The Future of ESnet

ASCAC Meeting, April, 2004

William E. Johnston, ESnet Manager and Senior Scientist

Michael S. Collins, Stan Kluz,
Joseph Burrescia, and James V. Gagliardi, ESnet Leads

and the ESnet Team

Lawrence Berkeley National Laboratory

Outline

- ➢ What is ESnet Today?
- Trends, Opportunities, and Pressures
- ESnet's Evolution over the Next 10-20 Years

What is ESnet Today?

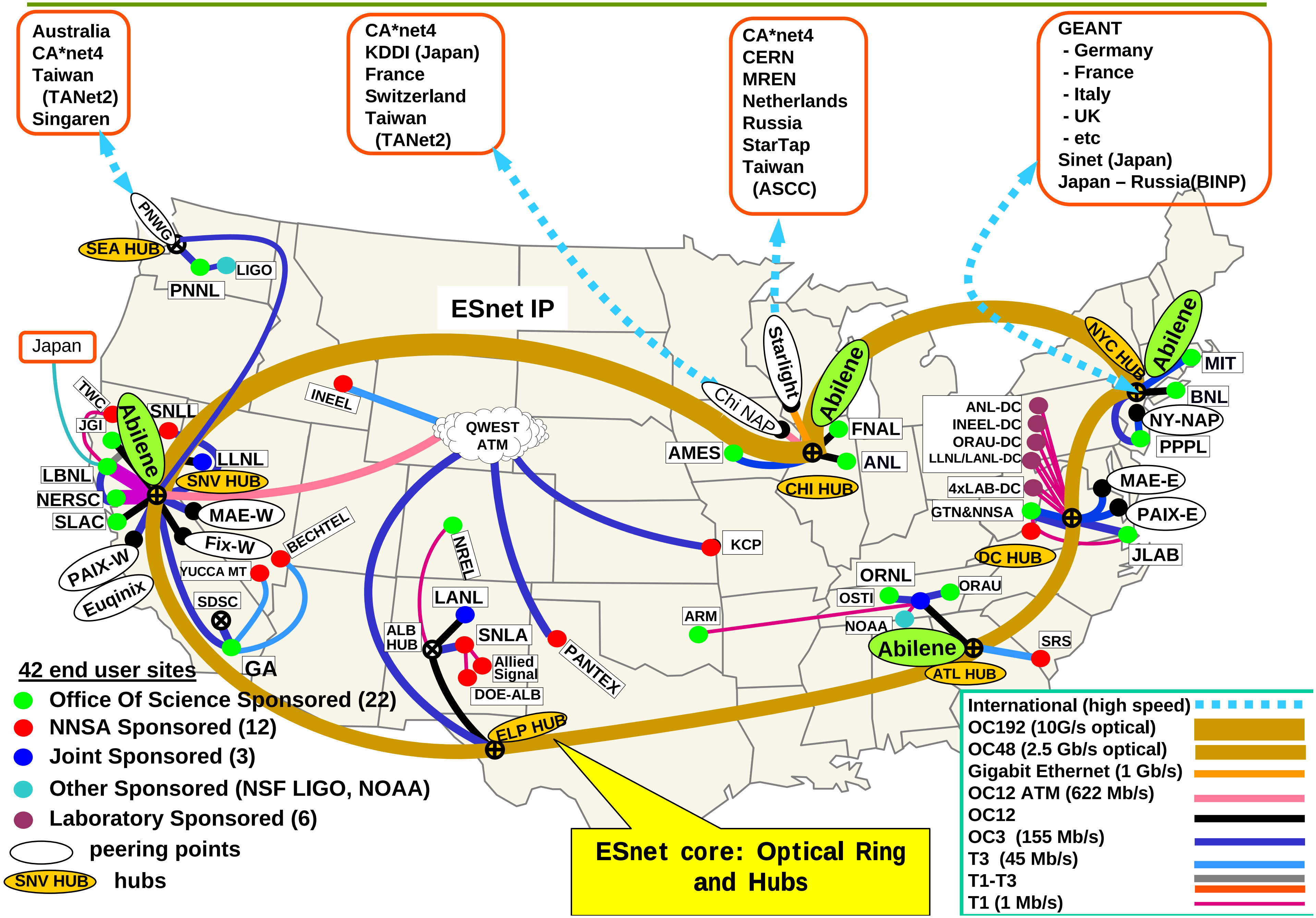
- Architected to move huge amounts of data between a small number of sites
- High bandwidth peering for access to the US, European, Japanese, and other Research and Education networks
- Access to the global Internet (managing 150,000 routes at 10 commercial peering points)
- Comprehensive user support, including “owning” all trouble tickets involving ESnet users (including problems at the far end of an ESnet connection) until they are resolved – 24x7 coverage
- Grid and collaboration services supporting science
  - trust, persistence, and science oriented policy
- Primarily OSC focused, but supports NNSA/Defense Programs, including SecureNet as an overlay network

What is ESnet Today?

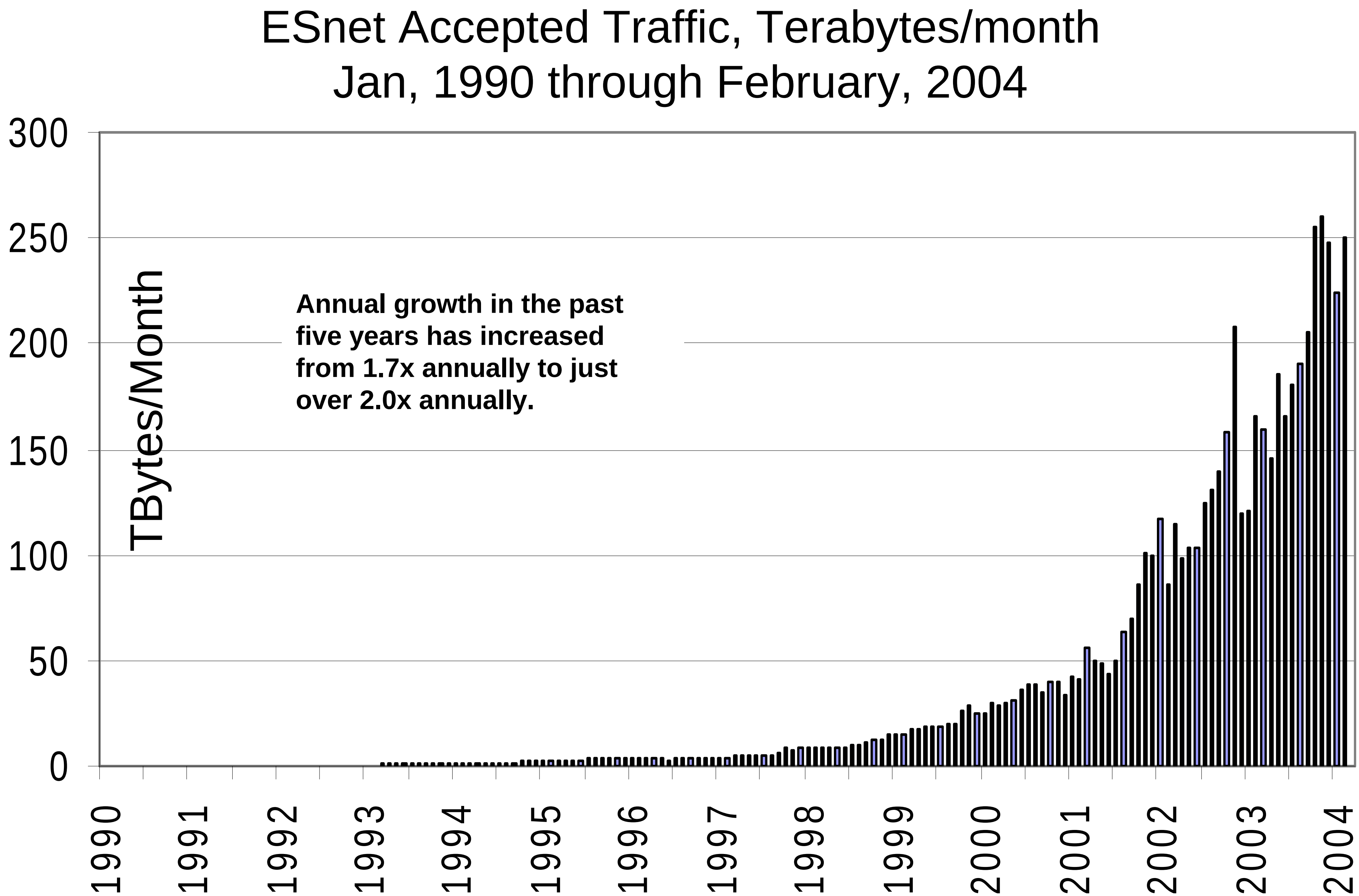
- A community endeavor
  - Strategic guidance from the OSC programs (ESSC)
  - Network operation is a shared activity with the community – via ESCC – which ensures the right operational “sociology” for success
- Complex and specialized – both in the network engineering and the network management – in order to provide its services to the Labs in an integrated support environment
- Extremely reliable in several dimensions

