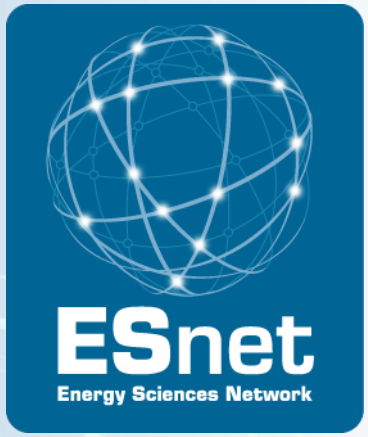


Research Networks: The "Other" Internet

R. Kevin Oberman
Senior Network Engineer

NANOG 52, June 13, 2011
Denver, Colorado

Overview

- What is a Research Network?
- A Brief History of Research Networks
- Current Research Networks (an incomplete list)
- Why Research Networks are “Another” Internet
- What Research Networks Will Be Doing in the Future

Definitions – What is a Research Network?

- Research (or Research and Educational) Networks are non-commercial networks created for the advancement of knowledge
  - Non-commercial
- Generally used to connect between research and education facilities
  - Avoids use of commercial service providers
  - Have adequate capacity to avoid congestion
  - Generally do NOT have “end-users”
    - End-users are regional networks and research facilities

History of Research networks

Before the mid-1990s, almost all networks on the Internet were "Research" networks of some sort

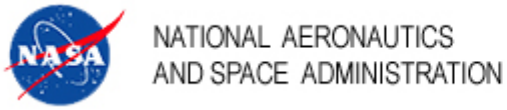
- Even commercial entities on the Internet were there primarily for research purposes
- Initially for research into networking (ARPANET)
- Later provided support for general scientific and engineering research (NASA Science Internet, NSFnet, ESnet (nee MFEnet)
- Multiprotocol nets were the norm in the 1980s
- By the mid-1990s IP was dominant
  - Primarily due to the work of MILNET and NSFnet

Commercialization of the Internet

In the mid-1990s the commercialization of the Internet, the "Information Superhighway", and the web changed the nature of the Internet

- Commercial Internet faced massive growth and many "growing pains"
  - Serious congestion became common
    - Downers Grove, Illinois
    - East and West Orange, New Jersey
  - Architectural issues
    - Hinsdale, Illinois fire
- University researchers became discouraged by poor performance and reliability
- Research networks started to develop a separate character

