Energy Sciences Network (ESnet) Futures

Chinese American Networking Symposium
November 30 – December 2, 2004

Joseph Burrescia, Senior Network Engineer

William E. Johnston, ESnet Dept. Head and Senior Scientist

What is the Mission of ESnet?

- Enable thousands of DOE, university and industry scientists and collaborators worldwide to make effective use of unique DOE research facilities and computing resources independent of time and geographic location
  - Direct connections to all major DOE sites
  - Access to the global Internet (managing 150,000 routes at 10 commercial peering points)
- Provide capabilities not available through commercial networks
  - Architected to move huge amounts of data between a small number of sites
  - High bandwidth peering to provide access to US, European, Asia-Pacific, and other research and education networks.

Objective: *Support scientific research* by providing seamless and ubiquitous access to the facilities, data, and colleagues

The STAR Collaboration at the Relativistic Heavy Ion Collider (RHIC)

Brookhaven National Laboratory

Brazil:
Universidade de Sao Paulo

China:
IHEP – Beijing
IMP - Lanzou
IPP – Wuhan
USTC
SINR – Shanghai
Tsinghua University

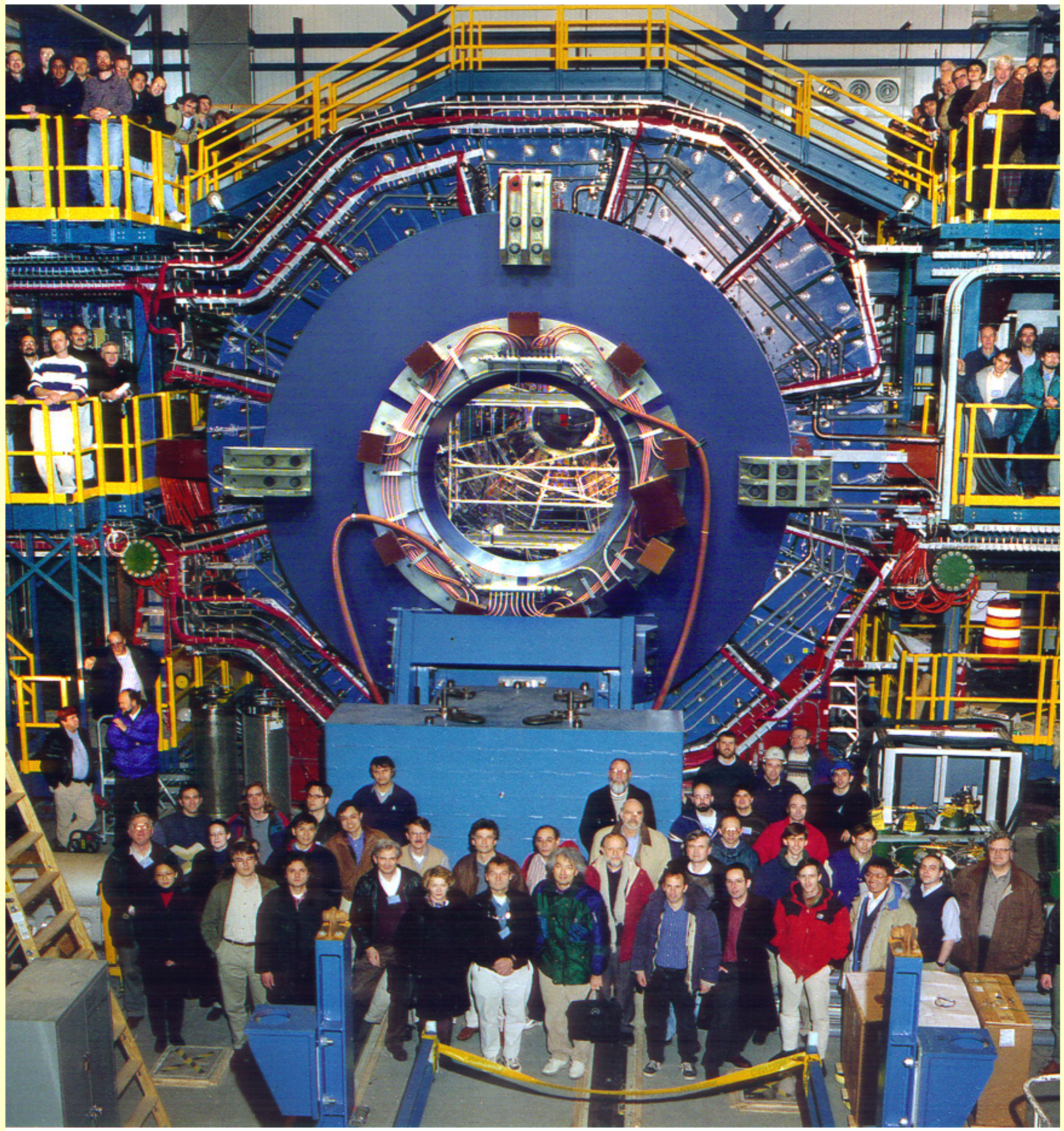
Great Britain:
University of Birmingham

France:
IReS Strasbourg
SUBATECH - Nantes

Germany:
MPI – Munich
University of Frankfurt

India:
IOP - Bhubaneswar
VECC - Calcutta
Panjab University
University of Rajasthan
Jammu University
IIT - Bombay
VECC – Kolcata