➢ Taken together these points make ESnet a unique facility supporting DOE science that is quite different from a commercial ISP or University network

ESnet Connects DOE Facilities and Collaborators



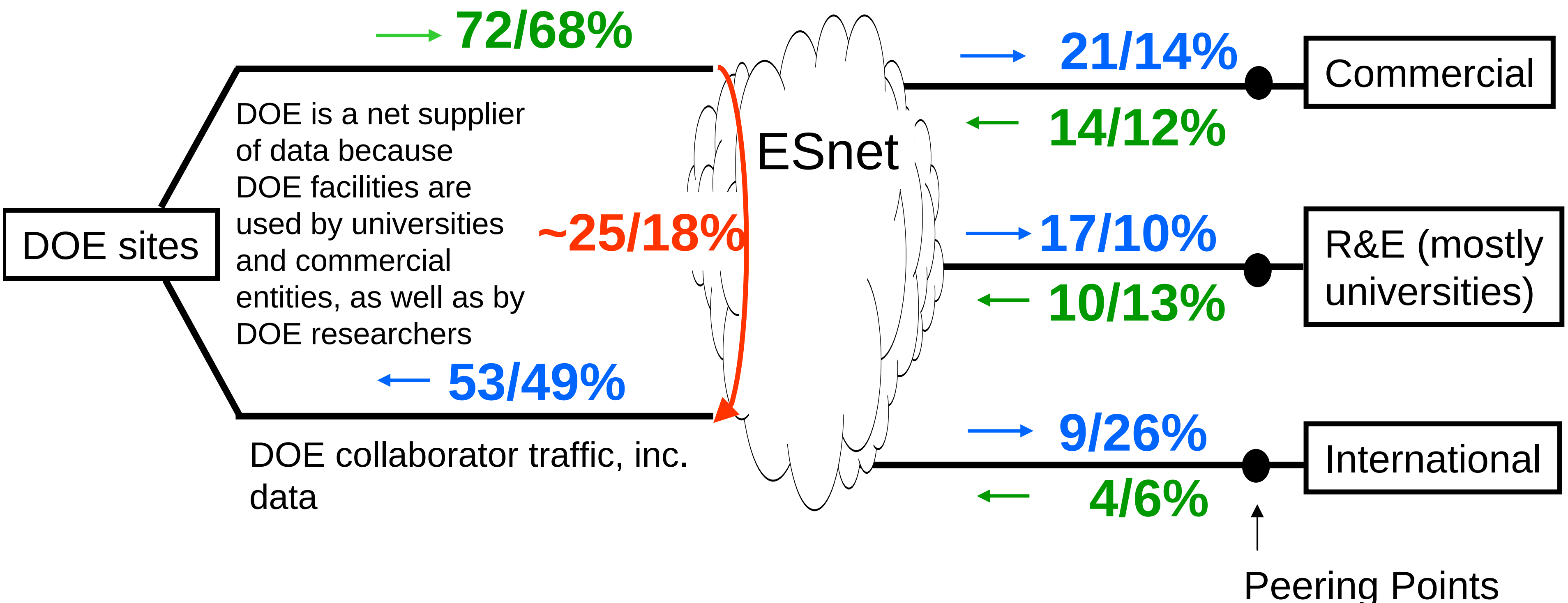
ESnet Accommodates Exponentially Increasing Traffic



Who Generates Traffic, and Where Does it Go?

ESnet Inter-Sector Traffic Summary, Jan 2003 / Feb 2004 (1.7X overall traffic increase, 1.9X OSC increase)

(the international traffic is increasing due to BABAR at SLAC and the LHC tier 1 centers at FNAL and BNL)



<u>Note that more that 90% of the ESnet traffic is OSC traffic</u>

<u>ESnet Appropriate Use Policy (AUP)</u>

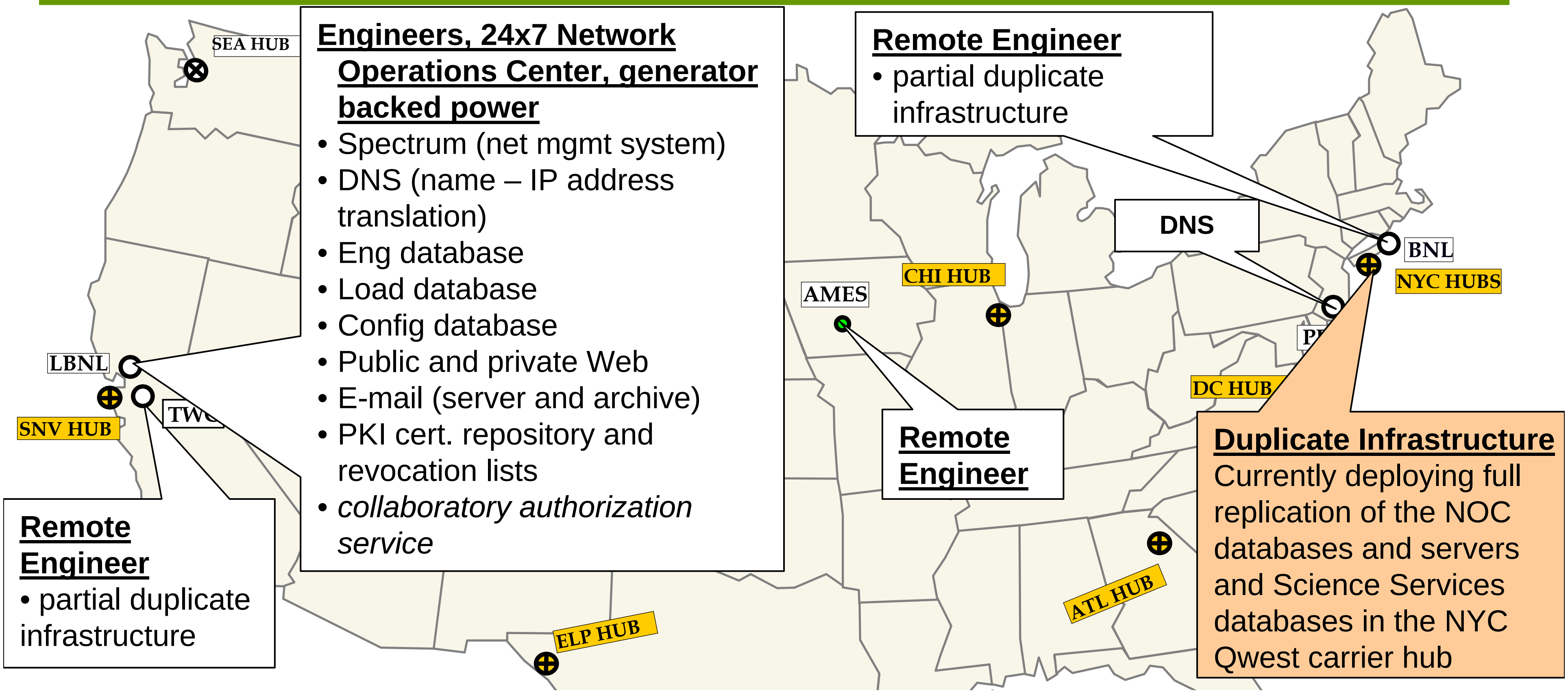
All ESnet traffic must originate and/or terminate on an ESnet an site (no transit traffic is allowed)

Traffic coming into ESnet = Green
Traffic leaving ESnet = Blue
Traffic between sites
% = of total ingress or egress traffic

Science Mission Critical Infrastructure

- ESnet is a visible and critical piece of DOE science infrastructure
  - if ESnet fails,10s of thousands of DOE and University users know it within minutes if not seconds
- <u>Requires high reliability and high operational security in both the network and in the ESnet infrastructure support</u> – the systems that support the operation and management of the network and services/
  - Secure and redundant mail and Web systems are central to the operation and security of ESnet
    - trouble tickets are by email
    - engineering communication by email
    - engineering database interface is via Web
  - Secure network access to Hub equipment
  - Backup secure telephony access to all routers
  - 24x7 help desk (joint w/ NERSC) and 24x7 on-call network engineers

