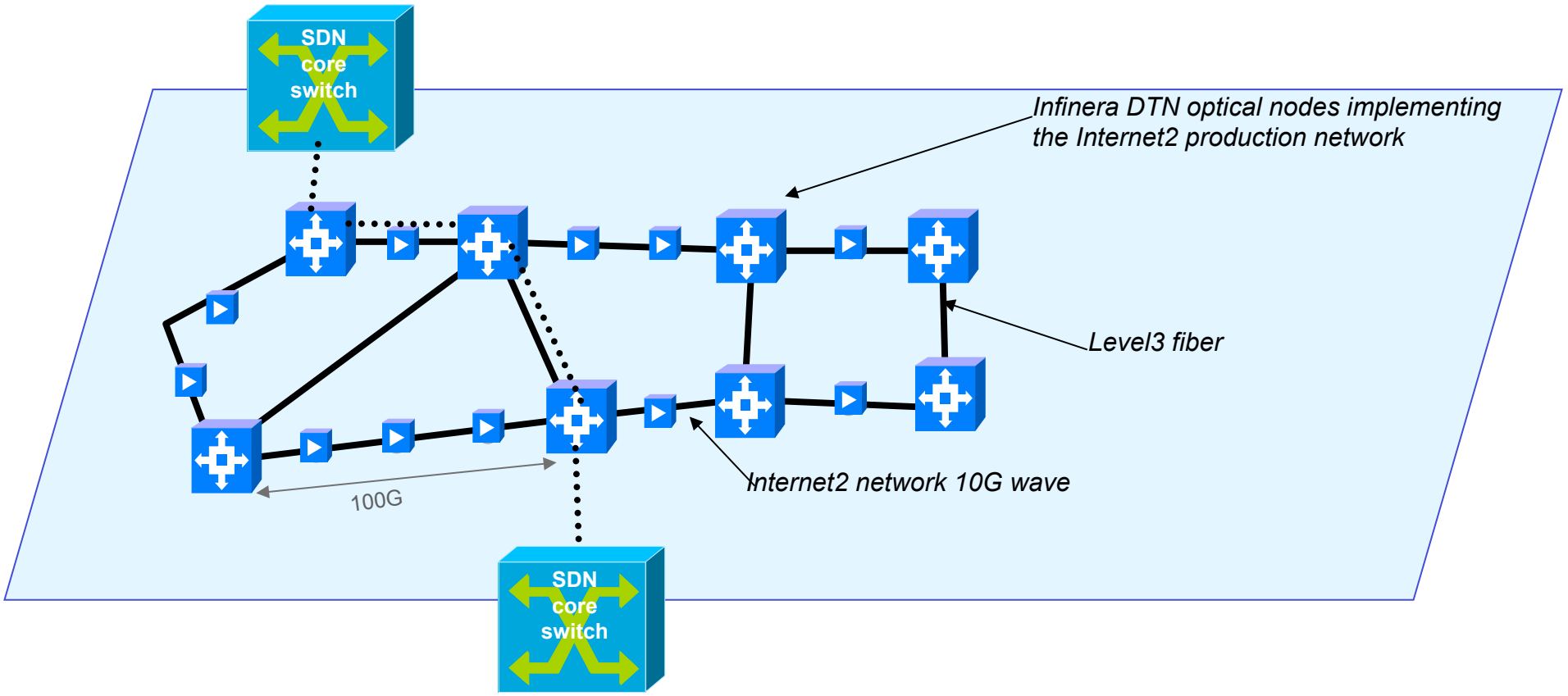
➢ We must change this situation and integrate the optical transport with the "network" and provide for dynamism / route flexibility at the optical level

- With completely flexible traffic management extending down to the optical transport level we ***should be able to extend the life of the current infrastructure*** by moving significant parts of the capacity to the specific routes where it is needed

Typical Internet2 and ESnet Optical Node Today



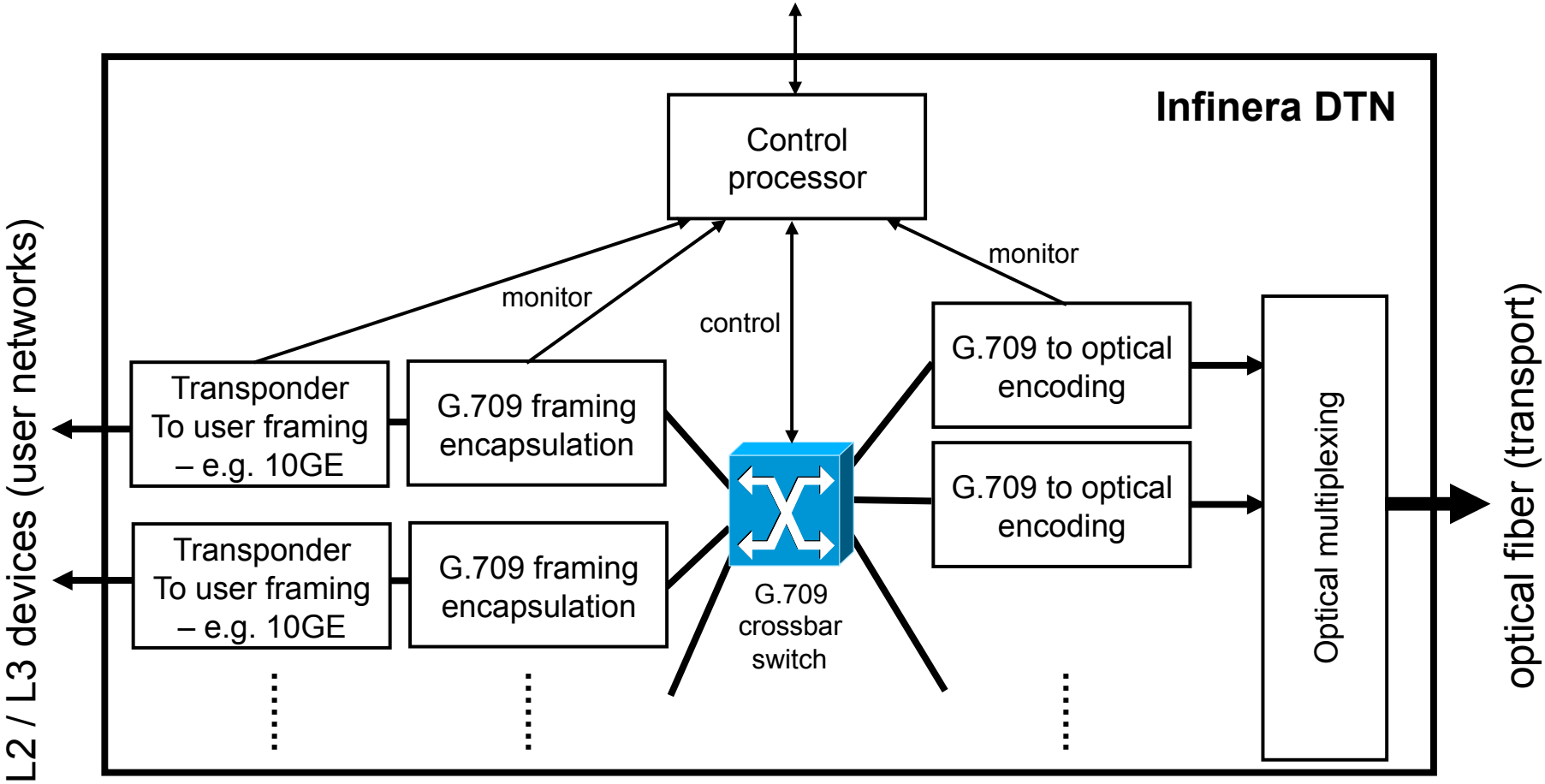
Today the Topology of the Optical Network as Seen by the Attached L2 and L3 Devices is Determined by a Static Wave Over Fiber Path Configuration