North American Research Networks



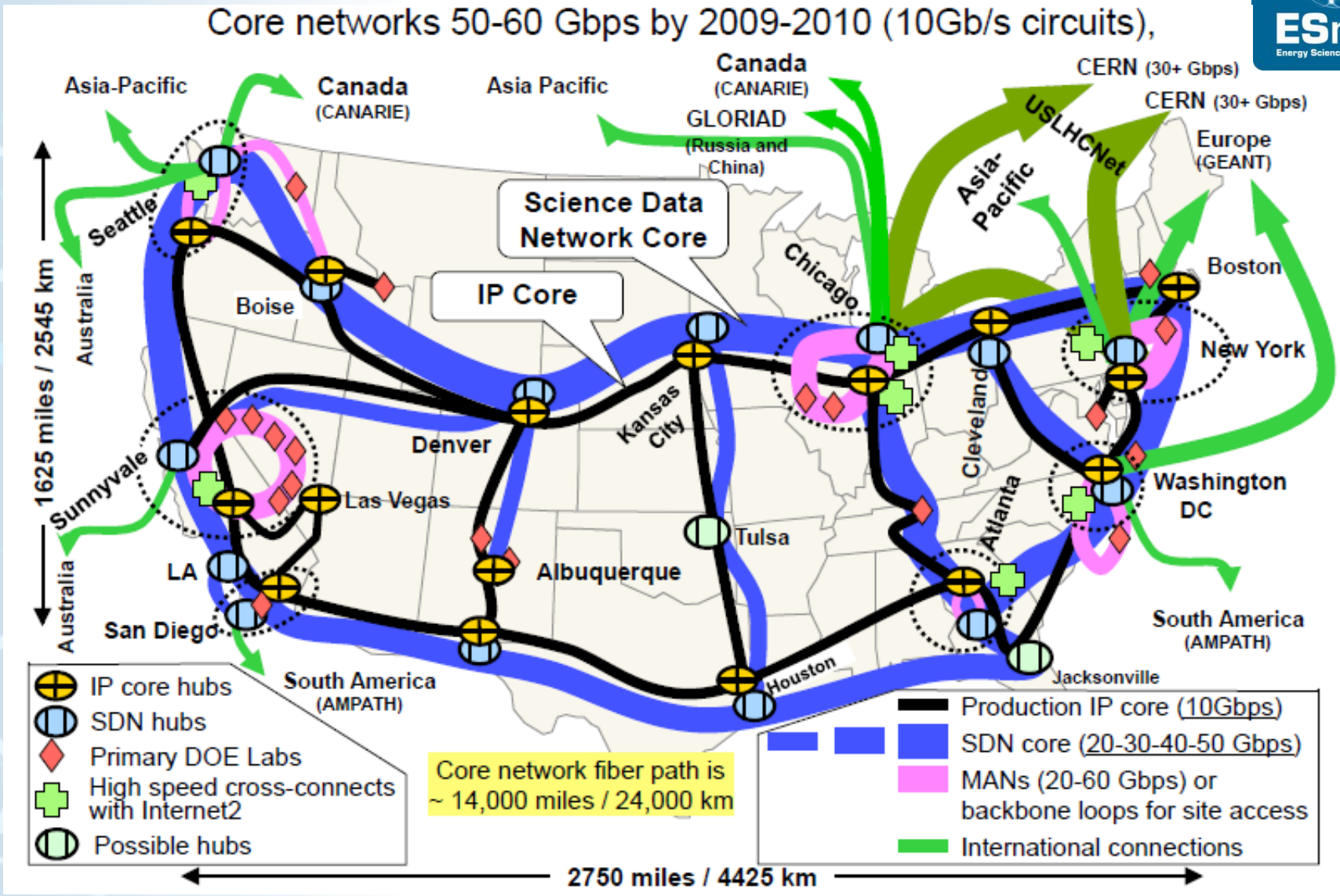
DREN

ESnet – 35 years of support for research

ESnet may be the oldest continuously operating national research network

- Started in 1976 as the MFEnet to support global fusion research
- Ran over 56K satellite links to US, Asian and European researchers
- Used its own protocols based on an early DEC network design document (referred to internally as “DECnet Phase 0”)
- In 1986 MFEnet merged with the DECnet based HEPnet and was renamed ESnet – The Energy Sciences NETwork and started transitioning to IP
  - Supports USDOE Office of Science funded facilities
    - The USDOE Office of Science is the largest funder of basic scientific research in the US
  - Also serves NNSA and other federal locations