Disaster Recovery and Stability

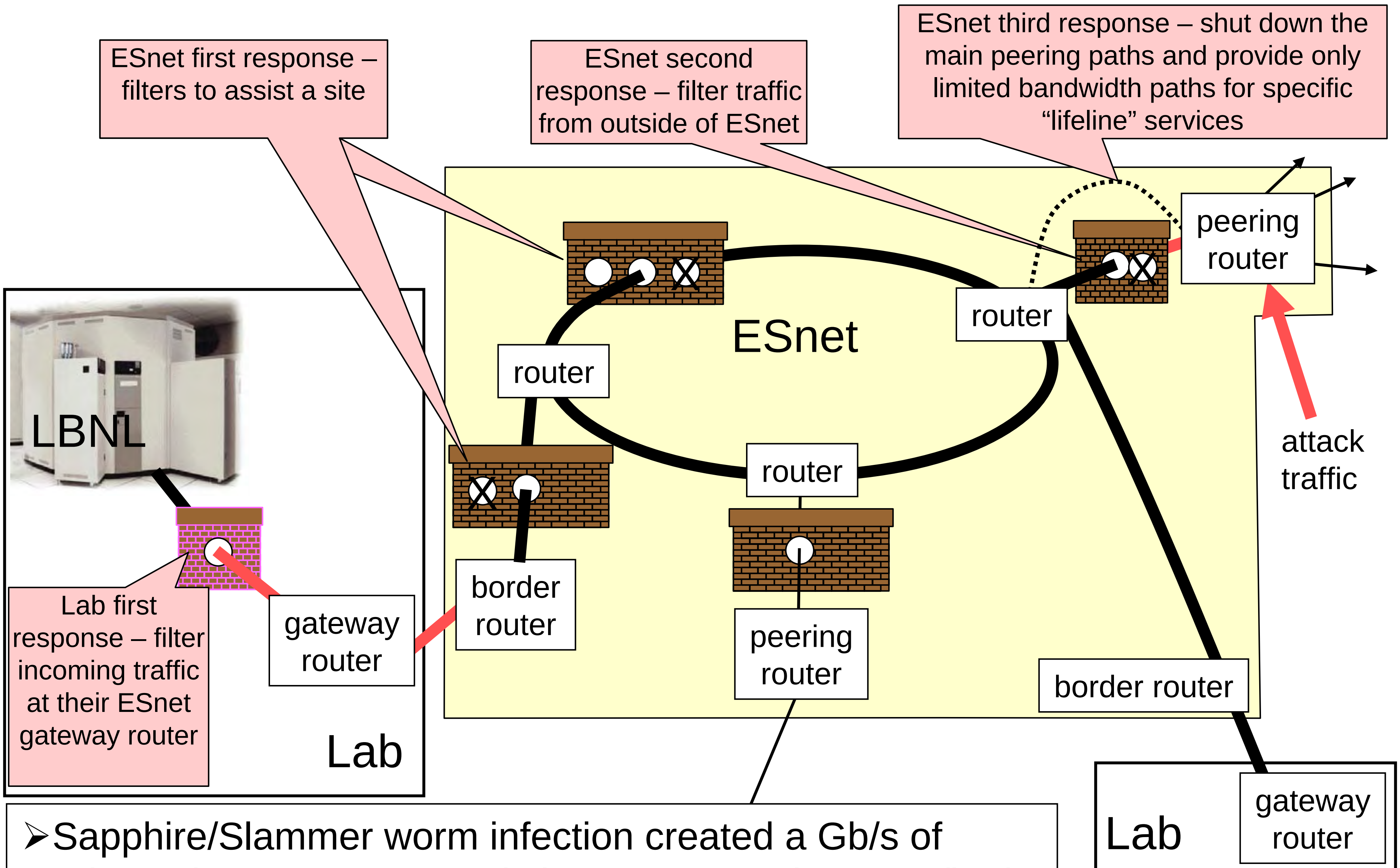


- The network must be kept available even if, e.g., the West Coast is disabled by a massive earthquake, etc.

Reliable operation of the network involves
- remote NOCs
- replicated support infrastructure
- generator backed UPS power at all critical network and infrastructure locations

- non-interruptible core - **ESnet core operated without interruption** through
  - N. Calif. Power blackout of 2000
  - the 9/11/2001 attacks, and
  - the Sept., 2003 NE States power blackout

Cyberattack Defense



➢Sapphire/Slammer worm infection created a Gb/s of traffic on the ESnet core until filters were put in place (both into and out of sites) to damp it out.

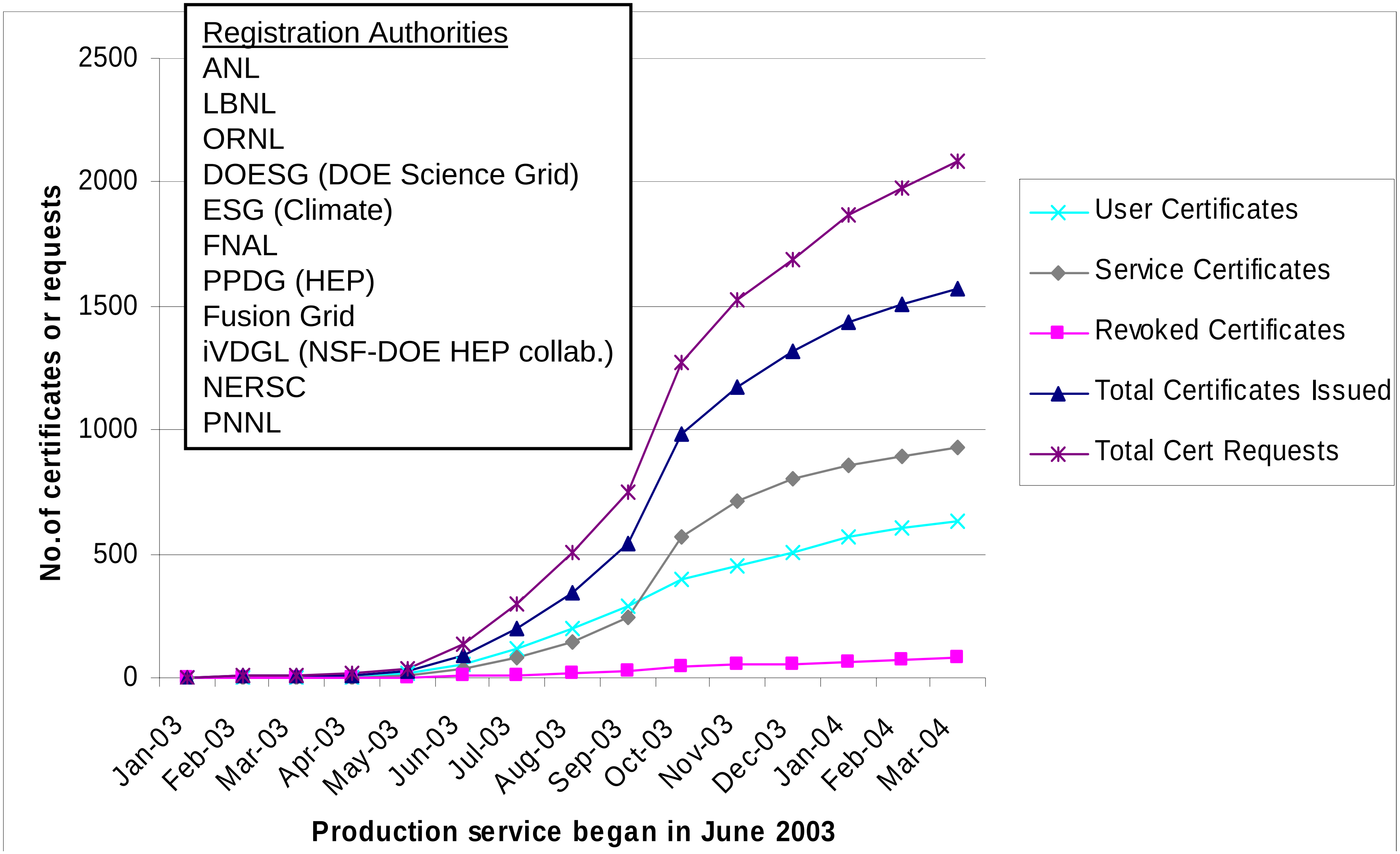
Typical Equipment of an ESnet Core Network Hub

Science Services: PKI Support for Grids

- X.509 identity certificates and Public Key Infrastructure provides the basis of secure, cross-site authentication of people and systems (www.doegrids.org)
  - Certification Authority (CA) issues certificates after validating request against policy
  - CA negotiates the cross-site, cross-organization, and international <u>trust relationships to provide policies that are tailored to collaborative science in order to permit sharing computing and data resources,</u> and other Grid services
  - This service was the basis of the first routine sharing of HEP computing resources between US and Europe
- Have recently deployed a second CA with a policy that supports secondary issuers that need to do bulk issuing of certificates with centralized private key management
  - NERSC will auto issue certs when accounts are set up – this constitutes an acceptable identity verification
  - A variant of this will also be set up to support security domain gateways such as Kerberos – X509 – e.g. KX509 – at FNAL