Dynamic Topology Management

- The Infinera optical devices (“DTN”) are capable of dynamic wave management
  - DTNs convert all user network traffic (Ethernet or SONET) to G.709 framing internally and the DTNs include a G.709 crossbar switch that can map any input (user network facing) interface to any underlying wave

Architecture of the Infinera DWDM System Used in the Internet2-ESnet network



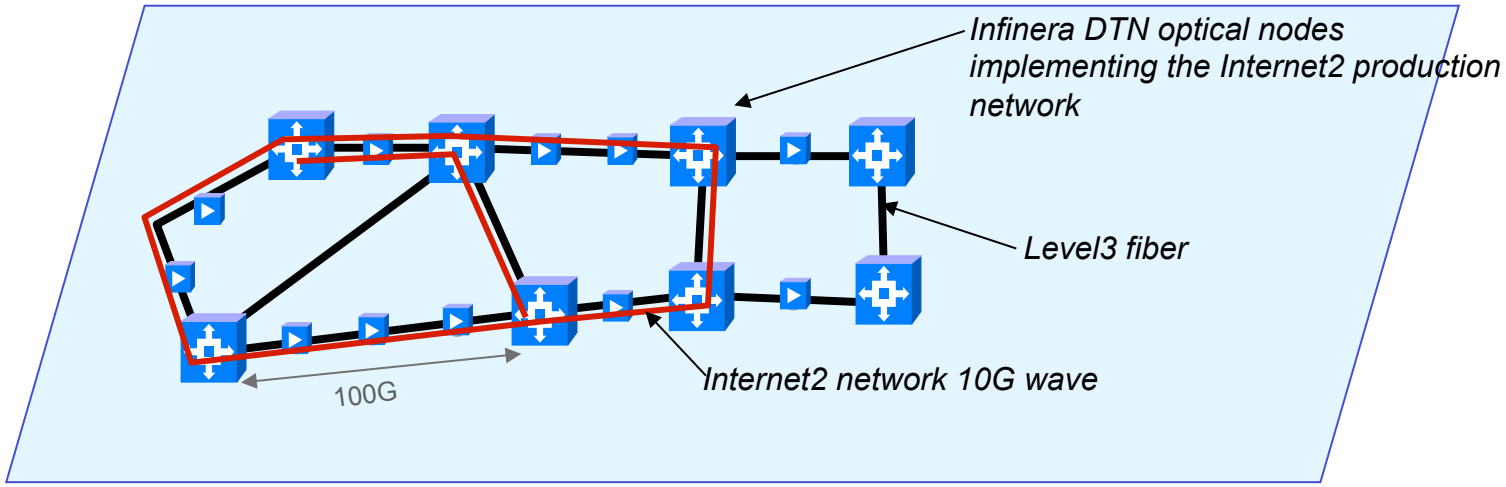
The crossbar switch determines what end points (e.g. layer 2 switch and layer 3 router interfaces) are connected together.
In other words, the topology of the optical network is entirely determined by the configurations of all of the crossbar switches in the optical network.

Dynamically Routed Optical Circuits for Traffic Engineering

- By adding a layer 1 (optical) control plane that is managed by Internet2 and ESnet, and that is integrated with the L2/3 control plane, the underlying topology of the optical network can be changed as needed for traffic engineering (management)
- An L1 control plane approach is in the planning phase and a testbed to do development and testing is needed
  - It is possible to build such a testbed as an isolated overlay on the production optical network and such a testbed has been proposed
- The control plane manager approach currently being considered is based on an extended version of the OSCARS [OSCARS] dynamic circuit manager – but a good deal of R&D is needed for the integrated L1/2/3 dynamic route management