Poland:
Warsaw University of Tech

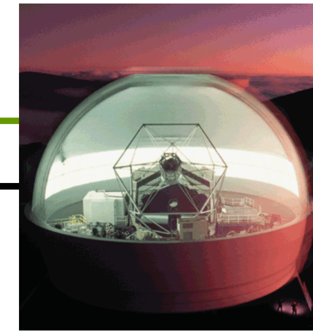


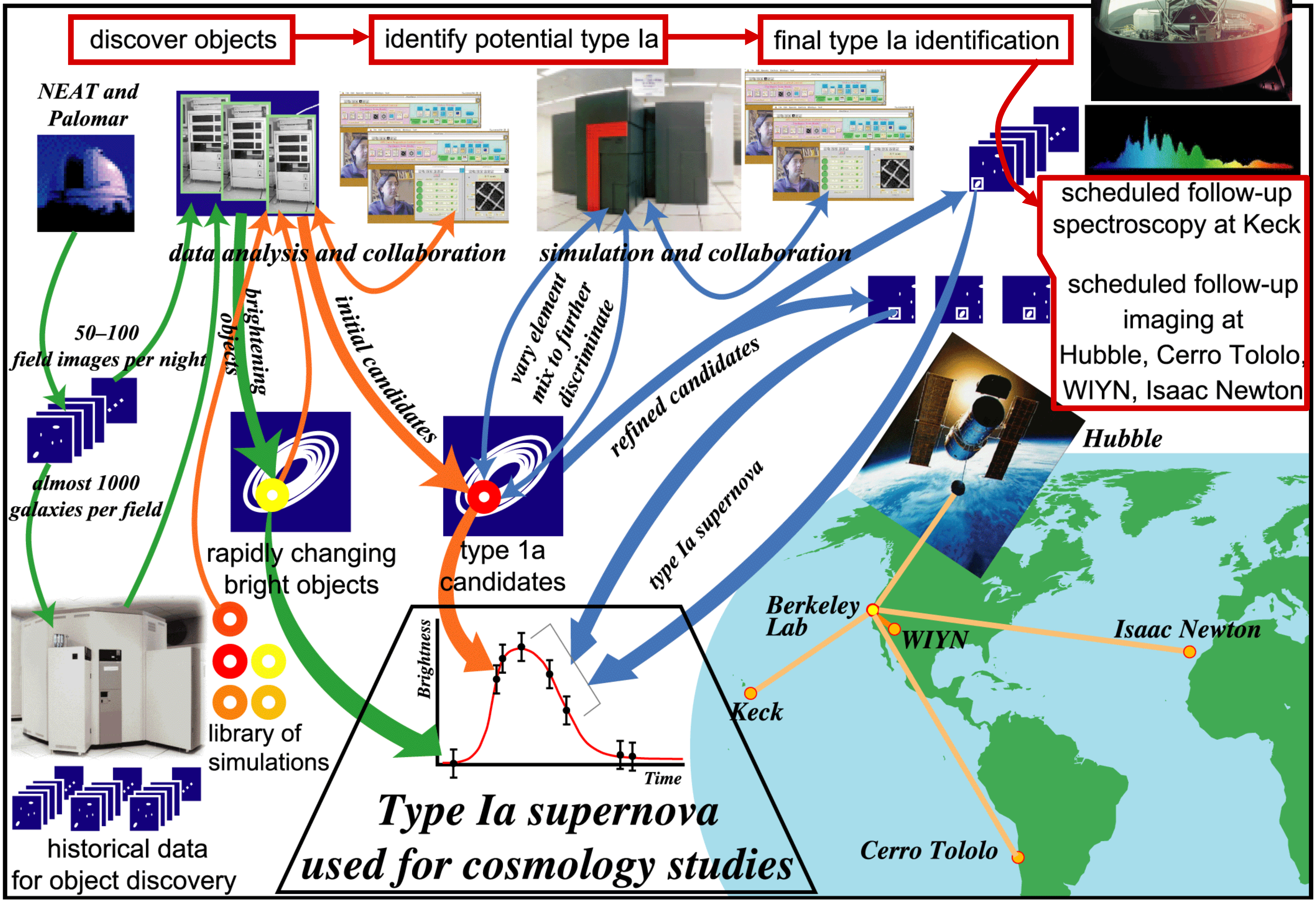
Russia:
MEPHI - Moscow
LPP/LHE JINR - Dubna
IHEP - Protvino

U.S. Laboratories:
Argonne
Berkeley
Brookhaven

U.S. Universities:
UC Berkeley
UC Davis
UC Los Angeles
Carnegie Mellon
Creighton University
Indiana University
Kent State University
Michigan State University
City College of New York
Ohio State University
Penn. State University
Purdue University
Rice University
Texas A&M
UT Austin
U. of Washington
Wayne State University
Yale University

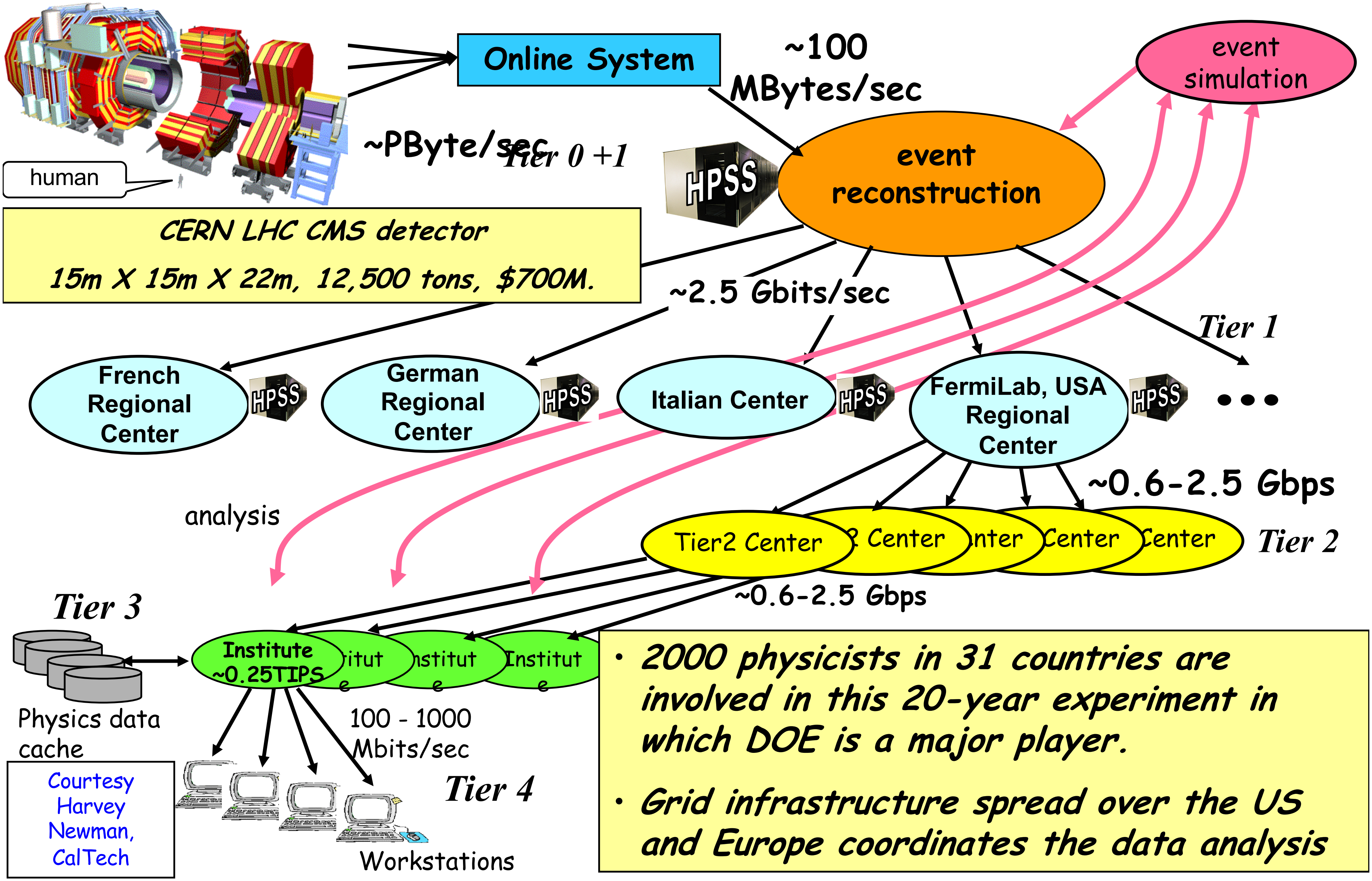
Supernova Cosmology Project





Supernova Cosmology Project, Perlmutter, et al. (http://www.supernova.lbl.gov)

CERN / LHC High Energy Physics Data Provides One of Science's Most Challenging Data Management Problems