Science Services: Public Key Infrastructure

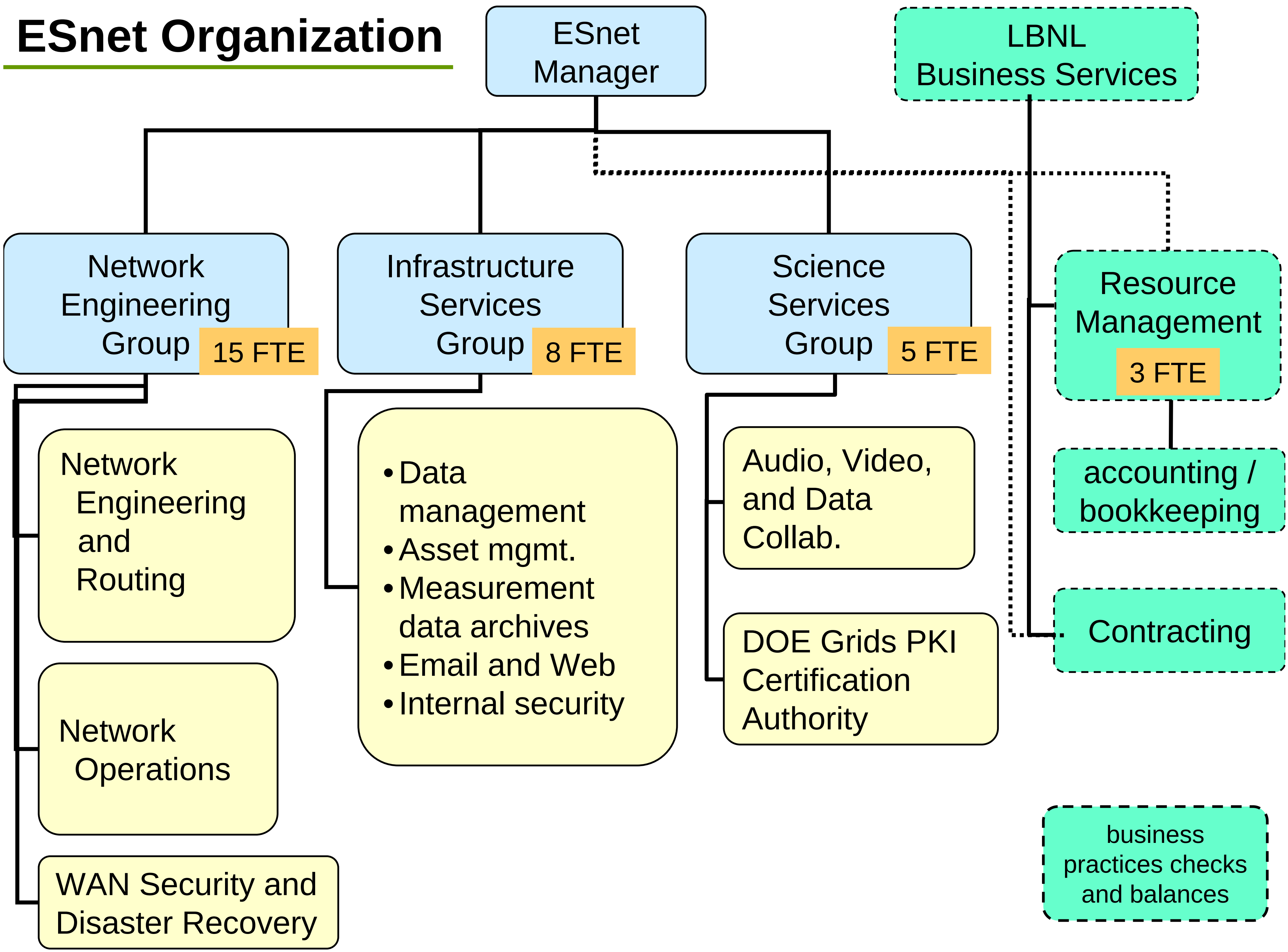
- The rapidly expanding customer base of this service will soon make it ESnet's largest collaboration service by customer count



Voice, Video, and Data Collaboration Service

- Another highly successful ESnet Science Service is the audio, video, and data teleconferencing service to support human collaboration
  - o Seamless voice, video, and data teleconferencing is important for geographically dispersed scientific collaborators
  - o ESnet currently provides to more than a thousand DOE researchers and collaborators worldwide
    - H.320 (ISDN) videoconferences (4600 port hours per month) (fading)
    - H.323 (IP) videoconferences (1100 port hours per month) (rising)
    - audio conferencing (2000 port hours per month) (constant)
    - data conferencing (100 port hours per month)
    - Web-based, automated registration and scheduling for all of these services

ESnet Organization



Overall FY04 Budget ($M)

Fiscal Year 2004 (approx.)	MICS	ICO/Other (non-MICS)	Total
Operating	5.4	0.6	6.0
Material and Services	2.3	0.3	2.6
Infrastructure Maintenance	0.9	0.1	1.0
Communications	11.0	2.7	13.7
Total	19.6	4.2	23.3

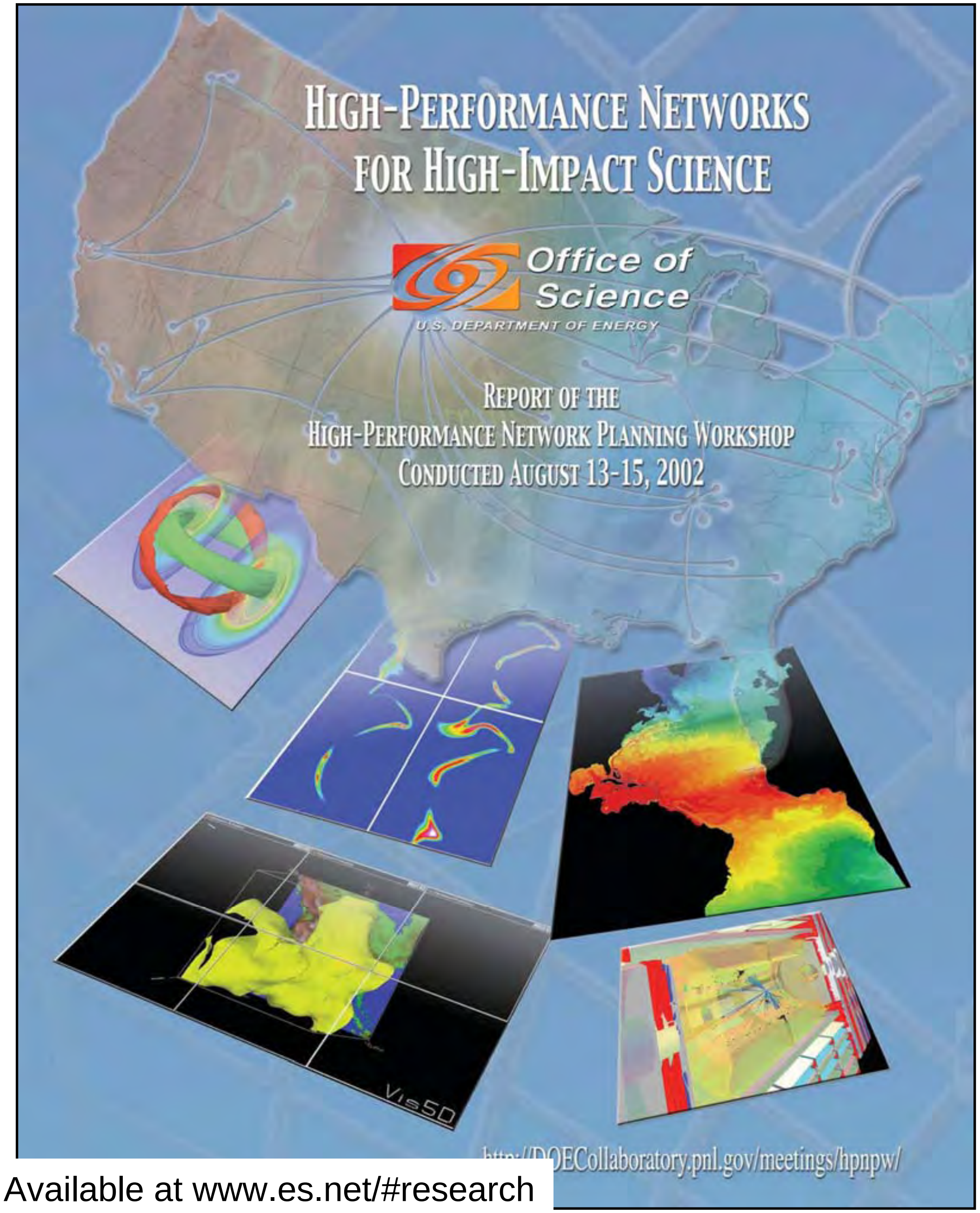
Outline

- Forward
- ➢Trends, Opportunities, and Pressures
- ESnet's Evolution over the Next 10-20 Years