R&D

Dynamically Routed Optical Circuits for Traffic Engineering

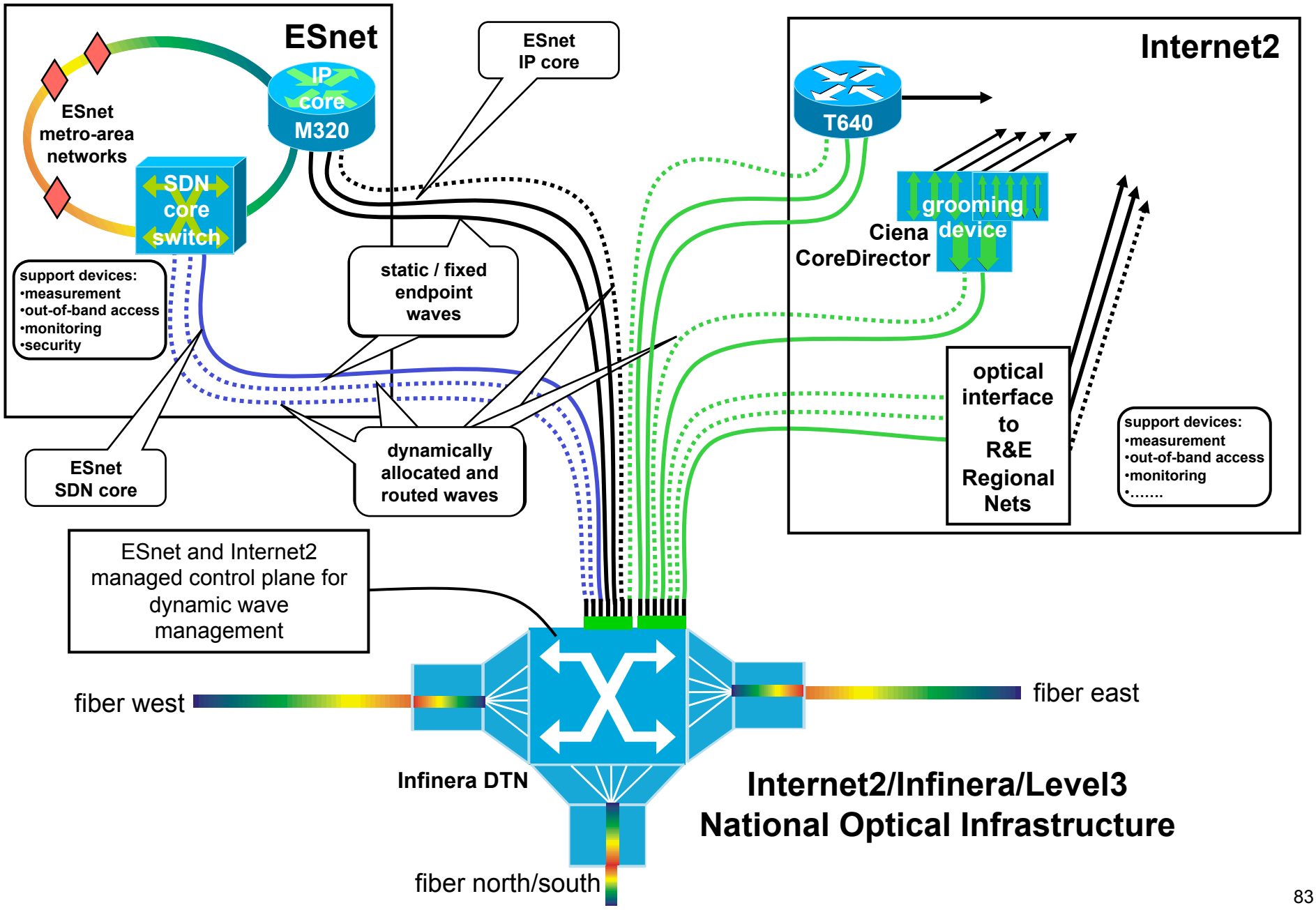


- The black paths can be though of both as the physical fiber paths and as the fixed wave allocations that provide a static configuration, especially for the layer 3 (IP) routers
- The red paths are dynamically switched layer 1 (optical circuit / wave) paths that can provide:

  1) Transient / reroutable paths between core network switch/router interfaces for more capacity, or

  2) Direct connections between sites if the intermediate networks can carry optical circuits

Internet2 and ESnet Optical Node in the Future



New Capability Requirements for Services

- The new service-oriented capabilities of the network, principally related to bandwidth reservation and end-to-end monitoring are also very important and have been discussed elsewhere
  - see, e.g.:
    - “Network Communication as a Service-Oriented Capability” and
    - “Intra and Interdomain Circuit Provisioning Using the OSCARS Reservation System.”
    - Both are available at http://www.es.net/pub/esnet-doc/index.html

IIIa. Federated Trust Services

- Remote, multi-institutional, identity authentication is critical for distributed, collaborative science in order to permit sharing widely distributed computing and data resources, and other Grid services
- Public Key Infrastructure (PKI) is used to formalize the existing web of trust within science collaborations and to extend that trust into cyber space
  - The function, form, and policy of the ESnet trust services are driven entirely by the requirements of the science community and by direct input from the science community
  - International scope trust agreements that encompass many organizations are crucial for large-scale collaborations
- The service (and community) has matured to the point where it is revisiting old practices and updating and formalizing them

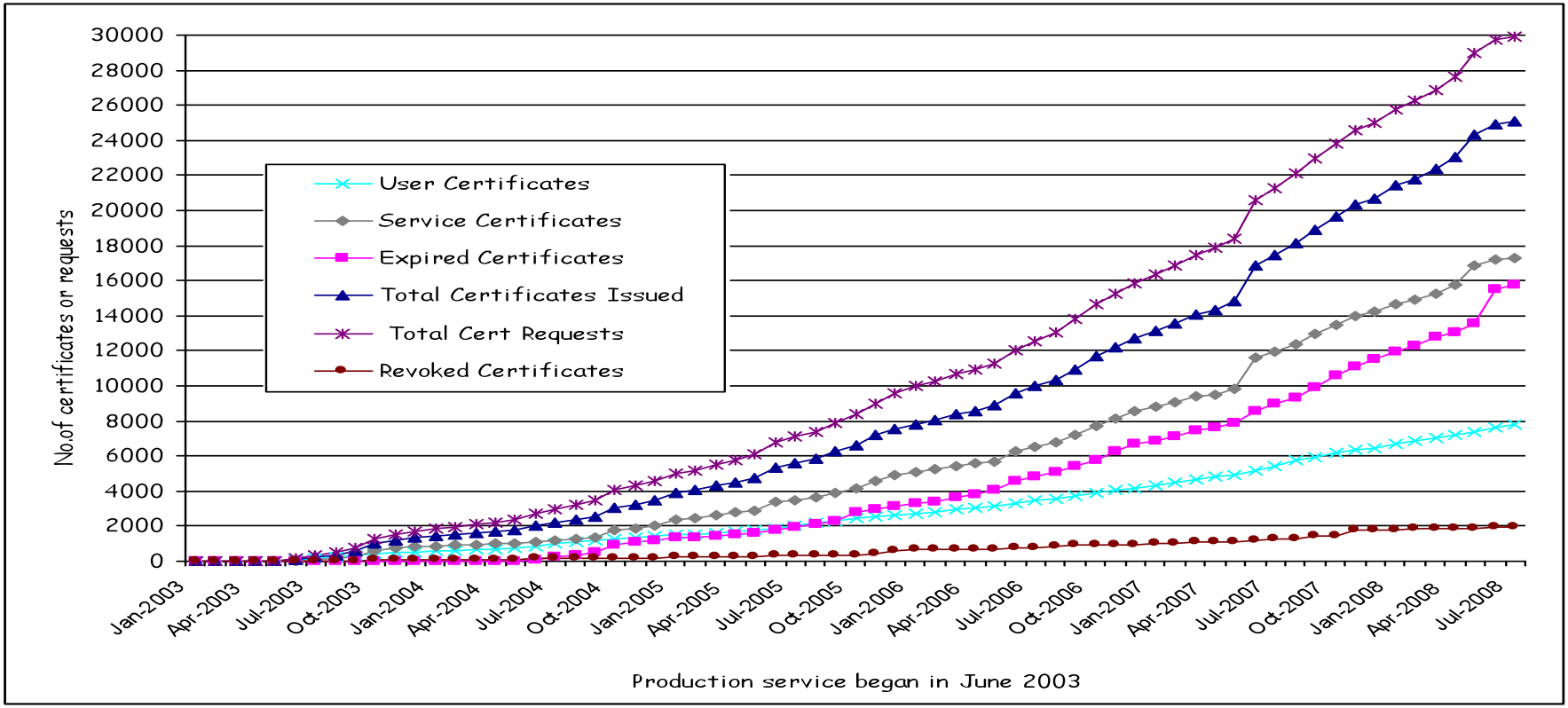
ESnet Grid CA and Federation Strategy

- ESnet operates the DOEGrids CA , which provides X.509 certificates to DOE SC funded (and related collaborators) projects; and actively supports IGTF, world-wide Grid CA federation
- Future technology strategies that are on the current ESnet roadmap
  - Make DOEGrids CA more robust
    - Multi-site cloned CA server and HSM (key management hardware)
      - Secure, disaster-resilient non-stop CA service (funded; in development)
  - Extend ESnet CA infrastructure to support Shibboleth-> X.509 certificates (Federated Identity CA)
    - Existing standards and ESnet hardware/software platform will provide X.509 certificates for DOE SC and related project members (limited funding)
    - Partnering with LBNL, NERSC, ANL to develop interoperability and policy

