Research and Commodity Networks: What are they and what's the difference?

Michael Sinatra, Network Engineer

with additional content from Kevin Oberman (retired Network Engineer) and Greg Bell (Director, ESnet)

NANOG On the Road, May 20, 2014
Los Angeles, California

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
  - Have adequate capacity to avoid congestion (this means a lot of headroom).
  - Generally do not have "end-users"
- We'll talk about differences between research and commodity networks a bit later…

Definitions – What is a Commodity Network?

- Commodity or commercial networks provide general internet service to individuals and companies.
  - Purpose is to provide access to Internet resources and/or content.
  - Not geared toward those interested in researching the network itself—the network is just a substrate for content.
- Emerged from the original (R&E) Internet in the late-1980s and early-1990s.
- Differentiated over the years:
  - Backbone networks
  - CDNs
  - (Broadband) access networks.

More definitions!

- Settlement-free peering: An agreement to interconnect networks directly in order to exchange *customer routes only.* Peering is done at (fixed) cost—cost of the router/switch interfaces, cross-connects, etc., but no charge for the traffic exchanged.
- Paid peering: Same as above, with an extra charge by one of the peers to the other for the “privilege” of peering. (E.g. Netflix and Comcast.)
- Transit: A connection in which one network agrees to provide connectivity beyond just the transit network’s own customers.

Still more definitions…

- Tier 1 ISP: An ISP (or NSP) which only peers and does not buy transit. In order to have a fully-connected Internet, all Tier 1s must peer with each other. I sometimes describe the Tier 1 Internet as being “butt-jointed.”
- Tier 2 ISP: An ISP that peers and purchases some transit (usually from Tier 1s).
- Tier 3 ISP: An ISP that only purchases transit (not necessarily only from Tier 2s).
- Back when the Internet WAS a research network, all networks were essentially Tier 1s.
- ESnet was a Tier 1 research network until 2011.

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, 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

History of Research networks

Early national research networks

- ARPANET
  - 1969: one of the first packet-switched networks (but not "the Internet" at this point).
  - 1983: ARPANET becomes the first network to adopt TCP/IP; this eventually allows it to interconnect with NSFnet, when it is formed in 1985.
- ESnet
  - 1976: MFEnet, connecting various US DOE fusion facilities.
  - 1980: HEPnet, connecting high-energy physics facilities.
  - 1986: MFEnet and HEPnet merge to create ESnet.

ESnet – 38 years of support for research (courtesy Kevin Oberman)