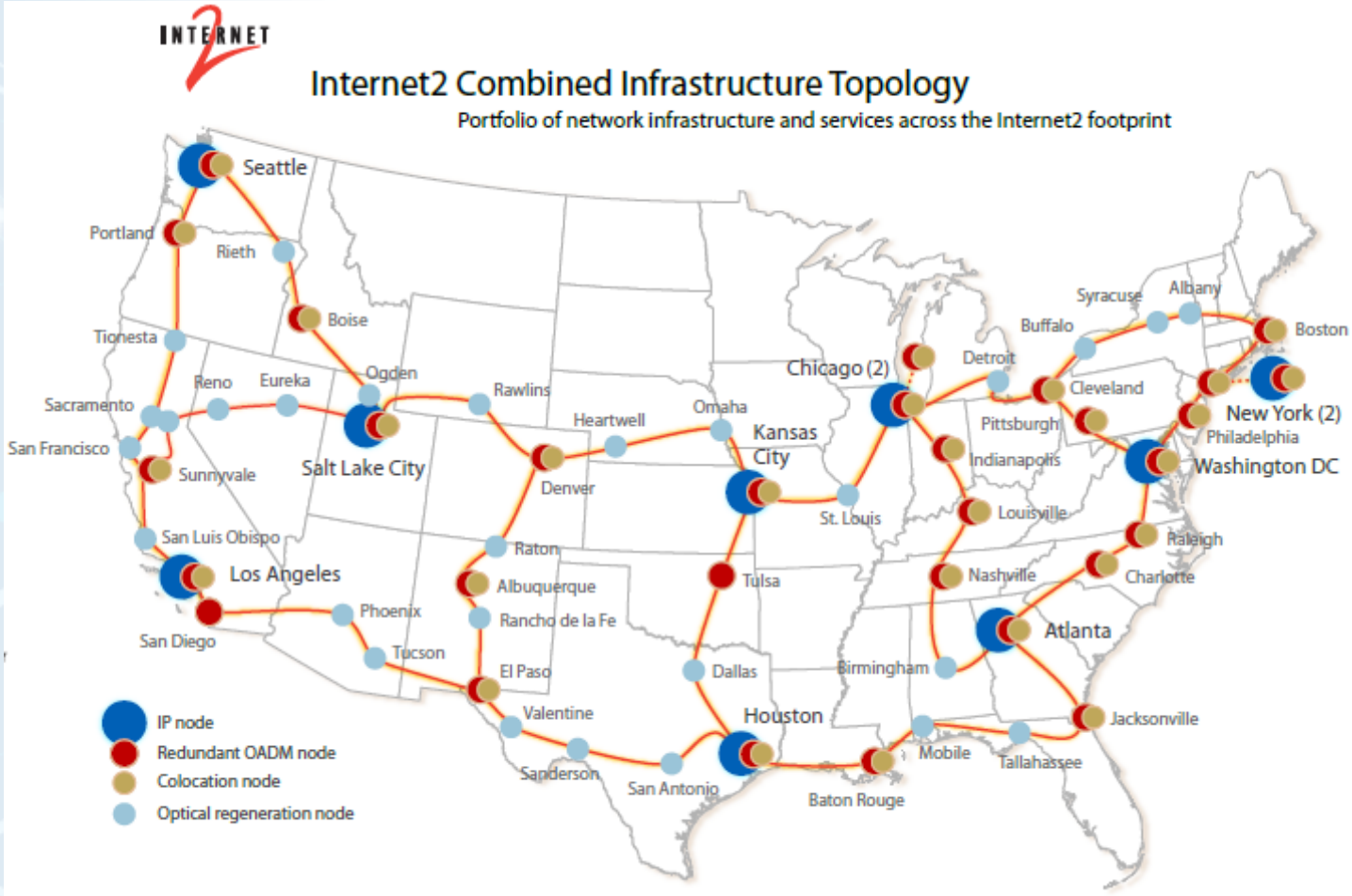
ESnet4 Backbone



Internet2 – Educational institutions do it themselves

- As the commercial Internet grew, major universities started seeing congestion and bandwidth constraints that impacted research
- In 1996 34 universities formed Internet2 to build a national backbone to connect those universities over state-of-the-art, OC-48 circuits
  - Built in partnership with ESnet (Common layer 1)
  - Has grown to over 200 member institutions over multiple 10G circuits
  - National backbone provides very high reliability and immunity to circuit failure
  - Primarily connect to regional educational consortia
    - Referred to as “GigaPOPs”