ESnet Grid CA and Federation Strategy

- Federation
  - ESnet has joined InCommon (as a service provider or SP)
    - InCommon is the US-wide academic Shibboleth federation
    - Enables ESnet to provide services, like CA gateways, using InCommon trust relationship, to sites recognizing InCommon
    - Studying integration with UCTrust and other regional federations
    - Usefulness TBD (see below)
  - Improving standards in IGTF
    - Grid Trust federation for CAs will recognize CAs providing gateways to Shibboleth federations
  - OpenID and Shibboleth service development
    - OpenID is a simple, web-based digital identity protocol from industry
    - OpenID consumer (clients) and Openid Provider (OP) for DOEGrids under study
    - “Retrofit” of Shibboleth and OpenID into existing ESnet services (non-CA)

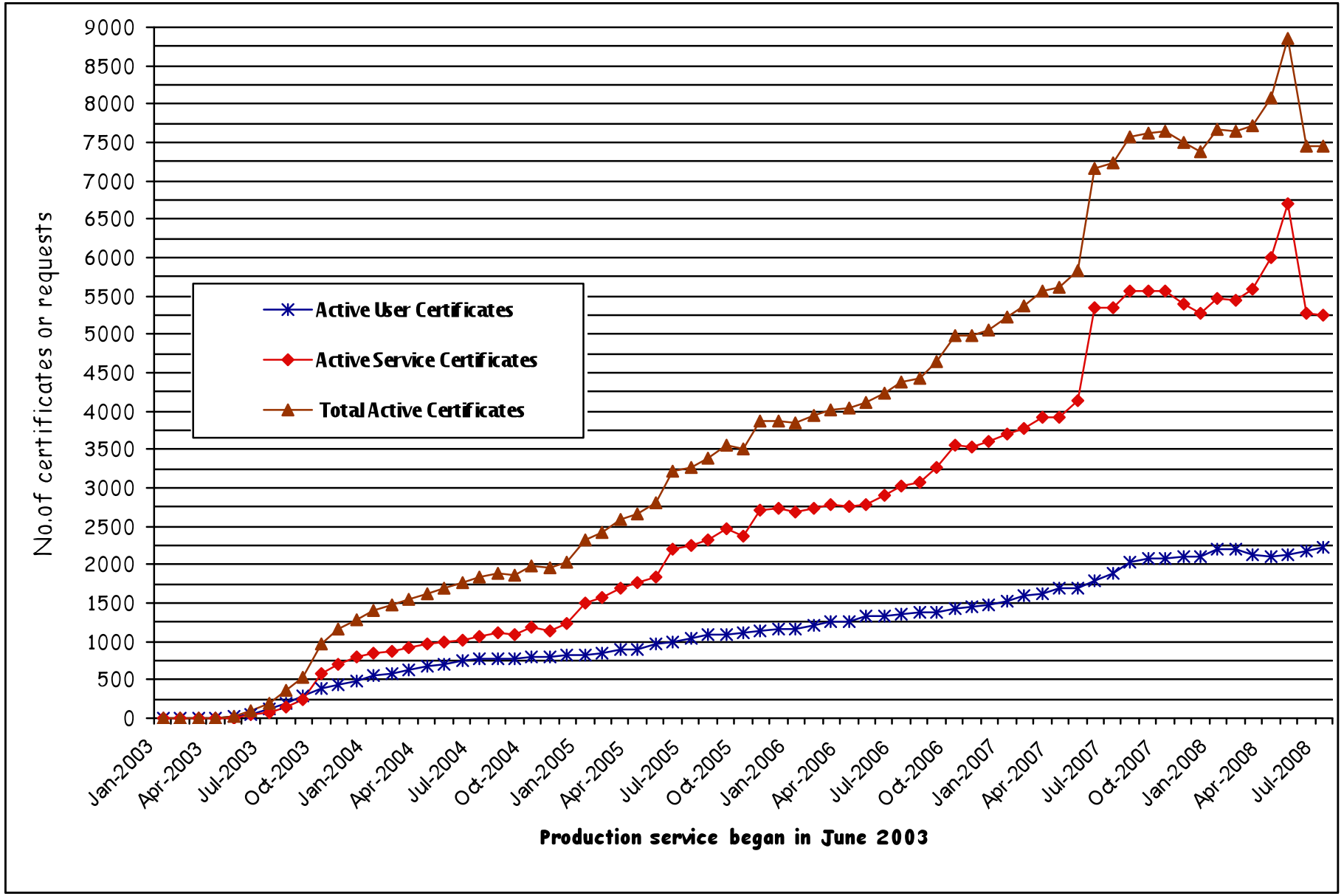
DOEGrids CA (one of several CAs) Usage Statistics July 15, 2008



User Certificates	7784	Total No. of Revoked Certificates	1936
Host & Service Certificates	17306	Total No. of Expired Certificates	15726
Total No. of Requests	29953	Total No. of Certificates Issued	25116
		Total No. of Active Certificates	7454
ESnet SSL Server CA Certificates			49
FusionGRID CA certificates			115