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

History of Research networks

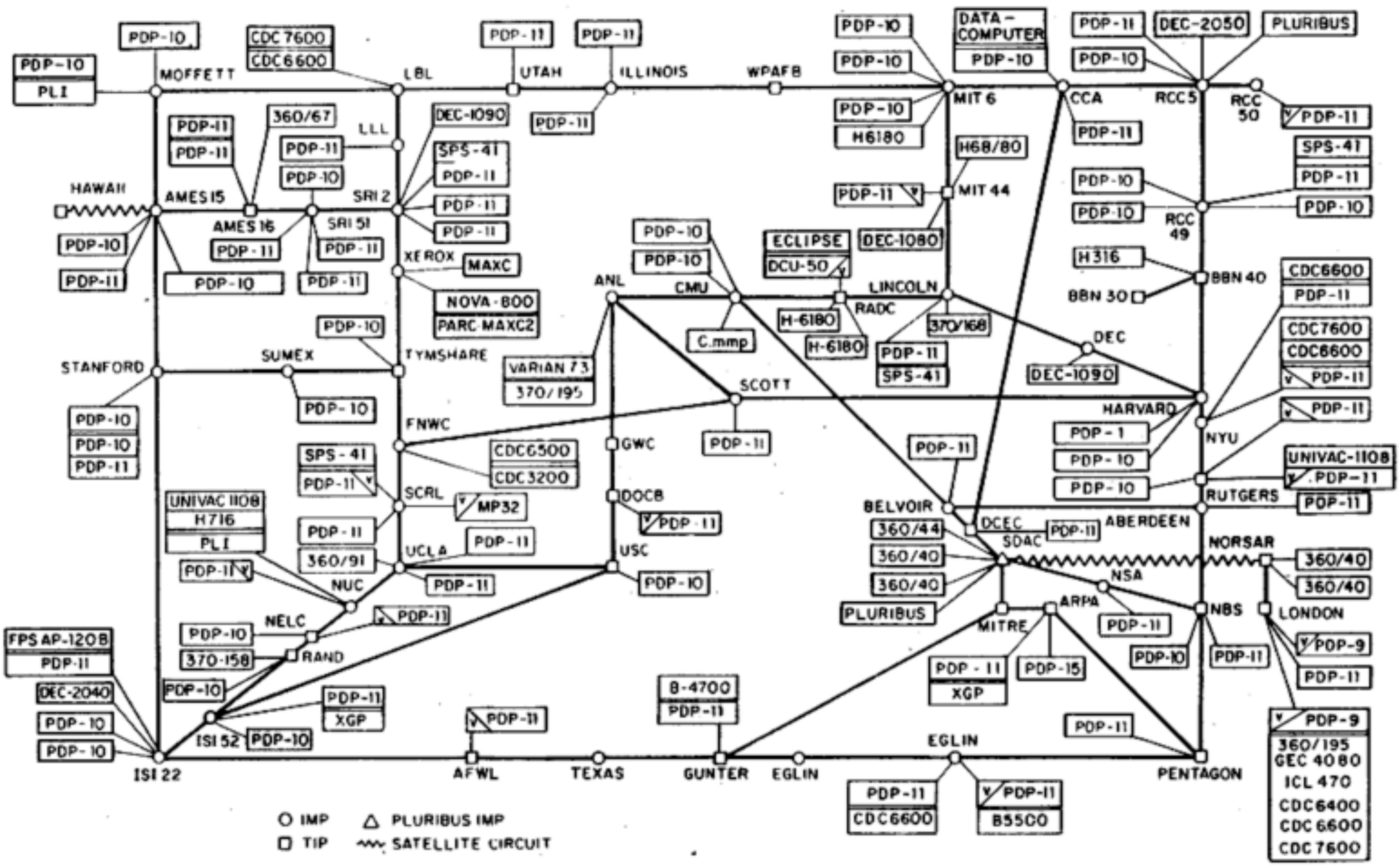
Early national research networks

- NSFNET
  - 1985: 56k leased-lines
  - 1988: T1 backbone (~1.5mbps)
  - 1991: T3/DS3 backbone (45mbps)
  - 1995: Assets sold, project ended, Internet fully commercialized.

History of Research networks: cool maps!

ARPANET (1977)