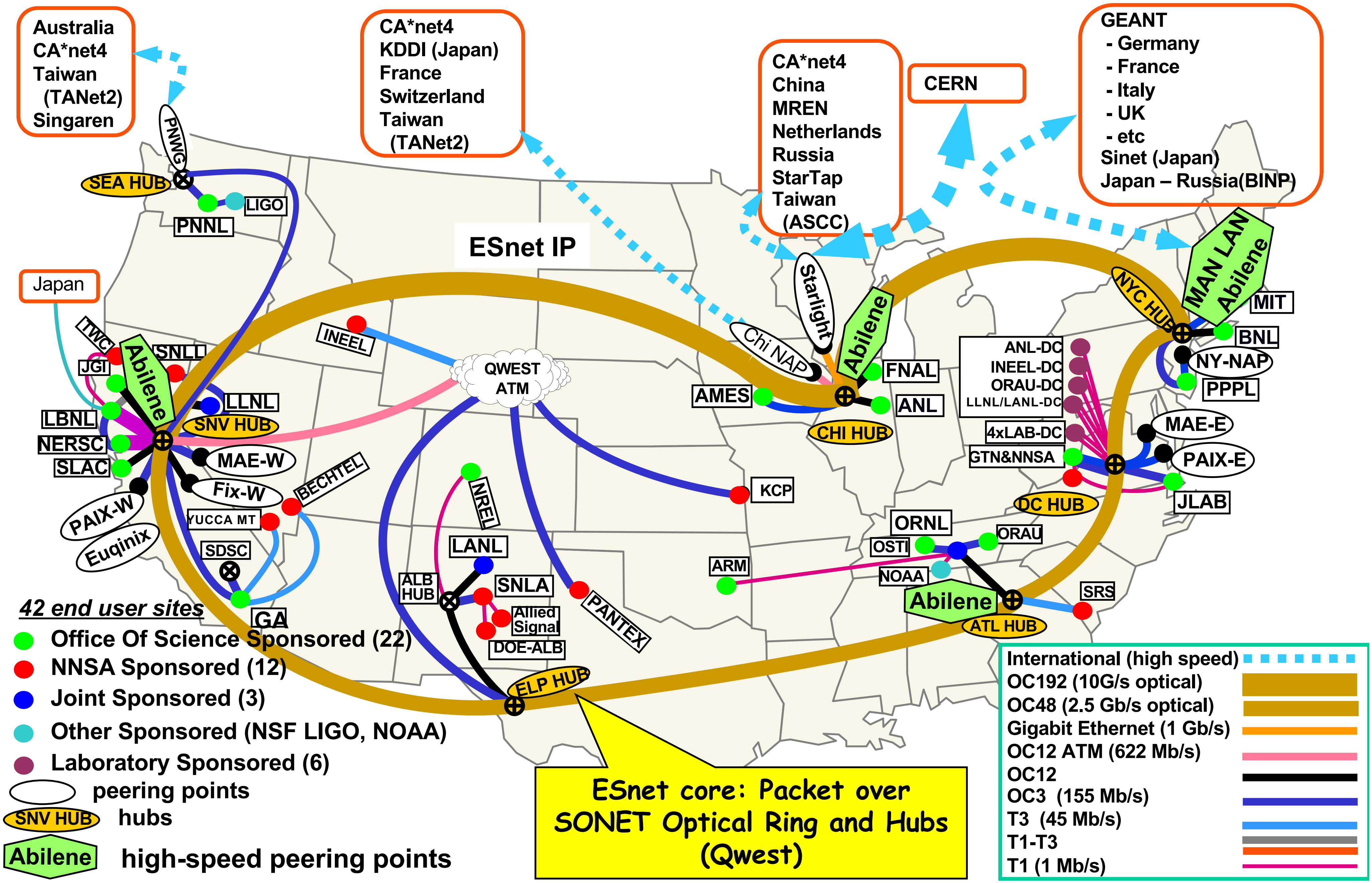


- *2000 physicists in 31 countries are involved in this 20-year experiment in which DOE is a major player.*
- *Grid infrastructure spread over the US and Europe coordinates the data analysis*

How is this Mission Accomplished?

- ESnet builds a comprehensive IP network infrastructure (routing, IPv4, IP multicast, IPv6) based on commercial circuits
  - ESnet purchases telecommunications services ranging from T1 (1 Mb/s) to OC192 SONET (10 Gb/s) and uses these to connect core routers and sites to form the ESnet IP network
  - ESnet peers at high speeds with domestic and International R&E networks.
  - ESnet peers with many commercial networks to provide full Internet connectivity

What is ESnet Today?



Science Drives the Future Direction of ESnet

- Modern, large-scale science is dependent on networks
  - Data sharing
  - Collaboration
  - Distributed data processing
  - Distributed simulation
  - Data management

Predictive Drivers for the Evolution of ESnet

August, 2002 Workshop Organized by Office of Science

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

Workshop Panel Chairs

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

- **The network and middleware requirements to support DOE science were developed by the OSC science community representing major DOE science disciplines**
  - **Climate simulation**
  - **Spallation Neutron Source facility**
  - **Macromolecular Crystallography**
  - **High Energy Physics experiments**
  - **Magnetic Fusion Energy Sciences**
  - **Chemical Sciences**
  - **Bioinformatics**
- **The network is essential for:**
  - **long term (final stage) data analysis**
  - **"control loop" data analysis (influence an experiment in progress)**
  - **distributed, multidisciplinary simulation**

Available at www.es.net/#research

The Analysis was Driven by the ***Evolving Process of Science***

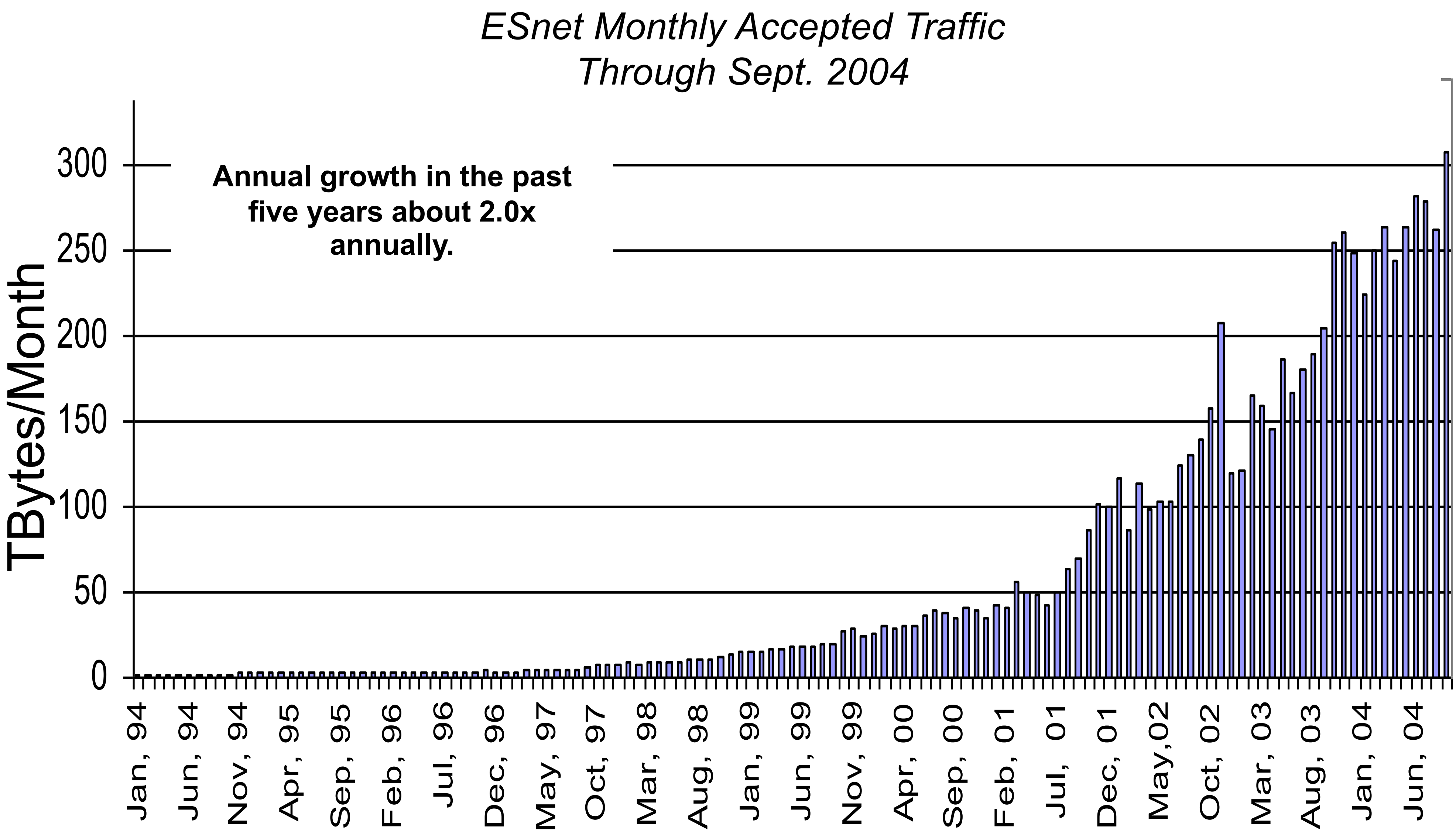
Feature / Discipline	Vision for the Future Process of Science (analysis was driven by)	Characteristics that Motivate High Speed Nets	Requirements: Networking	Requirements: Middleware
Climate (near term)	Analysis of model data by selected communities that have high speed networking (e.g. NCAR and NERSC)	• A few data repositories, many distributed computing sites • 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 Distribute large chunks of data to major users for post-simulation analysis	• 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 done with distributed, multidisciplinary simulation • Virtualized data to reduce storage load	• 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