Trends, Opportunities, and Pressures

- ESnet is on a cusp of change
- ESnet needs to evolve fairly rapidly from this point forward due to:
  - Trends:
    - rapidly changing requirements for supporting science require increased capabilities – bigger science requires larger networks
    - DOE has new facilities coming online – i.e. SNS, LHC etc.
  - Opportunities:
    - current availability of fiber provides opportunity for a new ESnet architecture that addresses several major Roadmap requirements
  - Pressures:
    - however, the opportunity decreases as the fiber market hardens after several soft years
    - ESnet’s current budget is flat thru FY2006

➢ESnet is Driven by the Needs of DOE Science



Available at www.es.net/#research

August 13-15, 2002

Organized by Office of Science

- Mary Anne Scott, Chair
- Dave Bader
- Steve Eckstrand
- Marvin Frazier
- Dale Koelling
- Vicky White

Workshop Panel Chairs

- Ray Bair and Deb Agarwal
- Bill Johnston and Mike Wilde
- Rick Stevens
- Ian Foster and Dennis Gannon
- Linda Winkler and Brian Tierney
- Sandy Merola and Charlie Catlett

Focused on science requirements that drive

- **Advanced Network Infrastructure**
- **Middleware Research**
- **Network Research**
- **Network Governance Model**

Eight Major DOE Science Areas Analyzed at the August '02 Workshop

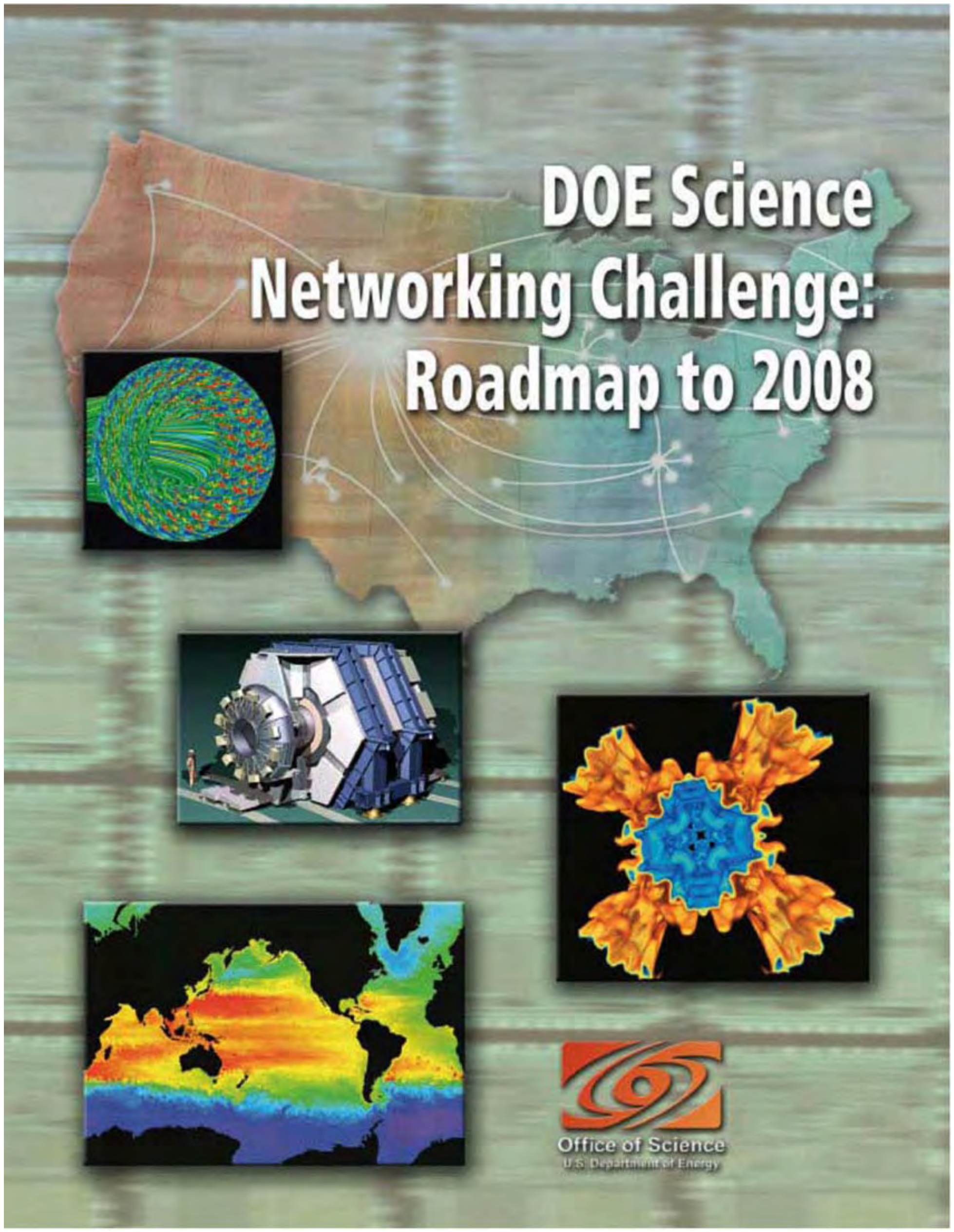
analysis was driven by

Feature / Discipline	Vision for the Future Process of Science	Characteristics that Motivate High Speed Nets	Requirements: Networking	Requirements: Middleware
Climate (near term)	Analysis of model data by selected communities that have high speed networking (e.g. NCAR and NERSC)	• <u>A few data repositories, many distributed computing sites</u> • NCAR - 20 TBy • NERSC - 40 TBy • ORNL - 40 TBy	• Authenticated data streams for easier site access through firewalls	• Server side data processing (computing and cache embedded in the net) • Information servers for global data catalogues
Climate (5 yr)	Enable the analysis of model data by all of the collaborating community	• Add many simulation elements/components as understanding increases • 100 TBy / 100 yr generated simulation data, 1-5 PBy / yr (just at NCAR) o <u>Distribute large chunks of data to major users for post-simulation analysis</u>	• Robust access to large quantities of data	• Reliable data/file transfer (across system / network failures)
Climate (5-10 yr)	Integrated climate simulation that includes all high-impact factors	• 5-10 PBy/yr (at NCAR) • Add many diverse simulation elements/components, including from other disciplines - this must be <u>done with distributed, multidisciplinary simulation</u> • <u>Virtualized data to reduce storage load</u>	• Robust networks supporting distributed simulation - adequate bandwidth and latency for remote analysis and visualization of massive datasets	• Quality of service guarantees for distributed, simulations • Virtual data catalogues and work planners for reconstituting the data on demand

Evolving Quantitative Science Requirements for Networks

Science Areas	Today *End2End* Throughput	5 years End2End Throughput	5-10 Years End2End Throughput	Remarks
High Energy Physics	0.5 Gb/s	100 Gb/s	**1000 Gb/s**	high bulk throughput
Climate (Data & Computation)	0.5 Gb/s	160-200 Gb/s	**N x 1000 Gb/s**	high bulk throughput
SNS NanoScience	Not yet started	1 Gb/s	**1000 Gb/s + QoS for control channel**	remote control and time critical throughput
Fusion Energy	0.066 Gb/s (500 MB/s burst)	0.198 Gb/s (500MB/ 20 sec. burst)	**N x 1000 Gb/s**	time critical throughput
Astrophysics	0.013 Gb/s (1 TBy/week)	N*N multicast	**1000 Gb/s**	computational steering and collaborations
Genomics Data & Computation	0.091 Gb/s (1 TBy/day)	100s of users	**1000 Gb/s + QoS for control channel**	high throughput and steering

New Strategic Directions to Address Needs of DOE Science



June 3-5, 2003

<u>Organized by the ESSC</u>

Workshop Chair

Roy Whitney, JLAB

Report Editors

Roy Whitney, JLAB

Larry Price, ANL

Workshop Panel Chairs

Wu-chun Feng, LANL
William Johnston, LBNL
Nagi Rao, ORNL
David Schissel, GA
Vicky White, FNAL
Dean Williams, LLNL

<u>Focused on what is needed to achieve the science driven network requirements of the previous workshop</u>

- **Both Workshop reports are available at <u>www.es.net/#research</u>**

Outline

- What is ESnet Today?
- Trends, Opportunities, and Pressures
- ➢ ESnet's Evolution over the Next 10-20 Years

ESnet's Evolution over the Next 10-20 Years

- Upgrading ESnet to accommodate the anticipated increase from the current 100%/yr traffic growth to 300%/yr over the next 5-10 years is priority number 7 out of 20 in DOE's "Facilities for the Future of Science – A Twenty Year Outlook"
- Based on the requirements of the OSC High Impact Science Workshop and Network 2008 Roadmap, ESnet must address

  I. Capable, scalable, and reliable production IP networking
    - University and international collaborator connectivity
    - Scalable, reliable, and high bandwidth site connectivity