ARPANET LOGICAL MAP, MARCH 1977

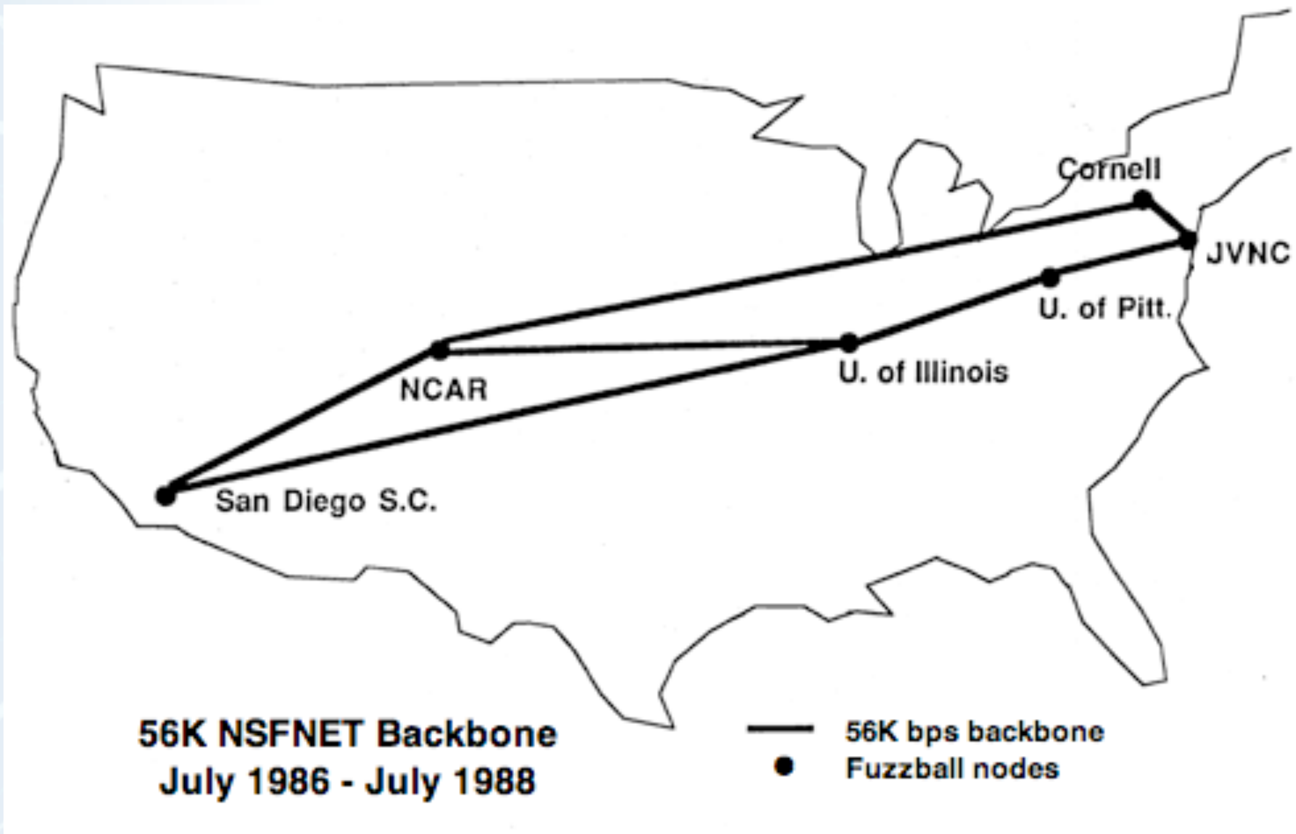


(PLEASE NOTE THAT WHILE THIS MAP SHOWS THE HOST POPULATION OF THE NETWORK ACCORDING TO THE BEST INFORMATION OBTAINABLE, NO CLAIM CAN BE MADE FOR ITS ACCURACY)

NAMES SHOWN ARE IMP NAMES, NOT (NECESSARILY) HOST NAMES

History of Research networks: cool maps!