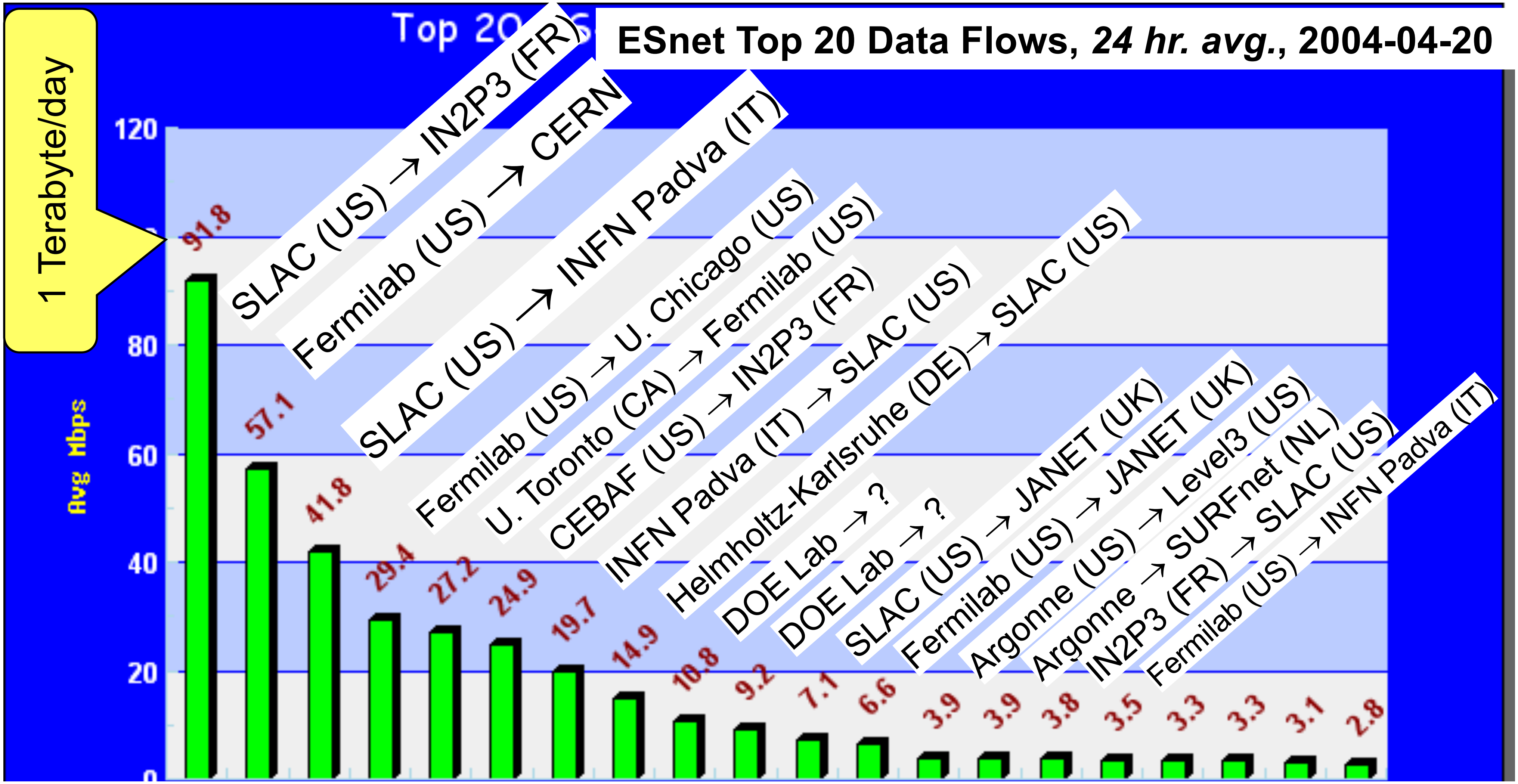
Science Areas	**Today *End2End* Throughput**	**5 years End2End Throughput**	**5-10 Years End2End Throughput**	**Remarks**
High Energy Physics	0.5 Gb/s	100 Gb/s	**1000 Gb/s**	high bulk throughput
Climate (Data & Computation)	0.5 Gb/s	160-200 Gb/s	**N x 1000 Gb/s**	high bulk throughput
SNS NanoScience	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

Observed Data Movement in ESnet

ESnet is currently transporting about 300 terabytes/mo.
(300,000,000 Mbytes/mo.)



The Science Traffic on ESnet is Increasing Dramatically: A Small Number of Science Users at Major Facilities Account for a Significant Fraction of all ESnet Traffic



- Since BaBar production started, the top 20 ESnet flows have consistently accounted for > 50% of ESnet's monthly total traffic (~130 of 250 TBy/mo)
- As LHC data starts to move, this will increase a lot (200-2000 times)

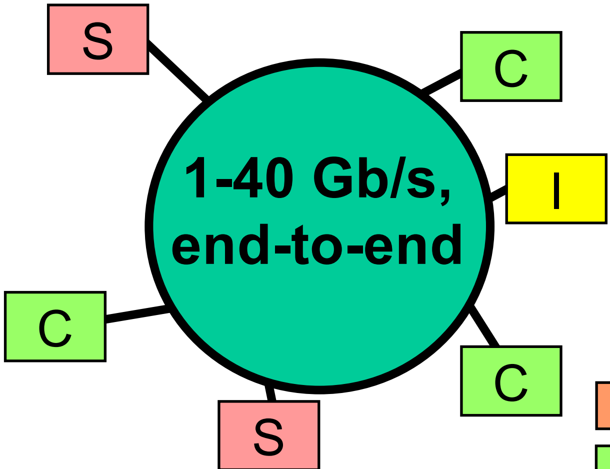
Enabling the Future: ESnet's Evolution Near Term and Beyond

- Based on the requirements of the OSC Network Workshops, ESnet networking must address
  - A reliable production IP Network
    - University and international collaborator connectivity
    - Scalable, reliable, and high bandwidth site connectivity
  - A network support of high-impact science (Science Data Network)
    - provisioned circuits with guaranteed quality of service (e.g. dedicated bandwidth)
  - Additionally be positioned to incorporate lessons learned from DOE's UltraScience Network, the Advanced Research Network.