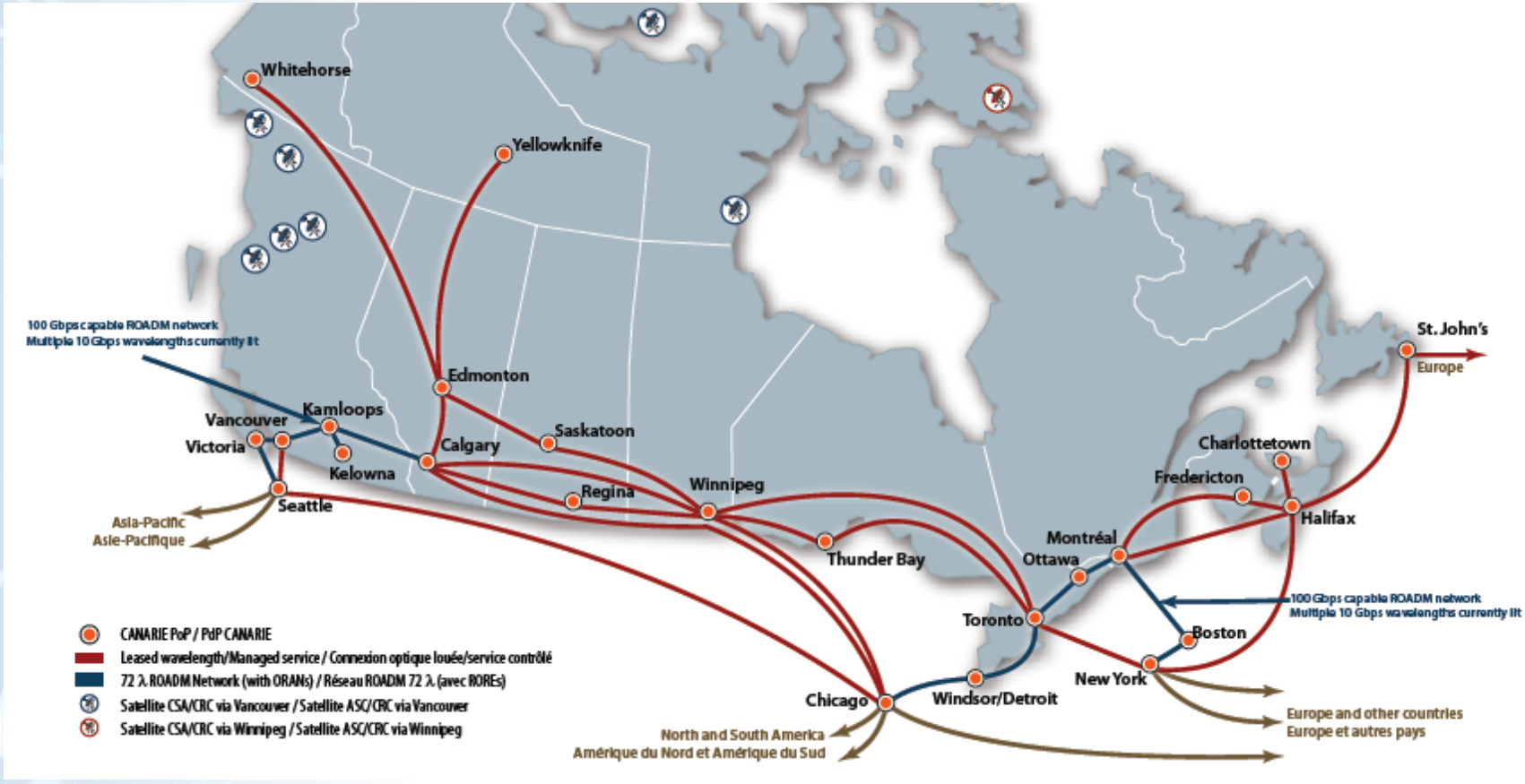
Internet2 Network Map



Canarie – Canadian research network

- Created in 1990 with low-speed (9.6K) links
- Upgraded to 56K in 1993
- Upgraded to 10 Mbps in 1995
- Now runs over WDM OC-192 circuits
- Over 19,000 km
- Nearly 200 colleges and universities
- 86 government laboratories
- Averages over 7 petabytes per quarter

CANARIE Network Map



National Lambda Rail (NLR)

- Founded in 2003 by a consortium of universities and Internet2
- Initial goal was a national footprint of 10G waves for use by researchers
- Footprint is over 12,000 miles of fiber
- Provides connectivity at layer 1 (Wavenet), Layer 2 (Framenet), and Layer 3 (Packetnet)
- 13 member organizations
- Connects over 200 colleges and universities through Regional Optical Networks (RONs)

National LambdaRail Network Map

Other research networks

- DREN – the Defense Research and Engineering Network
- N-WAVE – NOAA Research Network
- NREN – NASA Research and Engineering Network
- All provide networks for research of the sponsoring organization
- All are largely opaque as they are restricted to constituent use
  - Basically enterprise networks supporting research

Why Research networks are different

- Move very large scientific data sets (terabytes)
  - Large-scale science flows on ESnet typically account for half of all traffic volume
  - “Big science” is the largest driver
    - LHC
    - Supercomputer simulations (the “third leg of science”)
    - Astrophysics and astronomy
  - Some (e.g. large supercomputer centers) are likely to require terabit speed in a decade
- May require interactivity with significant data transfer requirements
  - Remote control of experiments (Real time)
  - Data analysis and tuning between scheduled runs

Example of “Big Science” – LHC

The Large Hadron Collider at CERN in Europe may be the largest single scientific experimentation facility ever built

- Runs multiple, independent experiments, each generating terabytes
- Multi-gigabit per second data flows may run for many hours
  - Dropped packets are devastating to performance at these speeds and latencies
  - Tuning of the full network and systems critical
  - Specialized cache and storage systems had to be designed
  - On-demand circuits assure that congestion will not impact transfers