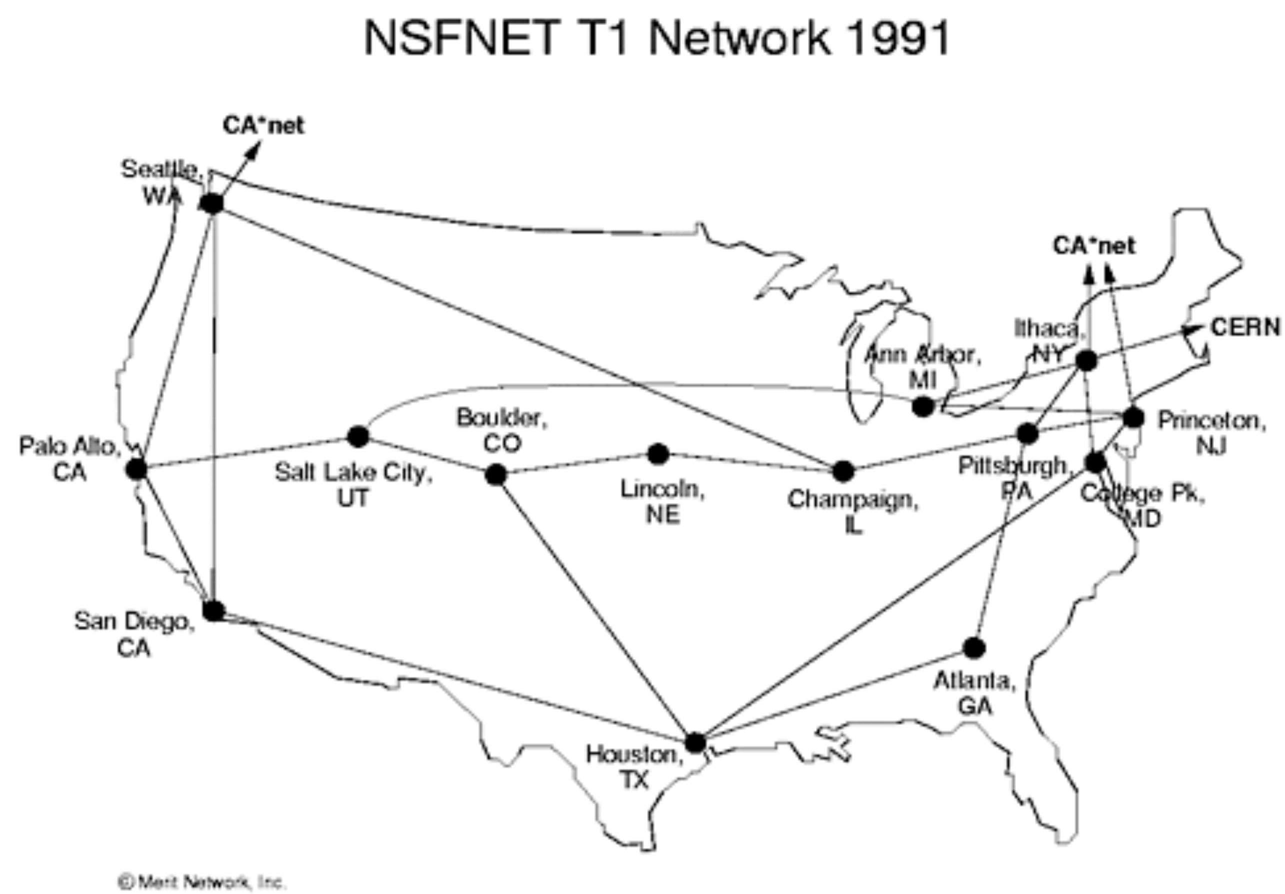
NSFNET 1985



03/29/11

History of Research networks: cool maps!