  II. Network support of high-impact science
    - provisioned circuits with guaranteed quality of service (e.g. dedicated bandwidth)

  III. Evolution to optical switched networks
    - Partnership with UltraScienceNet
    - Close collaboration with the network R&D community

  IV. Science Services to support Grids, collaboratories, etc

I. Production IP: University and International Connectivity

- Connectivity between any DOE Lab and any Major University should be as good as ESnet connectivity between DOE Labs and Abilene connectivity between Universities
  - Partnership with Internet2/Abilene
  - Multiple high-speed peering points
  - Routing tailored to take advantage of this
  - Continuous monitoring infrastructure to verify
  - Status: ***In progress***
    - 3 of 5 cross-connects are in place and carrying traffic
    - first phase monitoring infrastructure is planned for end of April
- 10 Gb/s ring in NYC to MANLAN for 10 Gb/s ESnet – for Abilene x-connect and for international links
- 10 Gb/s ring to StarLight for CERN link, etc.
  - Status: ***Both of these are in progress***

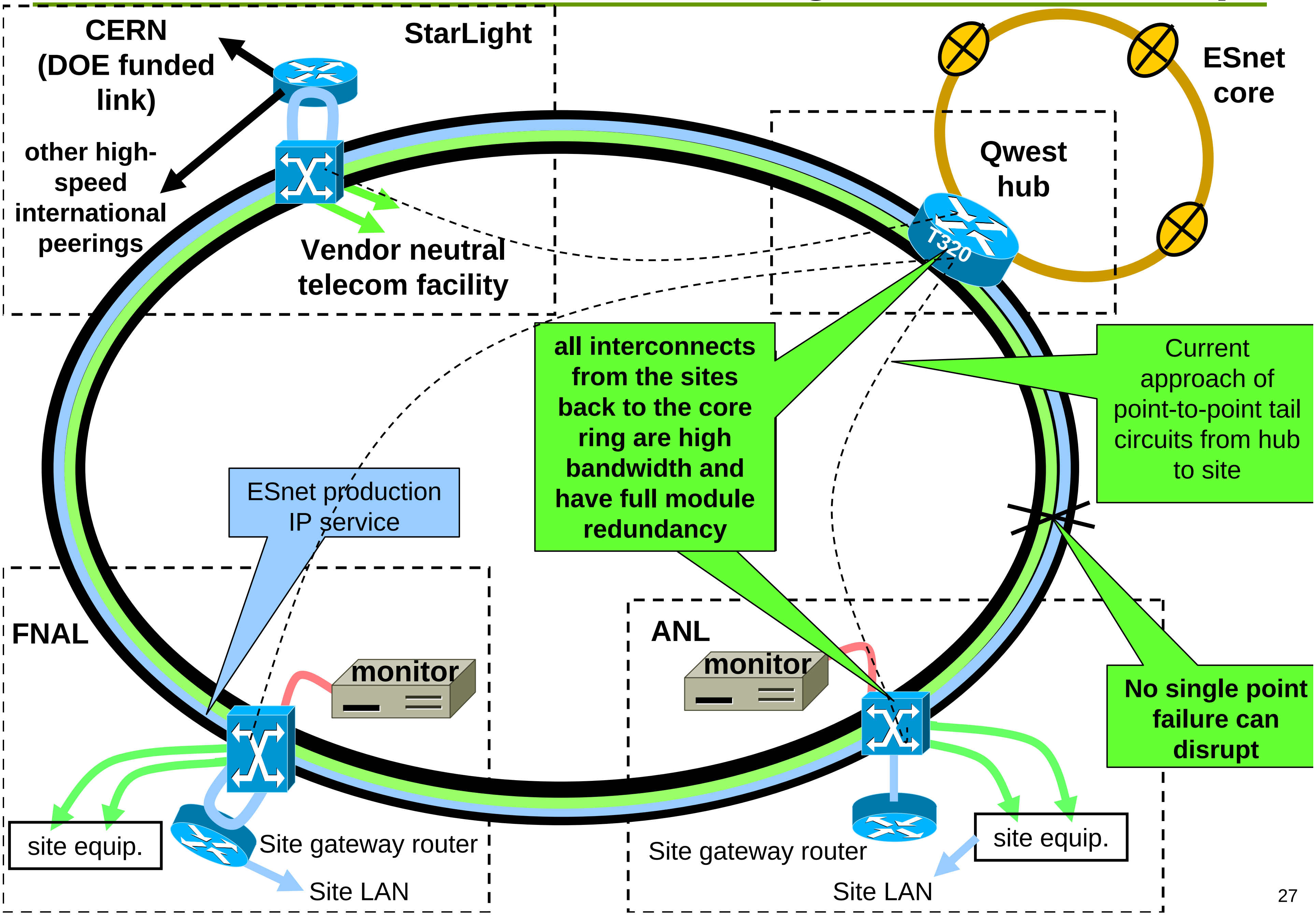
I. Production IP: A New ESnet Architecture

- Goal is local rings, like the core, that provide multiple paths and scalable bandwidth from the ESnet core to the sites – no single points of failure

- Fiber / lambda ring based Metropolitan Area Networks can be built in several important areas (SF Bay, Chicago, Long Island)

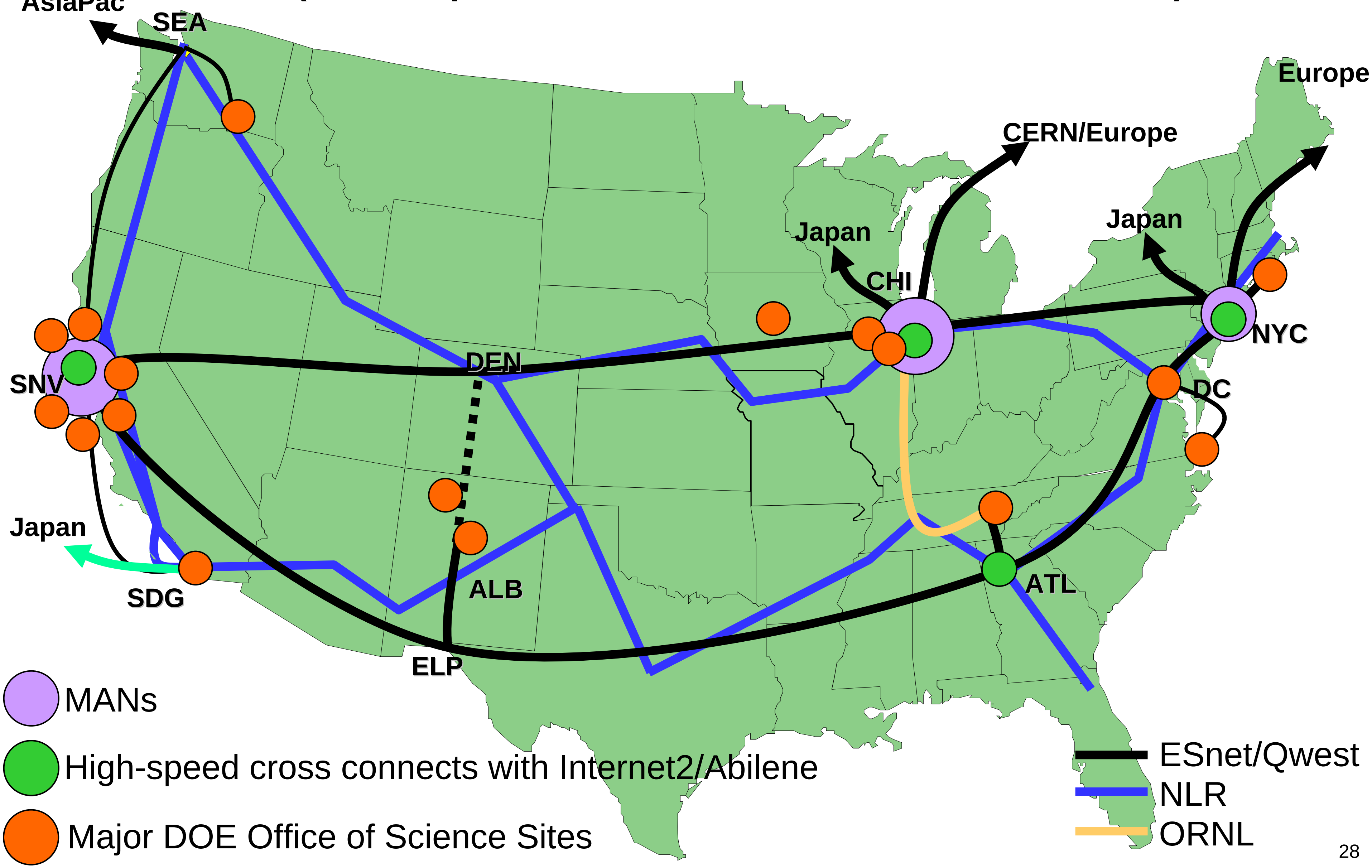
- Status: ***In progress***
  - Migrate site local loops to ring structured Metropolitan Area Networks and regional nets in some areas
  - Preliminary engineering study completed for San Francisco Bay Area and Chicago area
  - Proposal submitted
  - These will most likely be started this year