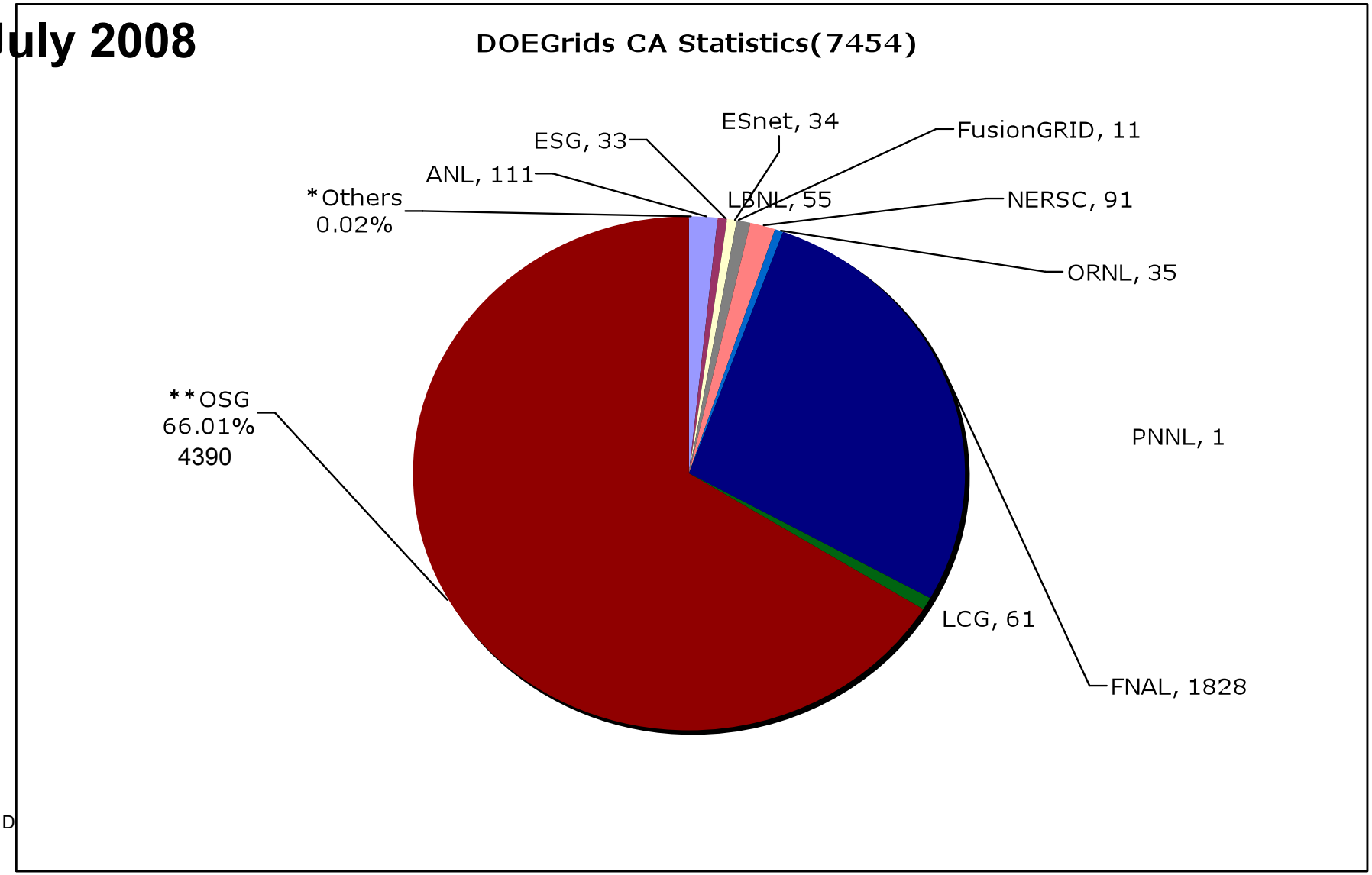
* Report as of July 15, 2008

DOEGrids CA (Active Certificates) Usage Statistics, July 15, 2008

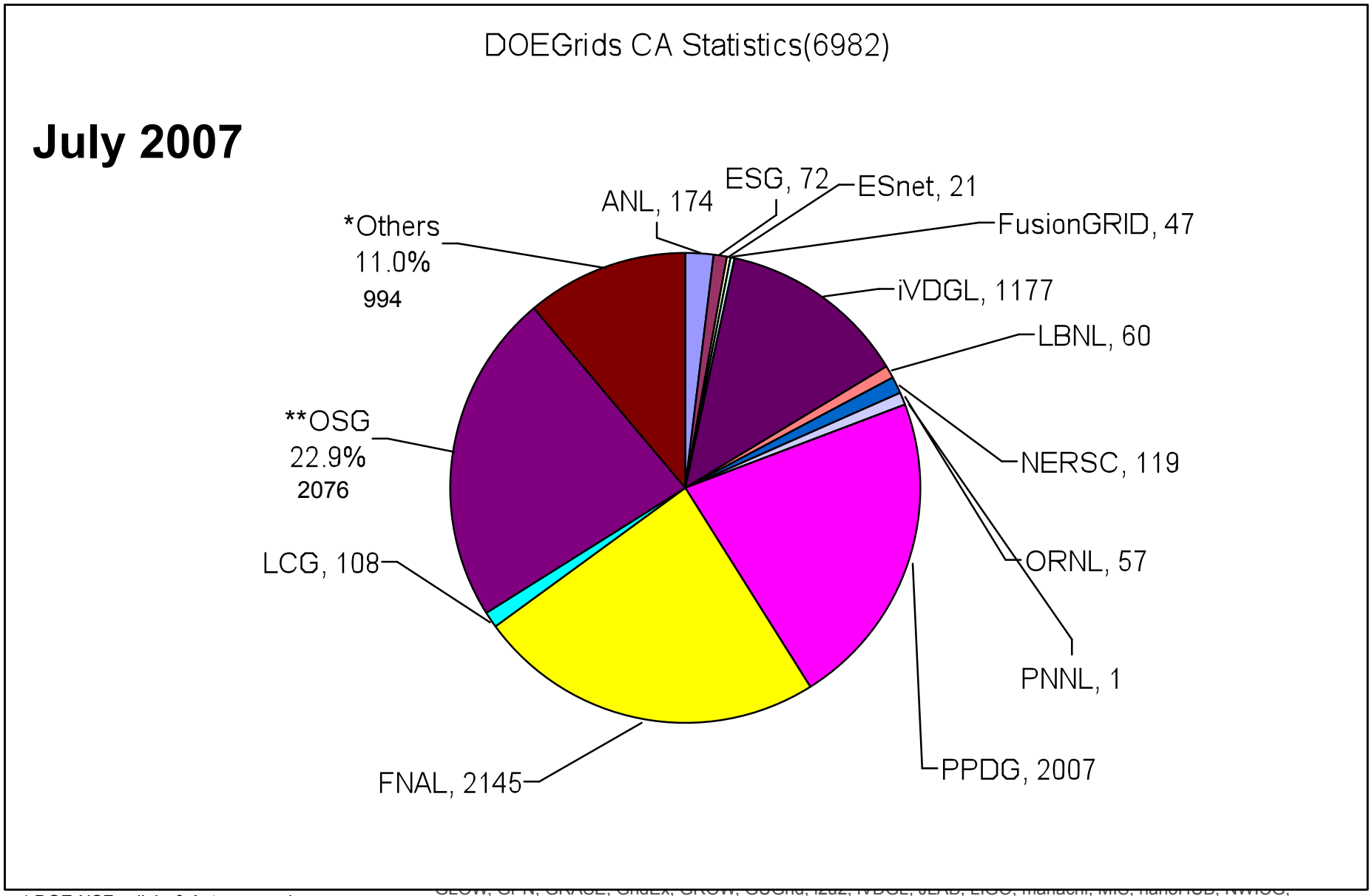


* Report as of July 15, 2008

DOEGrids CA Usage - Virtual Organization Breakdown



DOEGrids CA Usage - Virtual Organization Breakdown



* DOE-NSF collab. & Auto renewals

OSG, OSGEDU, SBGrid,SDSS, SLAC, STAR & USATLAS)