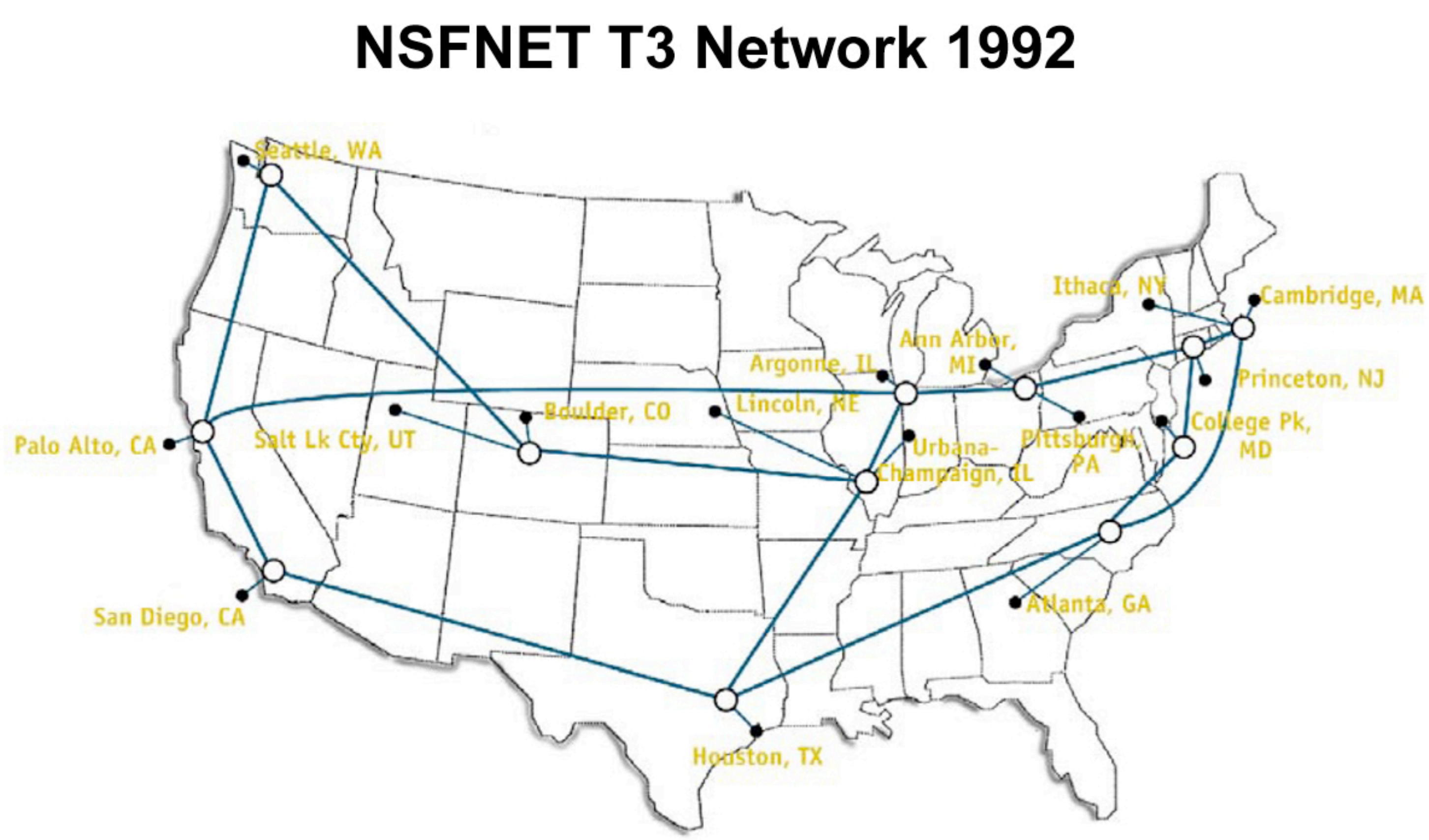
NSFNET 1991 (still T1)



03/29/11

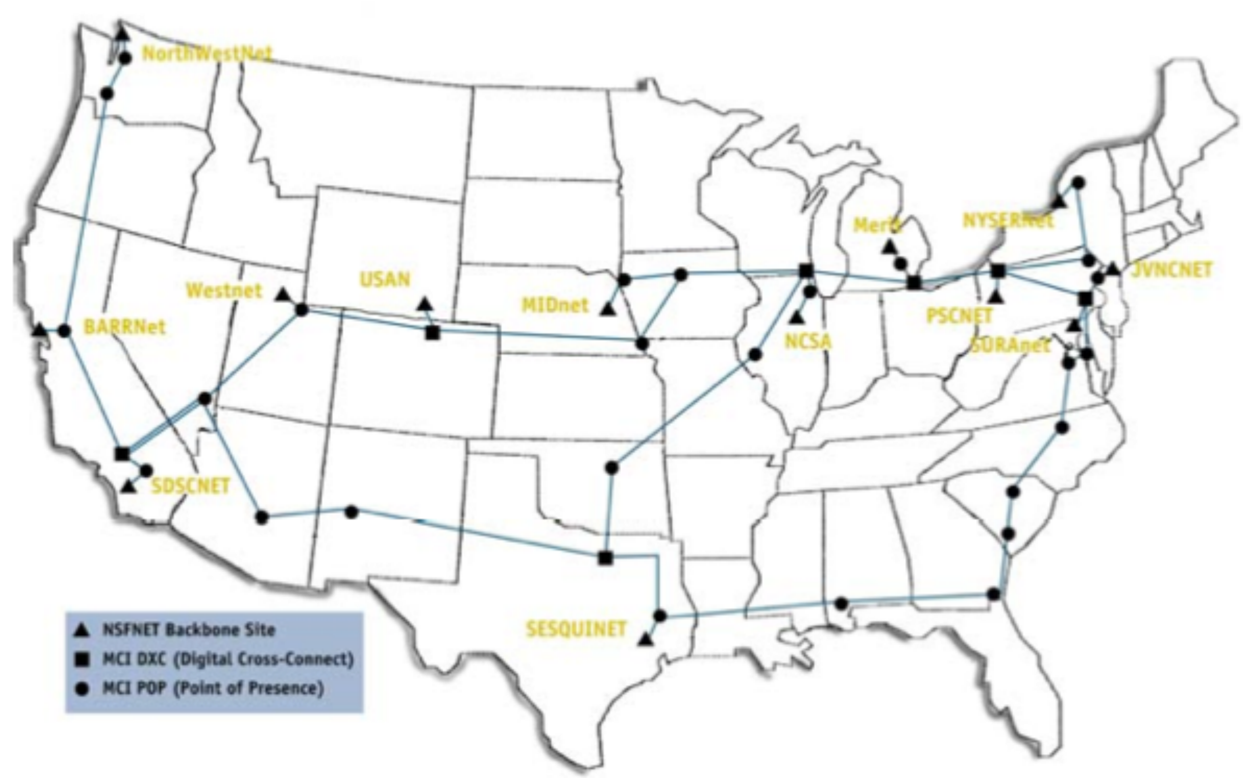
History of Research networks: cool maps!