New ESnet Architecture – Chicago MAN as Example



I. Production IP: Long-Term ESnet Connectivity Goal

- Connecting MANs with two cores to ensure against hub failure (for example, NLR is shown as the second core below)



I. Production IP: Scalable and Reliable Site Connectivity
Leverage and Amplify Non-ESnet Network Connectivity to Labs

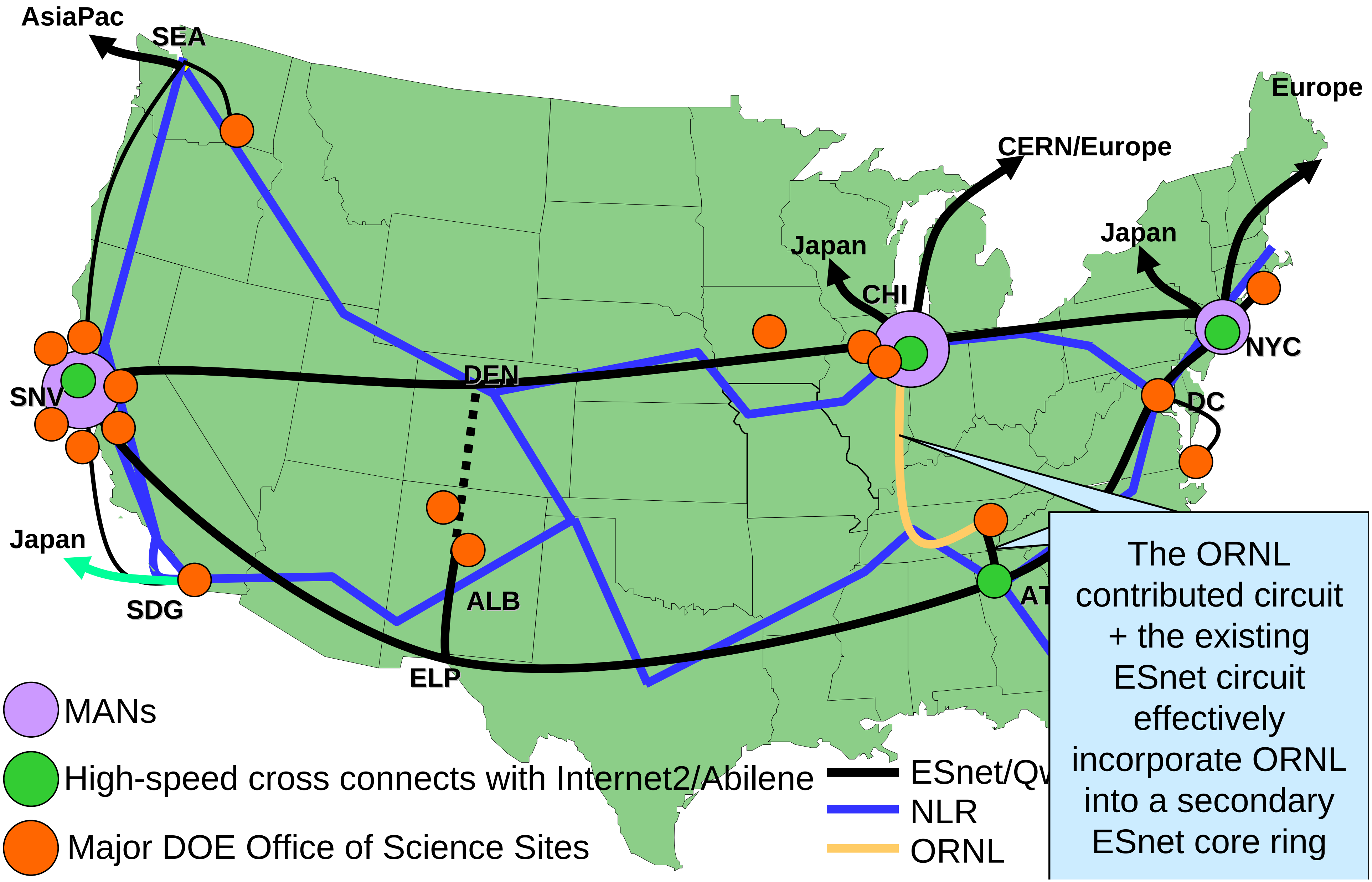
- When ESnet has not been able to afford to increase the site bandwidth, the Labs have sometimes gotten their own high-speed connections

- ESnet can take advantage of this to provide reliable, production high-speed access to the Labs

➢ When possible, incorporate the existing non-ESnet connections into the new ESnet architecture to provide a better and more capable service than the Labs can provide on their own

I. Production IP: Scalable and Reliable Site Connectivity

ORNL Connection to ESnet



I. Production IP: Scalable and Reliable Site Connectivity

Long-Term ESnet Bandwidth Goal

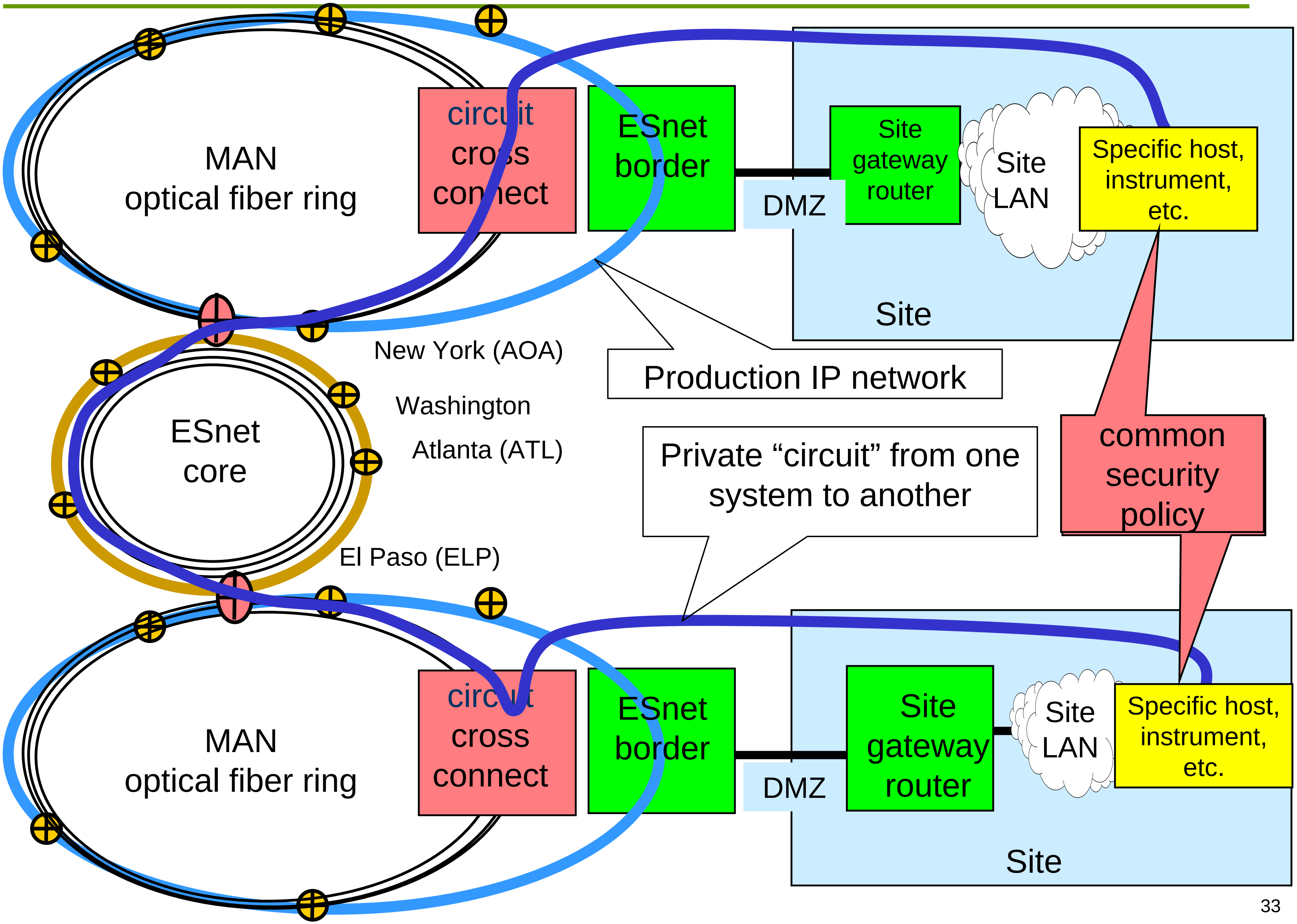
- Harvey Newman:
  "And what about increasing the bandwidth in the core?"
- Answer: technology progress
  - By 2008 (the next generation ESnet core) DWDM technology will be 40 Gb/s per lambda
  - And the core will be multiple lambdas
- Issues
  - How can applications use the high bandwidth networks – must address end-to-end performance (!!)