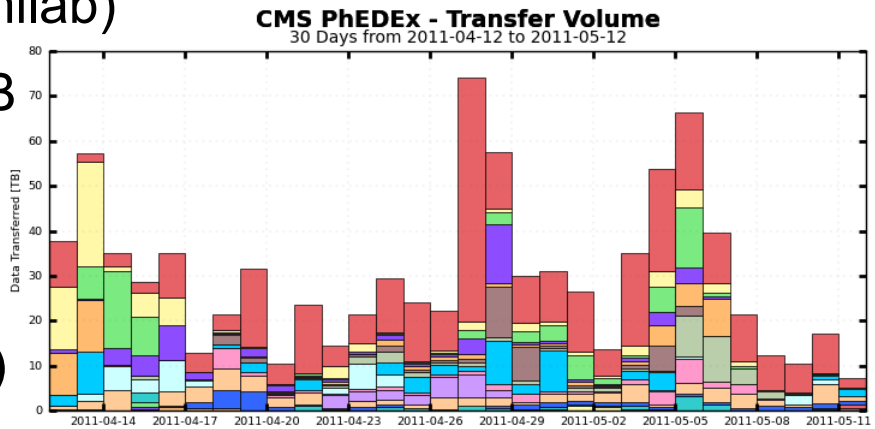
Under the Atlantic to FNAL, BNL & TRIUMF

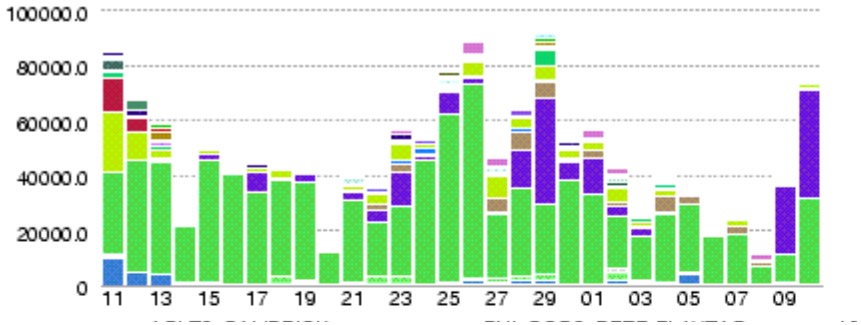
- Data is distributed to major research facilities around the world
  - Data sources from CERN (Tier 0)
  - Fermilab and Brookhaven in the US (Tier 1)
  - TRIUMF in Canada (Tier 1)
  - Stored and distributed on-demand to universities and other laboratories (Tier 2)
  - Moved between Tier 1 and Tier 2 sites
  - Processed data returned to Tier 1 sites
  - Not all data is available from every Tier 1
    - Results in additional trans-oceanic traffic

How much data are we talking about?

CMS and ATLAS are longest running, best understood experiments

- CMS annual traffic to the US (Fermilab)
  - Maximum daily data – 74.04 TB
  - Minimum daily data – 7.24 TB
  - Average daily data – 30.05 TB
- ATLAS 30 day traffic (Brookhaven)
  - Maximum daily data – 93 TB
  - Minimum daily data – 13 TB
  - Average daily data – 49 TB





Remote control

- Generally requires minimal delay
- Would typically use Expedited Forwarding (or Better than Best Effort)
  - Traffic volume would need to be constrained to maximize performance
  - Would be best on a low bandwidth reserved circuit
  - Latency must be considered in system design
- Example:
  - Health checks for automated protein crystallography
  - Control of fusion experiments from General Atomics in San Diego to the Princeton Plasma Physics Lab

Circuit reservation system (OSCARS)

On-demand Secure Circuits and Advance Reservation System

- Developed by ESnet, Internet2, and ISI East
- Web based interface
- L2VPNs or L3VPNs
- Uses RSVP-TE and OSPF-TE
  - Allows the selection of the fastest path with available bandwidth over the entire path
  - Reservations are time based to be built and torn down automatically