NSFNET 1992 (T3)



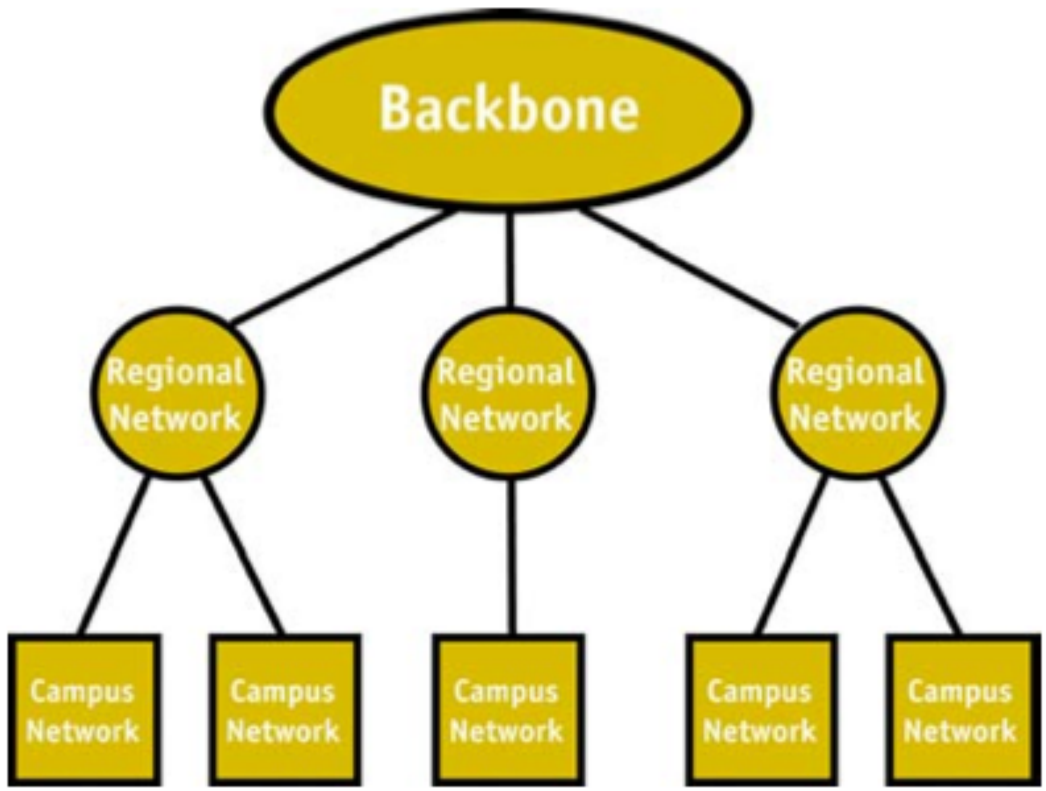
History of Research networks: cool maps!