II. Network Support of High-Impact Science

- Dynamic provisioning of private “circuits” in the MAN and through the core can provide “high impact science” connections with Quality of Service guarantees
  - A few high and guaranteed bandwidth circuits and many lower bandwidth circuits (e.g. for video, remote instrument operation, etc.)
  - The circuits are secure and end-to-end, so if the sites trust each other, and if they have compatible security policies, they should be able to establish direct connections by going around site firewalls to connect specific systems – e.g. HPSS <-> HPSS

- Status: ***Initial progress***
  - Proposal submitted to MICS Network R&D program for initial development of basic circuit provisioning infrastructure in ESnet core network (site to site)
  - Will work with UltraScience Net to import advanced services technology

II. Hi-Impact Science Bandwidth



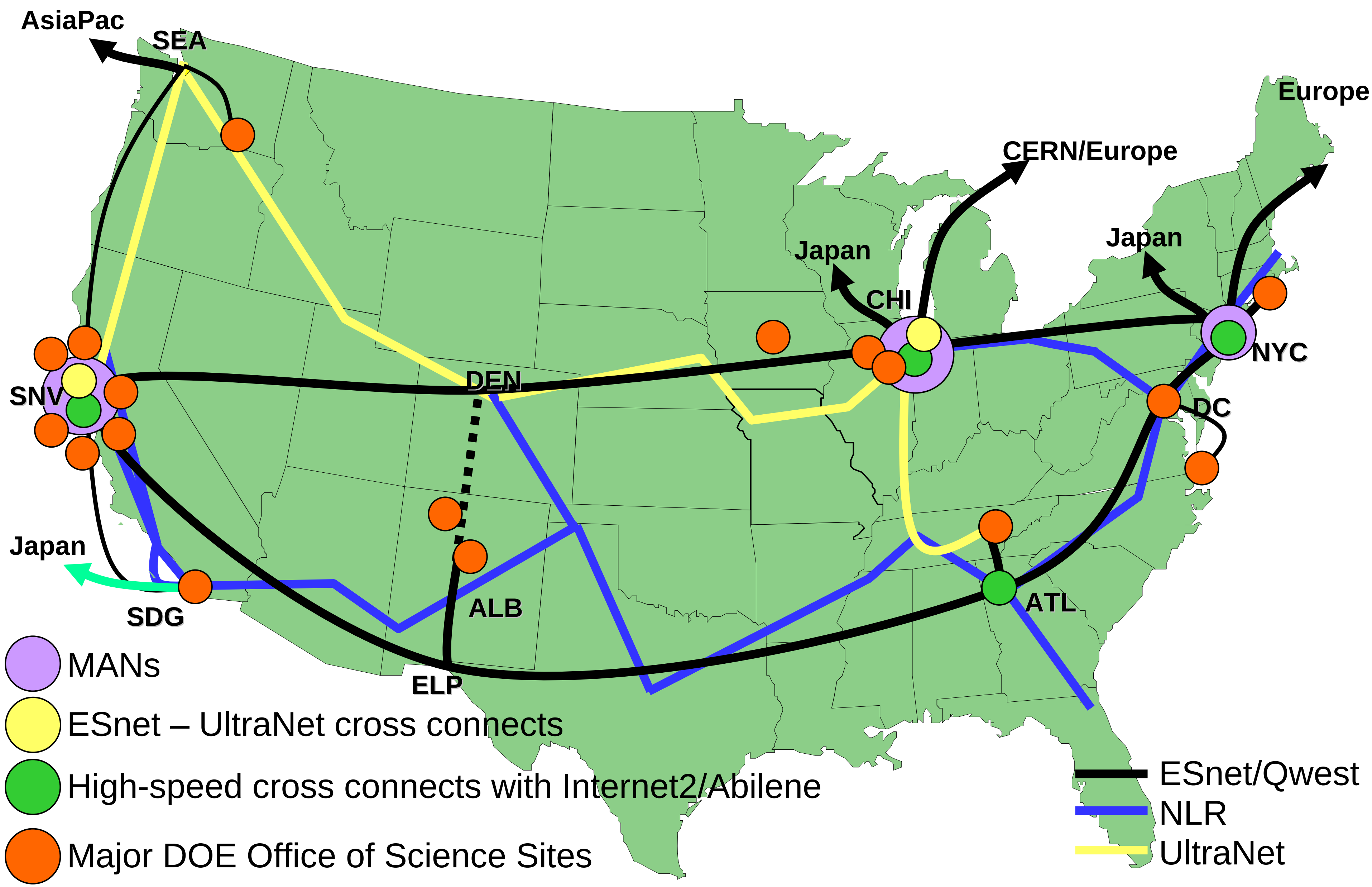
III. Evolution to Optical Switched Networks

- Partnership with DOE's network R&D program
  - ESnet will cross-connect with UltraNet / National Lambda Rail in Chicago and Sunnyvale, CA
  - ESnet can experiment with UltraScience Net virtual circuits tunneled through the ESnet core (up to 5 Gb/s between UltraNet and appropriately connected Labs)
  - <u>One important element of importing DOE R&D into ESnet</u>
  - Status: <u>***In progress***</u>
    - Chicago ESnet – NLR/UltraNet x-connect based on the IWire ring is engineered
      - This is also critical for DOE lab connectivity to the DOE funded LHCNet 10 Gb/s link to CERN
    - Qwest – ESnet Sunnyvale hub x-connect is dependent on Qwest permission, which is being negotiated (almost complete)

III. Evolution to Optical Switched Networks

- ESnet is building partnerships with the Federal and academic R&D networks in addition to DOE network R&D programs and UltraScienceNet
  - o Internet2 Hybrid Optical Packet Internet (HOPI) and National Lambda Rail for R&D on the next generation hybrid IP packet – circuit switched networks
  - o Status: ***<u>Initial progress</u>***
    - ESnet co-organized a Federal networking workshop on the future issues for interoperability of Optical Switched Networks
    - ESnet is participating in the Internet2 HOPI design team (where UltraScience Net also participates)
- These partnerships will provide ESnet with direct access to, and participation in, next generation technology for evaluation and early deployment in ESnet

III. Evolution to Optical Switched Networks UltraNet – ESnet Interconnects



IV. Science Services Strategy

- The Roadmap Workshop identified twelve high priority middleware services, and several of these fit the criteria for ESnet support. These include, for example
  - long-term PKI key and proxy credential management (e.g. an adaptation of the NSF's MyProxy service)
  - directory services that virtual organizations (VOs) can use to manage organization membership, member attributes and privileges
  - end-to-end monitoring for Grid / distributed application debugging and tuning
  - perhaps some form of authorization service
  - knowledge management services that have the characteristics of an ESnet service are also likely to be important (future)
- ESnet will seek the addition funding necessary to develop, deploy, and support these types of middleware services

Conclusions

- ESnet is an infrastructure that is critical to DOE's science mission
- Focused on the Office of Science Labs, but serves many other parts of DOE
- ESnet is working hard to meet the current and future networking need of DOE mission science in several ways:
  - Evolving a new high speed, high reliability, leveraged architecture
  - Championing several new initiatives which will keep ESnet's contributions relevant to the needs of our community

William E. Johnston Bio

- Formerly Department Head of LBNL Distributed Systems Department
- Long history in High Performance Networking Community
- 1980s -1998 PI or Co-PI for
  - LBL Network Advisory Group
  - Advised NSF on NSF backbone transition to commercial service
  - Chaired the ESnet Site Coordinating Committee for 5 yrs
  - Blanca/XUnet - first x-country ATM network (w/ATT Bell Labs)
  - BAGnet – first OC3 (155 Mb/s) ATM net around the SF Bay Area (w/ Pac Bell)
  - MAGIC – DARPA testbed, 1st Sprint OC48 ATM wide area network that worked
  - NGI QoS – DOE bandwidth reservation network, w/ ESnet
  - Clipper – first sustained transfer of terabyte files for HEP, filling an OC12 circuit
- 1998-2003 NASA project manager for an $18M/yr Grids project
  - $6M/yr in external subcontracts
- 2000-2003 PI, DOE Science Grid
- March 2002, Co-Author of LBNL/ANL, <u>A Vision for DOE Scientific Networking driven by High Impact Science</u>
- August 2002, Co-Author of Office of Science Workshop, <u>High Performance Networks for High Impact Science</u>
- June 2003, Co-Author of Office of Science Workshop, <u>DOE Science Networking Challenge: Roadmap to 2008</u>