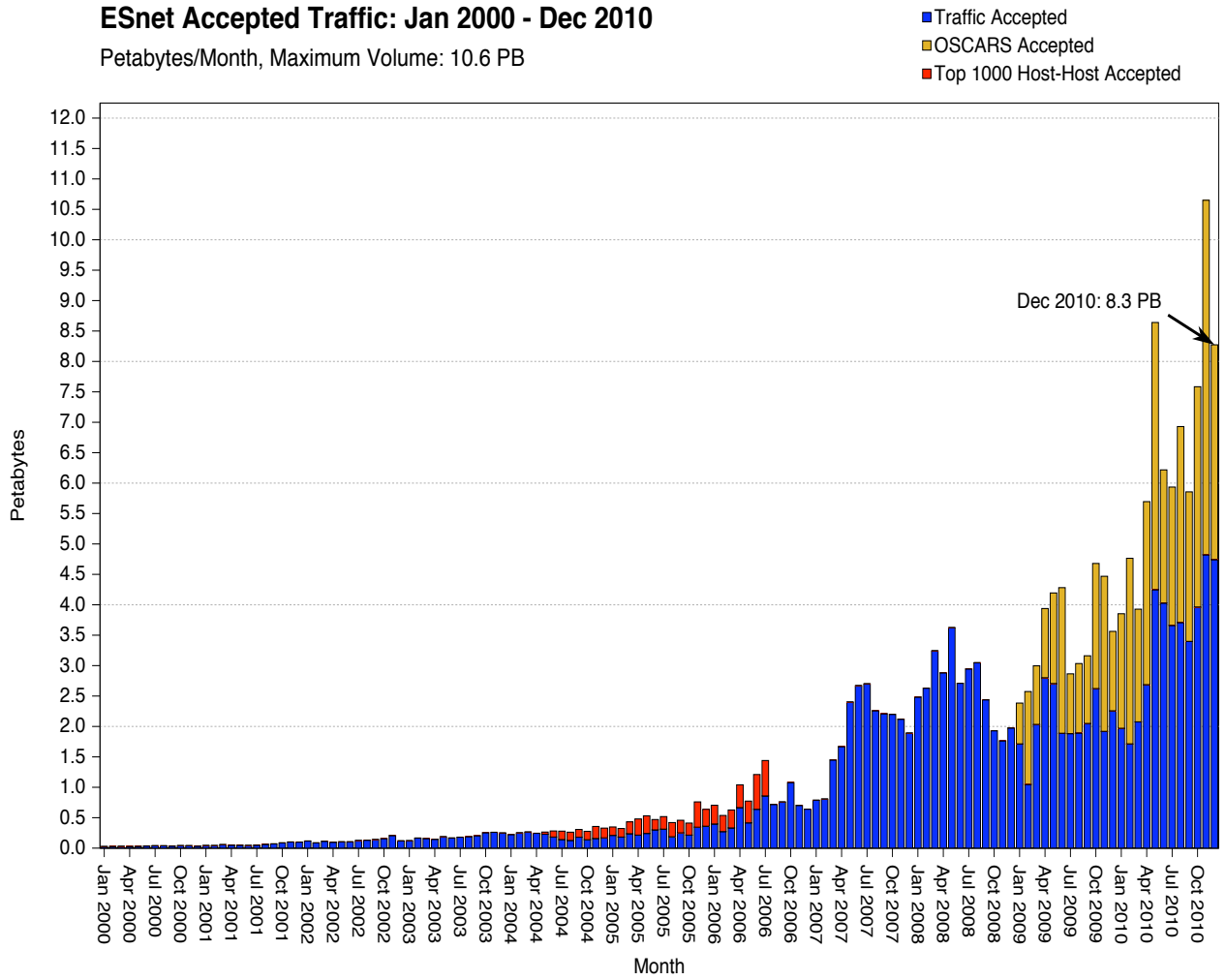
ESnet4 Network

Peaked at 10.6 PB in Nov

More than 50% go over OSCARS circuits

Instrumenting with perfSONAR all connections 1G and up

99.9983% network availability to Office of Science labs



Reliability – 5-9s (or 4-9s and an 8) and near zero drops!

- CERN has a limited ability to store data during network failures
- Rapid troubleshooting is critical
- perfSONAR (perfSONAR.net) is a key tool
  - Servers at all backbone sites and connected facilities
  - Run regular tests of network performance
  - Allow more intrusive tests (up to 10G) for pin-pointing problems
  - Have proven very effective in locating problems
  - Use standard protocols and can test across networks
  - Widely deployed on ESnet, Internet2, and in Europe
  - When you need large buffers to provide elasticity for fan-in, it's not "buffer bloat!

The future of research networks

- More bandwidth
  - Most are moving to 100Gbps connectivity
- More efficient use a spectrum
- More agility to better utilize available bandwidth
  - Enhancements to middleware to improve host-to-host performance
  - Improvements to OSCARS to better schedule bandwidth
    - Includes automatic L3VPN creation
  - Tune end systems
    - Fasterdata (fasterdata.es.net) provides a lot of data on system tuning

Advanced Networking Initiative
100Gbps Waves Across the US

Summary

- Research and commercial network requirements have diverged
- Still have a great deal in common
- Techniques and tools developed by the research community may prove useful to commercial providers
  - More awareness of what the research networks are doing could benefit commercial providers