DOEGrids CA Audit

- Audit was conducted as part of the strategy to strengthen ties between DOEGrids CA and European counterparts
- The audit report and response will be available shortly
  - Will be released through DOEGrids PMA (www.doegrids.org)
- Major issues:
  - US Science ID verification – EU Grid organizations have agreed to accept "peer sponsored" or NFS-allocation as alternative to face-to-face + ID check
  - Renewals – US Science is resistant to re-verification of IDs. We will address this in part by improving information flow in DOEGrids so RAs have better oversight, but this is only a step in the right direction. NB: Adopting a "federation" approach in the early stages will gradually diminish this issue.
  - RFC 3647 – we will rewrite (translate) the DOEGrids CPS in RFC 3647 format (our auditors insist, IGTF does not require)
  - Audit burden – we are organizing our documentation and security management along the lines of NIST SP 800-53. This will align us with the expectations of auditors from a variety of interested parties.
- Other findings:
  - Largely documentation omissions and errors. We will document our actual practice in a new revision of the DOEGrids CPS (this will appear before the rewrite to RFC 3647).
  - Update of certificate formats will occur gradually in step with disclosure and quality control checks in our communities
- We will schedule another audit perhaps in early 2009.
  - More focused, with additional institutional by-in

DOEGrids Continuity of Operations

- Also discussed in DOEGrids Audit
- Focus on cloning CA and HSM (as noted earlier) to reduce coupling to local (LBNL) infrastructure issues
- Clone and distribute CRL distribution machines
  - Waiting on manpower (2Q 09)
- Local (LBNL) Infrastructure
  - Some improvements have been made to local infrastructure (see elsewhere for details)
  - A more intelligent power management service is in the planning stages
    - The initial step of developing criteria for critical/non critical servers has been done, and a first pass at classification

Policy Management Authorities

- DOEGrids PMA (DOEGrids.org)
  - New chairman: John Volmer, ANL
  - John Volmer is commissioning a “strategy committee” to look at how DOEGrids CA should evolve
  - DOEGrids has added one new RA: US Philips Research
  - DOEGrids is about to add one new RA: ESnet (!)
- TAGPMA.org (The Americas Grid PMA)
  - Web site still at CANARIE (this may need to be addressed soon)
  - Email is stabilized at ESnet (waiting for ESnet mail system transition)
  - Developing Twiki for PMA use
- IGTF.net
  - “gridpma.org” website has transitioned to igtf.net
  - gridpma.org/igtf.net email stabilized at ESnet (waiting for ESnet mail system transition)
  - Wiki waiting on wiki – federation integration

Federation Protocols and Services

- Why is ESnet's InCommon membership interesting?
  - Aligns us with US Academic Shibboleth federation
  - We can offer federation-aware ("Shibbolized") services to InCommon members, as well as other federations that recognize InCommon
  - What kind of services?
    - Network management and resource allocation
    - A-V collaboration services
    - Gateway certification authorities
    - Or ….
  - DOE laboratories have many collaboration issues and requirements
  - Shibboleth can provide a useful platform for normalizing authentication, reducing burdens on projects, and improving security
  - Federation allows sites to maintain local infrastructure and autonomy
  - Cost and technology barriers are non-existent
  - Want to talk about "DOETrust"? Talk to Mike Helm or Adam Stone
- Why is OpenID interesting?
  - OpenID is a simple, small footprint federation protocol
  - Industry response to the problem – one could interpret it as a recasting of Shibboleth/Liberty Alliance
    - Yahoo, Microsoft, Google, AOL, VeriSign & others support this
- There is room for PKI-based, Shibboleth-based (really SAML2-based), and OpenID-based services
  - NOT mutually exclusive
  - Possibly providing different "levels of assurance"

PGP Key Server

- The ESnet PGP key server to be updated
  - Current service based on PKS
  - Replace with one based on SKS
    - SKS is supported better
    - Update protocol works much better (not email – dependent)
  - Email service for PGP key server may not be available immediately
  - Update will allow us to support the service better, and retire some antique equipment

Federation Services: Wikis

- Experimenting with Wiki and client cert authentication
  - Motivation: Eliminate manual registration, scale to a large community
  - **Result**: Difficult to manage Apache configuration supporting both client certificate and non-certificate authentication (seems like an esoteric problem but it's an issue at ESnet).
  - **Result**: client cert authentication is sometimes too limiting (not everyone has a usable certificate).
  - **Result**: We often want more than yes/no: we want groups, roles. We need an IdP that can serve up more attributes.
- Experimenting with Federation protocols and services
  - Shibboleth and OpenID
  - Jan Durand (our summer intern) is building an OpenID Provider for us
    - OpenID 2.0; also studying interoperability issues with OpenID 1.0
    - Integration with DOEGrids certificates and other IdPs
  - Plan to use this as an OP for a wiki service, perhaps for one of the PMAs