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"

Evolving Requirements for DOE Science Network Infrastructure

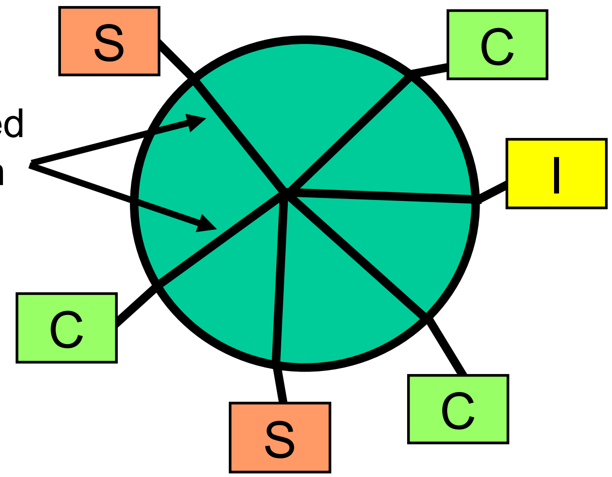
1-3 yr Requirements



S storage
C compute
I instrument
C&C cache & compute

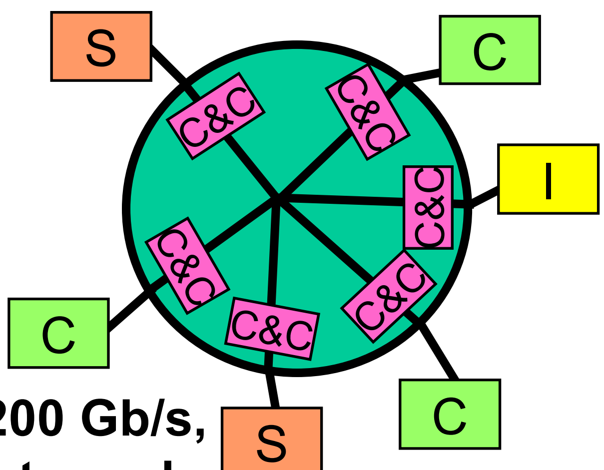
- In the near term applications need higher bandwidth

2-4 yr Requirements



- high bandwidth
- QoS

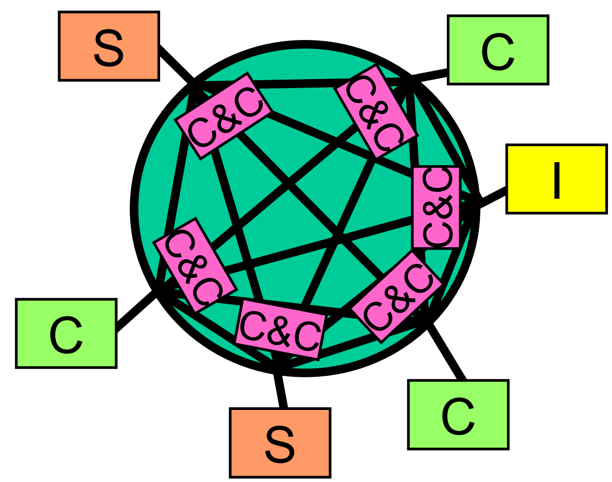
3-5 yr Requirements



100-200 Gb/s,
end-to-end

- high bandwidth and QoS
- network resident cache and compute elements

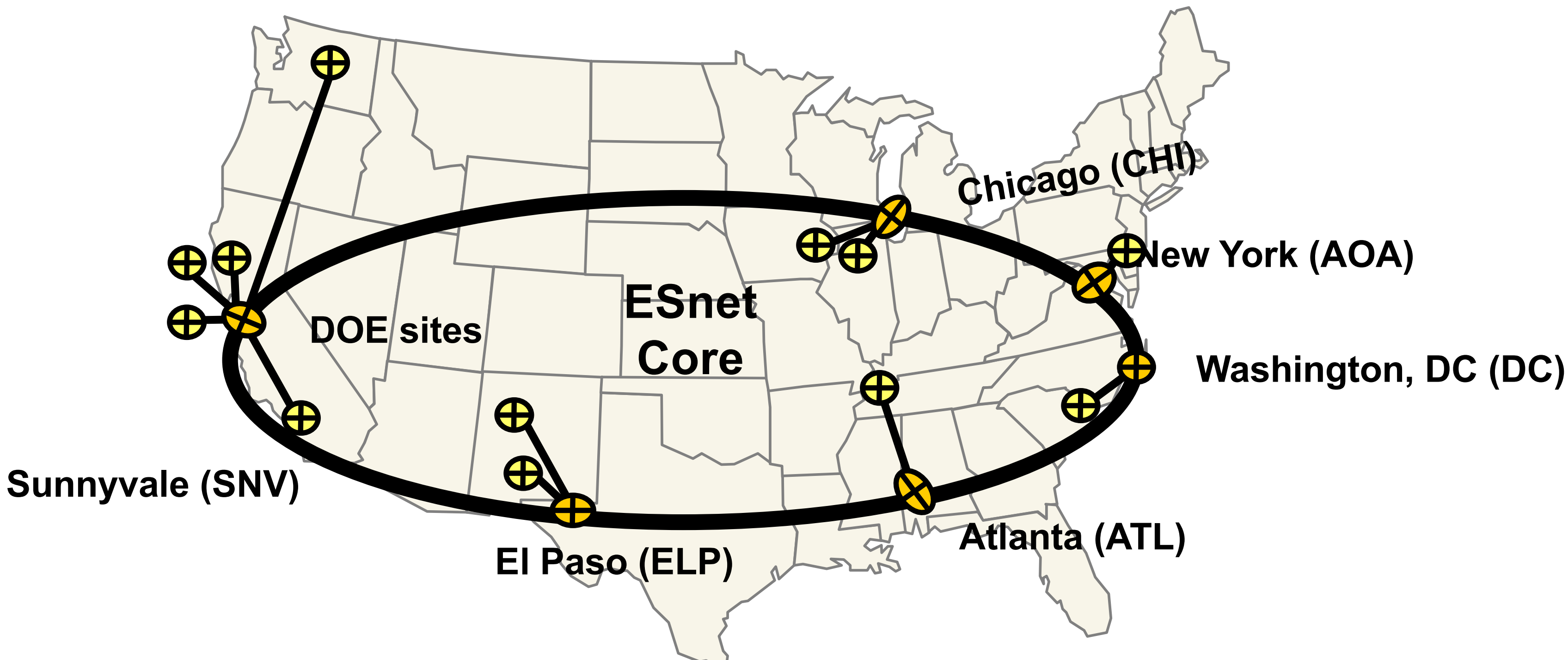
4-7 yr Requirements



- high bandwidth and QoS
- network resident cache and compute elements
- robust bandwidth (multiple paths)

New ESnet Architecture Needs to Accommodate OSC

- The essential requirements cannot be met with the current, telecom provided, hub and spoke architecture of ESnet



- The core ring has good capacity and resiliency against single point failures, but the point-to-point tail circuits are neither reliable nor scalable to the required bandwidth