NSFNET “physical”



History of Research networks: cool maps!

NSFNET 3-tier model



History of Research networks

Early regional research networks

- Merit
  - 1966: one of the first regionals—connected three universities in Michigan.
  - 1987: wins contract to manage NSFnet, along with MCI and IBM.
  - Today: Continues to operate MichNet.
- Los Nettos
  - 1988: founded to provide better connectivity to NSFnet for LA-area entities.
  - Today: Still connects a many LA-area R&E institutions to CENIC and national R&E networks.

North American National Research Networks

ESnet

NLR

Internet2 Network

canarie

n-wave
NOAA RESEARCH NETWORK

NATIONAL AERONAUTICS
AND SPACE ADMINISTRATION

DREN

03/29/11

Confusion! The name of the R&E *organization* isn't always the same as the R&E *network!*

Organization	Network name	DNS name
CANARIE	CA*netN (old) CANARIE (new)	canet4.net (still used) canarie.ca
Internet2	Abilene (old)	ucaid.edu (sometimes)
	Internet2 network	internet2.edu
CENIC	CalREN-DC/HPR/XD	cenic.net
ESnet	ESnetN (curr ESnet5)	es.net
Merit	MichNet	mich.net
DANTE	GEANT	geant.net**

**Actually, the relationship between DANTE and GEANT is even more complicated

Commercialization of the Internet

Why?

- Many R&E networks had AUPs which required usage to be consistent with the funding source's desires and regulations. Some R&E networks still have AUPs (e.g. ESnet).
- There were private companies attached to NSFNET, but they were connected mainly for R&E purposes.
- Growing demand for commercial and personal use of things like e-mail, USENET forums, etc.
- New services like Gopher, WAIS, and that new-fangled WWW were putting increasing demand on the fledgling Internet.
- Increasing tension between the R&E, Government, and commercial communities.

Commercialization of the Internet

Why?

- Increasing bandwidth demands were causing other problems. For example, early networks were interconnected in what Milo Medin called “haphazard ways.”
  - These interconnections sometimes happened through regional networks or sometimes even campus networks!
  - Time to invent peering!
  - Exchanges formed: FIX-East and FIX-West.
- Commodity providers wanted their own exchanges (CIX) and they also wanted access to the FIXes.
- → Need for new exchange points: MAE-East and MAE-West.
- → Need for dedicated commodity networks.

Commercialization of the Internet

Early commodity networks

- UUNET
  - 1988: Founded with a grant from Usenix to provide commodity access to USENET and e-mail services.
  - Today: Core part of Verizon Business.
- PSInet
  - 1989: Founded by NYSERnet staff.
  - Today: Purchased and integrated into Cogent.
- ANS CO+RE
  - 1991: Spun off as a for-profit entity of ANS (the non-profit formed by Merit, et al to manage NSFNET).
  - 1995: Sold networking assets to AOL.

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

Internet2 – Educational institutions do it themselves (courtesy K. Oberman)

- 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 10/100G circuits
  - National backbone provides very high reliability and immunity to circuit failure
  - Primarily connect to regional educational consortia
    - Referred to as “GigaPOPs”—similar to the NSFNET model.
  - Different from commodity networks in some of the services it supported, e.g. Multicast, IPv6.

History of Research networks

Later regional research networks (GigaPOPs or RONs)

- CENIC
  - 1996: Founded by five major California universities
    - To provide consolidated connectivity to the nascent Internet2 network among California R&E institutions.
    - To overcome difficulties after the loss of gen 1 regionals (e.g. Berkeley-Stanford connectivity after BARRnet was privatized).
  - 1997: CalREN-2 network begun with NSF seed money.
  - 2000 – 2001: ONI – Optical Network Initiative: CENIC one of the first research networks to build out its own fiber network.
- Research networks began to realize advantages of operating Layer 1 optical transport gear.