Federation Services: ECS

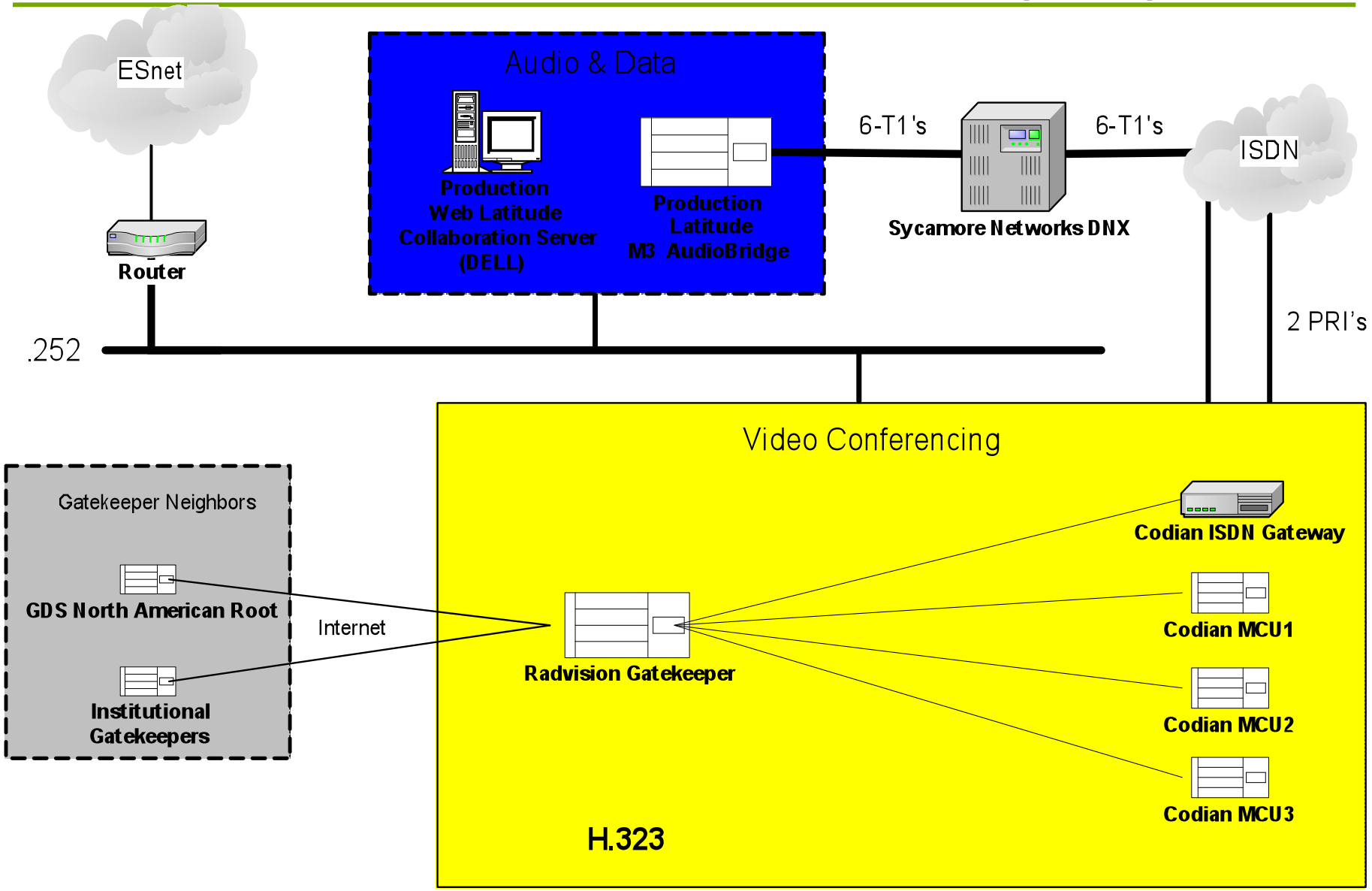
- Exploring applicability of Federation Services to ECS
  - Client certificates, OpenID, Shibboleth, or ?
  - About 21% of the ECS audio bridge registrants appear to have DOEGrids certificates
  - Use cases and comprehensive coverage are difficult
  - ECS is technology driven (AAI necessarily has to follow hardware/architecture choices).

Questions? Consult Stan Kluz or Mike Helm

IIIb. ESnet Conferencing Service (ECS)

- An ESnet Science Service that provides audio, video, and data teleconferencing service to support human collaboration of DOE science
  - Usage is a little higher than last year
  - ECS video serves about 1800 DOE researchers and collaborators worldwide at 270 institutions
    - Videoconferences - 3900 port hours per month, year average
    - Data conferencing - about 230 port hours per month
  - ECS audio serves about 800 DOE researchers and collaborators worldwide at 170 institutions
    - Audio conferencing - about 2500 port hours per month
  - Web-based, automated registration and scheduling for all of these services

ESnet Collaboration Services (ECS)



ECS Video Collaboration Service

- High Quality videoconferencing over IP and ISDN
- Reliable, appliance based architecture
- Ad-Hoc H.323 and H.320 multipoint meeting creation
- Web Streaming options on 3 Codian MCU's using Quicktime or Real
- 3 Codian MCUs with Web Conferencing Options
- 120 total ports of video conferencing on each MCU (40 ports per MCU)
- 384k access for video conferencing systems using ISDN protocol
- Access to audio portion of video conferences through the Codian ISDN Gateway

ECS Voice and Data Collaboration

- 144 usable ports
  - Actual conference ports readily available on the system.
- 144 overbook ports
  - Number of ports reserved to allow for scheduling beyond the number of conference ports readily available on the system.
- 108 Floater Ports
  - Designated for unexpected port needs.
  - Floater ports can float between meetings, taking up the slack when an extra person attends a meeting that is already full and when ports that can be scheduled in advance are not available.
- Audio Conferencing and Data Collaboration using Cisco MeetingPlace
- Data Collaboration = WebEx style desktop sharing and remote viewing of content
- Web-based user registration
- Web-based scheduling of audio / data conferences
- Email notifications of conferences and conference changes
- 800+ users registered to schedule meetings (not including guests)

ECS Futures

- ESnet is still on-track (funds budgeted) to upgrade the teleconferencing hardware currently located at LBNL and provide a replicate at a different location (FNAL) – though this is somewhat delayed
  - Video bridge upgrade is waiting on Tandberg/Codian high port density model which has now been announced (Model 8000) and will be available by Q4 2008 / Q1 2009
  - Once the model 8000 is installed at LBNL, the 3 existing MCUs and the gatekeeper, plus maybe ISDN/Pots gateway, will go to FNAL
  - This will provide for redundant operation and, if California is cut off from the rest of the world, people still can use 3 MCUs at Fermi
- The new equipment is intended to provide at least comparable service to the current (upgraded) ECS system
  - Also intended to provide some level of backup to the current system
  - A new Web based registration and scheduling portal may also come out of this

TANDBERG Codian MSE 8000 Series

Key Differentiators

- High availability – Designed for ETSI/NEBS 3
- All blades, power supplies and fans are hot-swappable
- AC or DC power inputs and multiple power supplies
- MCU, ISDN GW and VCR blades in the same chassis
- Up to 360 MCU video or 720 audio ports, 90 IP VCR ports, or 72 ISDN PRI ports in a single chassis
- Current technology same as MCU 4200 Series
- Scheduled and managed through TMS 11.8+