Basis for a New Architecture

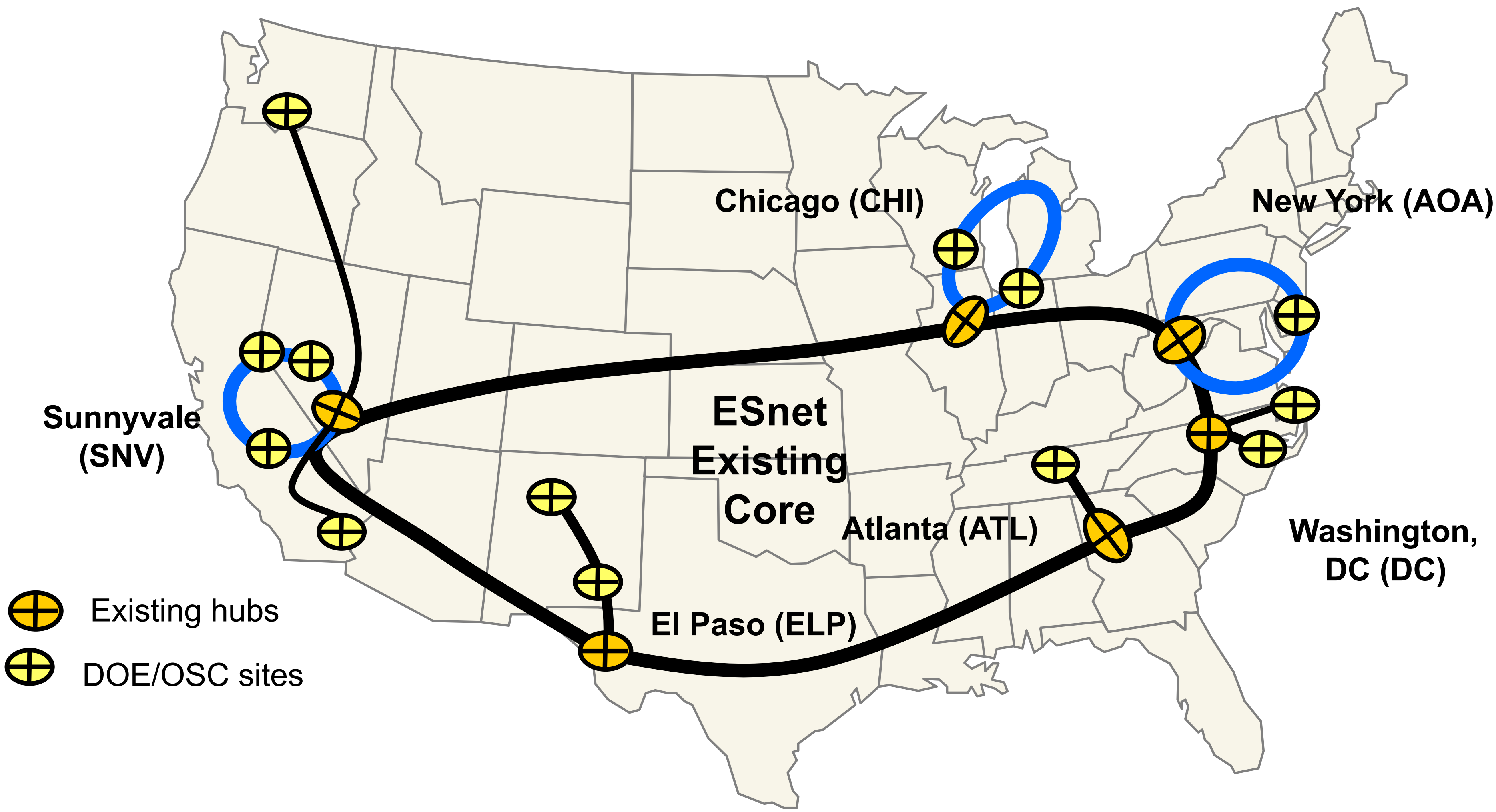
Goals for each site:

- fully redundant connectivity
- high-speed access to the backbone

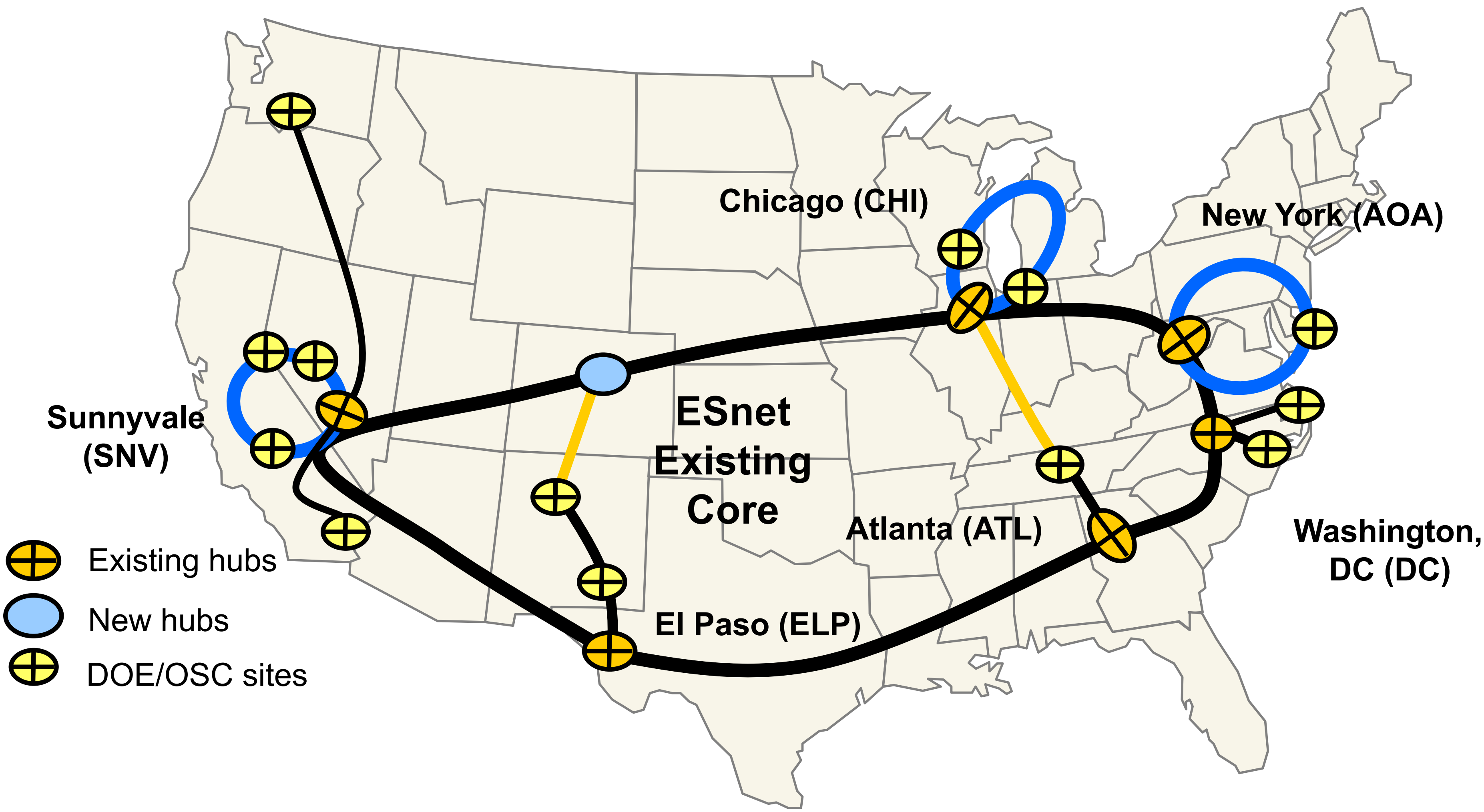
Meeting these goals requires a two part approach:

- Connecting to sites via a MAN ring topology to provide:
  - Dual site connectivity and much higher site bandwidth
- Employing a second backbone to provide:
  - Multiply connected MAN ring protection against hub failure
  - Extra backbone capacity
  - A platform for provisioned, guaranteed bandwidth circuits
  - An alternate path for production IP traffic
  - Access to carrier neutral hubs

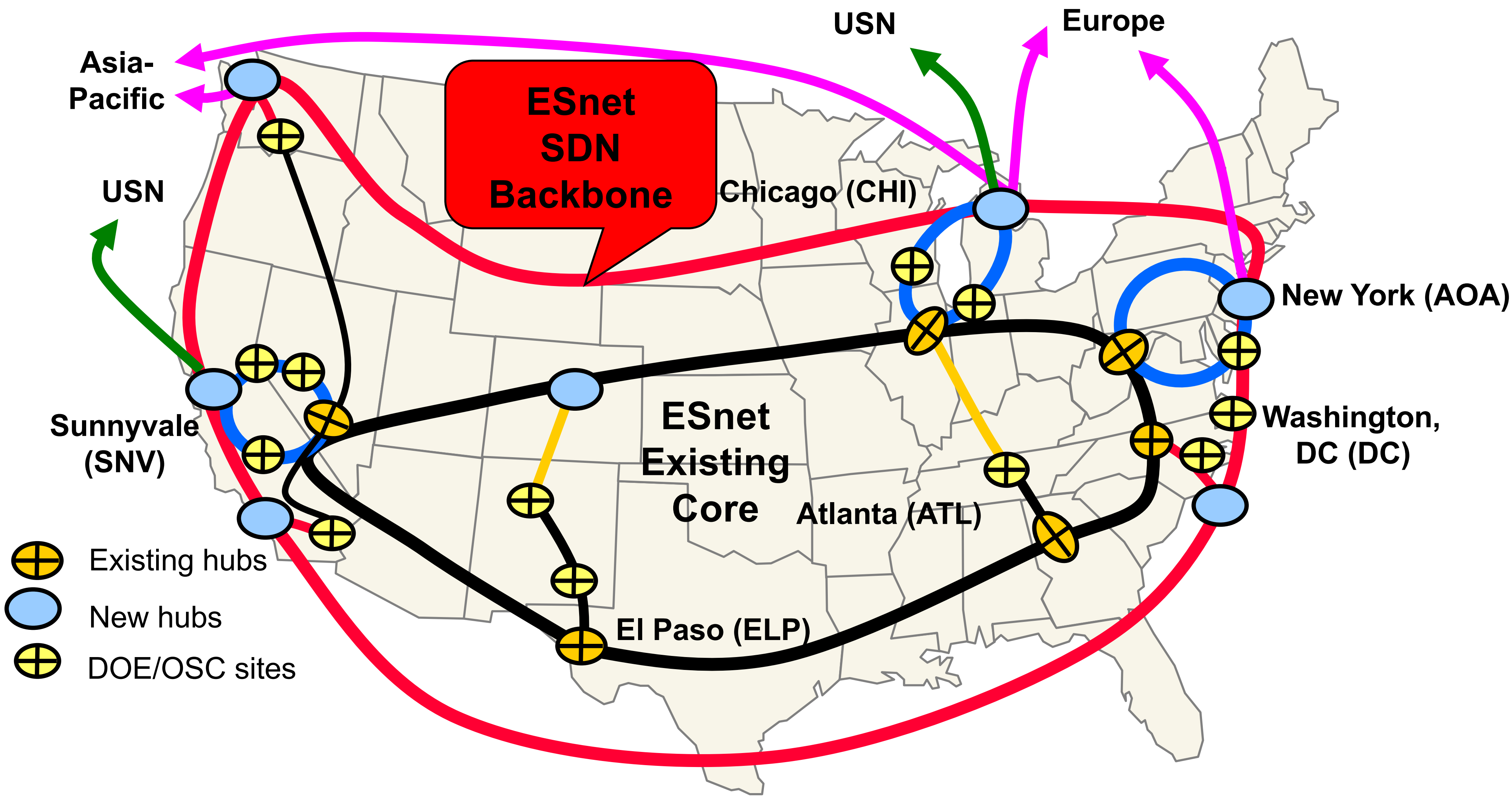
Bay Area, Chicago, New York MANs



Enhancing the Existing Core

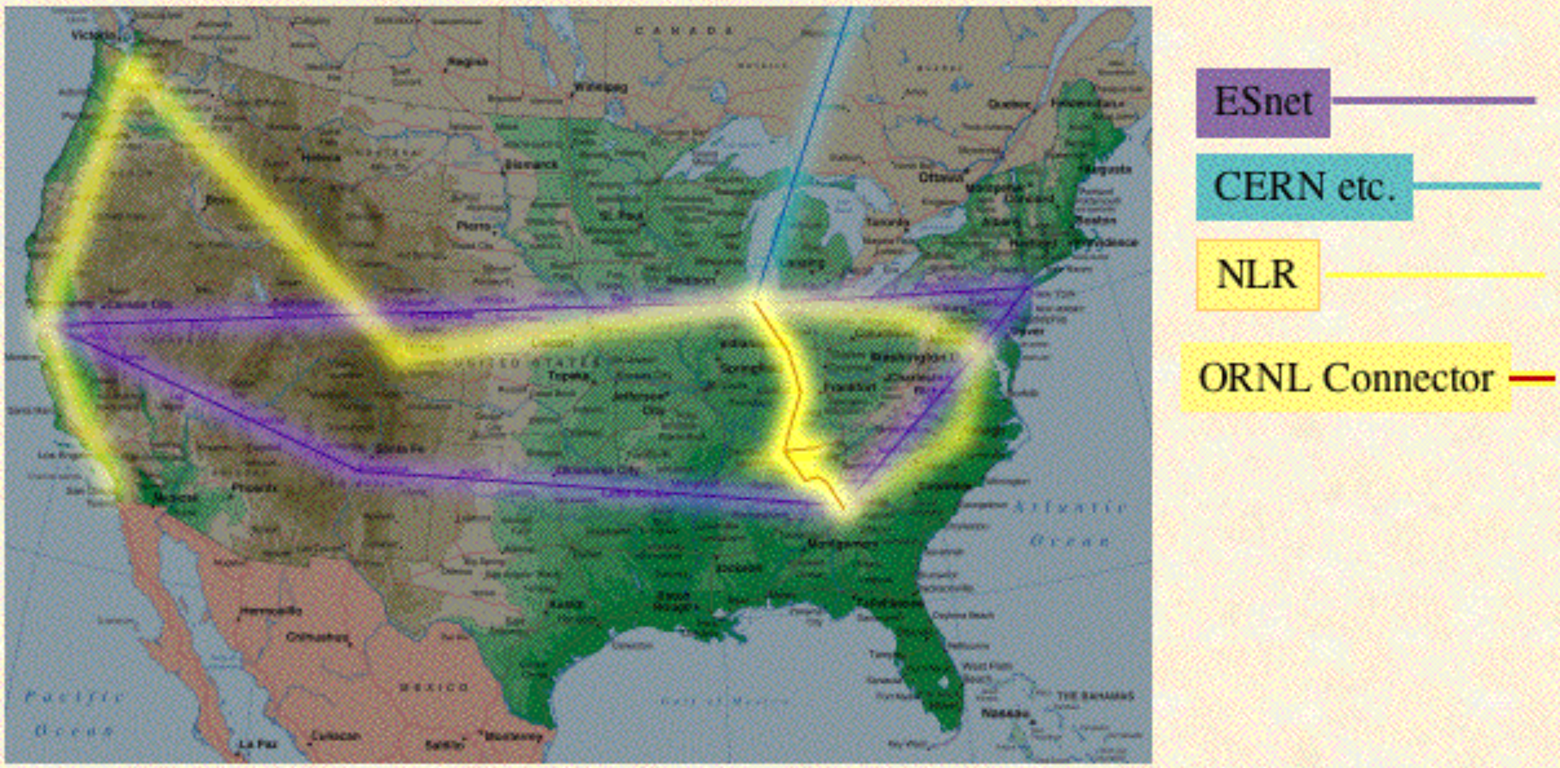


Addition of Science Data Network (SDN) Backbone



ESnet – UltraScience Net Topology

The Physical View From 50,000 feet



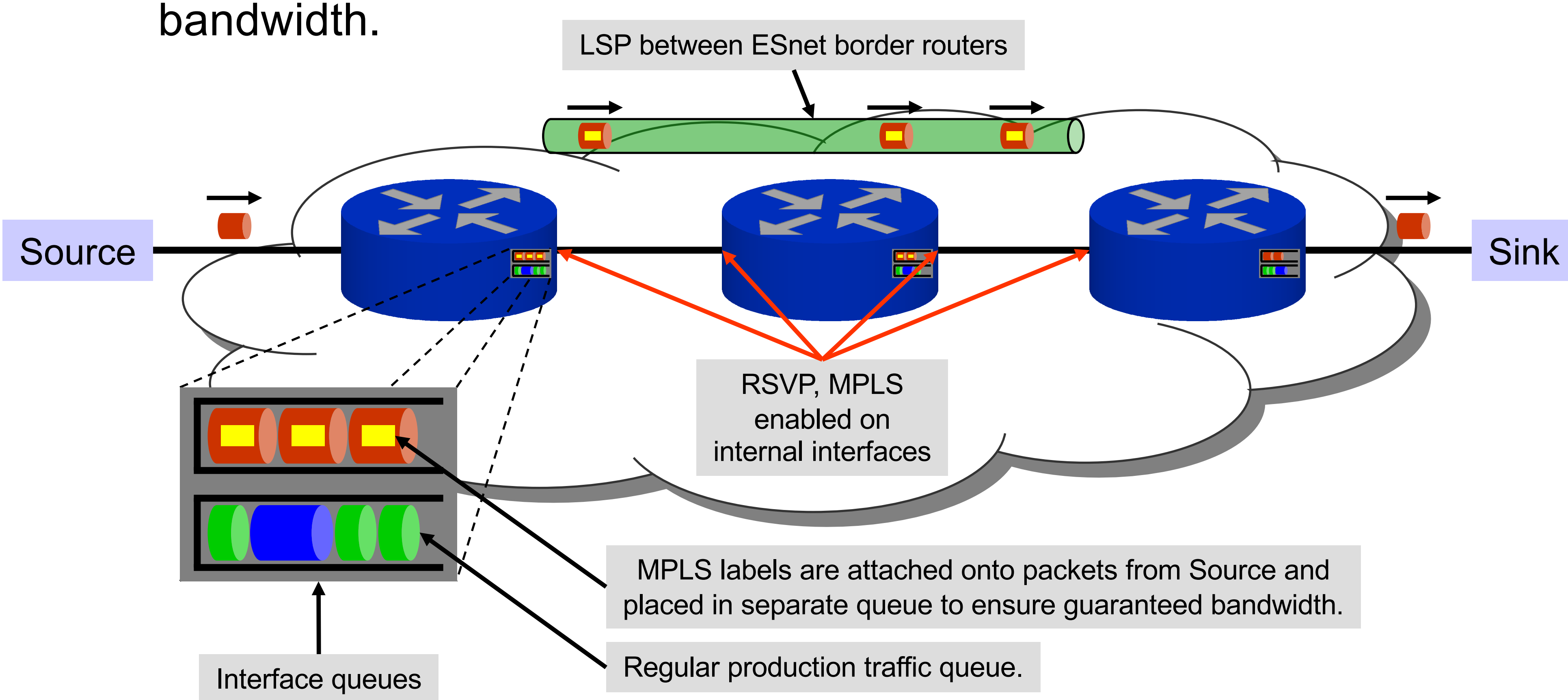
OAK RIDGE NATIONAL LABORATORY
U. S. DEPARTMENT OF ENERGY

UT-BATTELLE

On-Demand Secure Circuits and Advance Reservation System (OSCARS)

Guaranteed Bandwidth Circuit

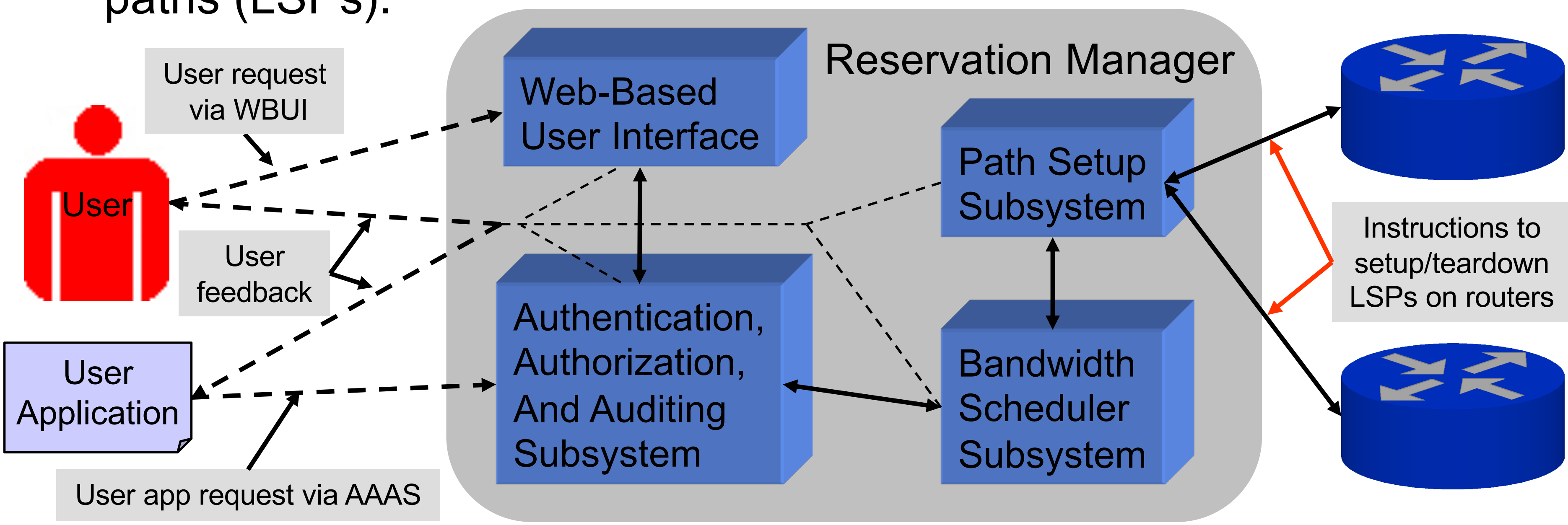
- Multi-Protocol Label Switching (MPLS) and Resource Reservation Protocol (RSVP) is used to create a Label Switched Path (LSP)
- Quality-of-Service (QoS) level is assign to the LSP to guarantee bandwidth.



On-Demand Secure Circuits and Advance Reservation System (OSCARS)

Reservation Manager

- Web-Based User Interface (WBUI) will prompt the user for a username/password and forward it to the AAAS.
- Authentication, Authorization, and Auditing Subsystem (AAAS) will handle access, enforce policy, and generate usage records.
- Bandwidth Scheduler Subsystem (BSS) will track reservations and map the state of the network (present and future).
- Path Setup Subsystem (PSS) will setup and teardown the on-demand paths (LSPs).



ESnet Meeting Science Requirements – 2007/2008

- 10 Gbps enterprise IP traffic
- 40-60 Gbps circuit based transport