TANDBERG provided slide

MSE 8510 Media2 Blade

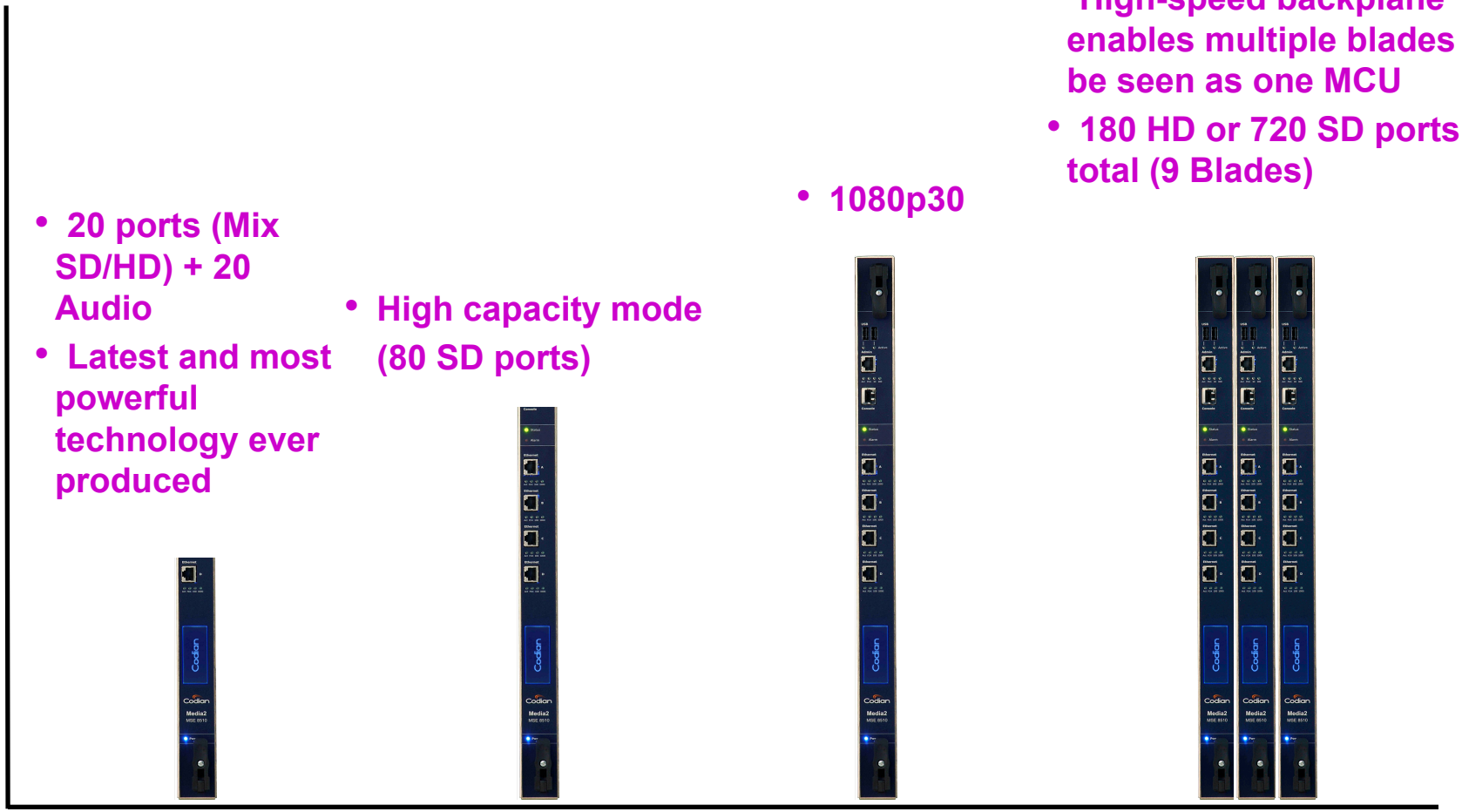
Provide unrivaled HD and Multipoint scalability and is an investment that will grow in capabilities and scale.

- 20 ports (Mix SD/HD) + 20 Audio
- Latest and most powerful technology ever produced

- High capacity mode (80 SD ports)

- 1080p30

- High-speed backplane enables multiple blades to be seen as one MCU
- 180 HD or 720 SD ports in total (9 Blades)



Q3-2008 Q4-2008 Q1-2009 1H-2009

TANDBERG provided slide

ECS Service Level

- ESnet Operations Center is open for service 24x7x365.
- A trouble ticket is opened within15 to 30 minutes and assigned to the appropriate group for investigation.
- Trouble ticket is closed when the problem is resolved.
- ***ECS support*** is provided Monday to Friday, 8AM to 5 PM Pacific Time excluding LBNL holidays
  - Reported problems are addressed within 1 hour from receiving a trouble ticket during ECS support period
  - ESnet does NOT provide a real time (during-conference) support service

Real Time ECS Support

- A number of user groups have requested “real-time” conference support (monitoring of conferences while in-session)
- Limited Human and Financial resources currently prohibit ESnet from:

  A) Making real time information available to the public on the systems status (network, ECS, etc) This information is available only on some systems to our support personnel

  B) 24x7x365 real-time support

  C) Addressing simultaneous trouble calls as in a real time support environment.

  - This would require several people addressing multiple problems simultaneously

Real Time ECS Support

- Solution
  - A fee-for-service arrangement for real-time conference support
  - Available from TKO Video Communications, ESnet's ECS service contractor
  - Service offering could provide:
    - Testing and configuration assistance prior to your conference
    - Creation and scheduling of your conferences on ECS Hardware
    - Preferred port reservations on ECS video and voice systems
    - Connection assistance and coordination with participants
    - Endpoint troubleshooting
    - Live phone support during conferences
    - Seasoned staff and years of experience in the video conferencing industry
    - ESnet community pricing

IIIc. Enhanced Collaboration Services

- The Fusion community has outlined the need for significant advances in collaboration technology
  - The challenge is to effectively collaborate with a remote Tokamak control room
  - Current Tokamak control rooms (e.g. DIII-D at General Atomics) already employ any technology they can get their hands on – VRVS, H.323, Instant Messaging, Access Grid, Skype, etc.
  - PIs still travel to the instrument to run experiments
    - current remote collaboration technology still not good enough
    - ITER assumes this will be solved – there is no plan for a large central control room for ITER
  - Collaboration tools need to be integrated with a federated security framework
  - This defines a clear and present research priority

R&D

➢Summary

- Transition to ESnet4 is going smoothly
  - New network services to support large-scale science are progressing
  - OSCARS virtual circuit service is being used, and the service functionality is adapting to unforeseen user needs
  - Measurement infrastructure is rapidly becoming widely enough deployed to be very useful
- Revaluation of the 5 yr strategy indicates that the future will not be qualitatively the same as the past – and this must be addressed
  - R&D, testbeds, planning, new strategy, etc.
- New ESC hardware and service contract are working well
  - Plans to deploy replicate service are delayed to early CY 2009
- Federated trust - PKI policy and Certification Authorities
  - Service continues to pick up users at a pretty steady rate
  - Maturing of service - and PKI use in the science community generally

References

[OSCARS]

For more information contact Chin Guok (chin@es.net). Also see

- http://www.es.net/oscars

[Workshops]

see http://www.es.net/hypertext/requirements.html

[LHC/CMS]

http://cmsdoc.cern.ch/cms/aprom/phedex/prod/Activity::RatePlots?view=global

[ICFA SCIC] "Networking for High Energy Physics." International Committee for Future Accelerators (ICFA), Standing Committee on Inter-Regional Connectivity (SCIC), Professor Harvey Newman, Caltech, Chairperson.

- http://monalisa.caltech.edu:8080/Slides/ICFASCIC2007/

[E2EMON] Geant2 E2E Monitoring System –developed and operated by JRA4/WI3, with implementation done at DFN
http://cnmdev.lrz-muenchen.de/e2e/html/G2_E2E_index.html
http://cnmdev.lrz-muenchen.de/e2e/lhc/G2_E2E_index.html

[TrViz] ESnet PerfSONAR Traceroute Visualizer
https://performance.es.net/cgi-bin/level0/perfsonar-trace.cgi