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

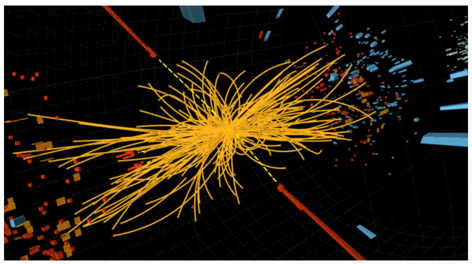
Other research networks

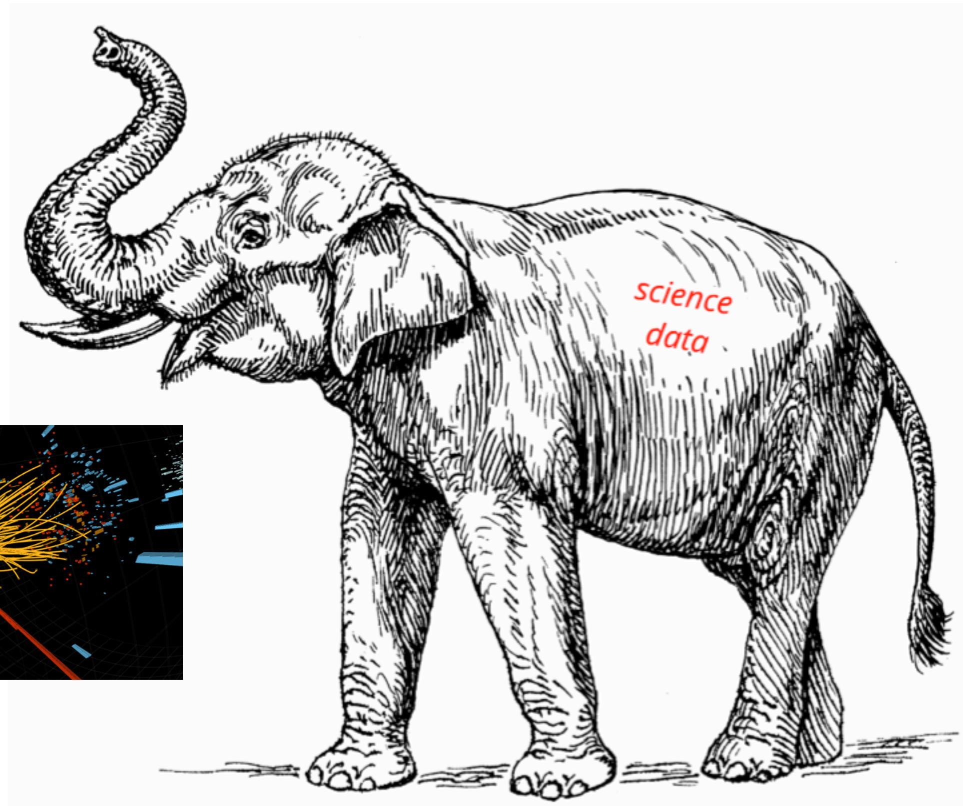
- DREN – the Defense Research and Engineering Network
- N-WAVE – NOAA Research Network
- NREN – NASA Research and Engineering Network
- Provide networks for research of the sponsoring organization
- Largely opaque as they are restricted to constituent use
  - Basically enterprise networks supporting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

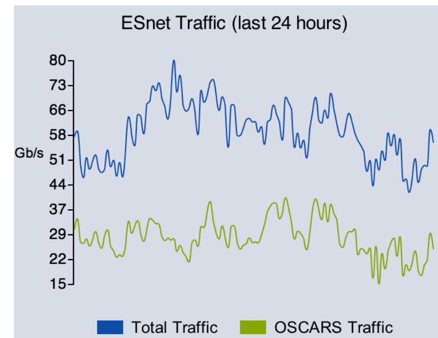
Why Research networks are different

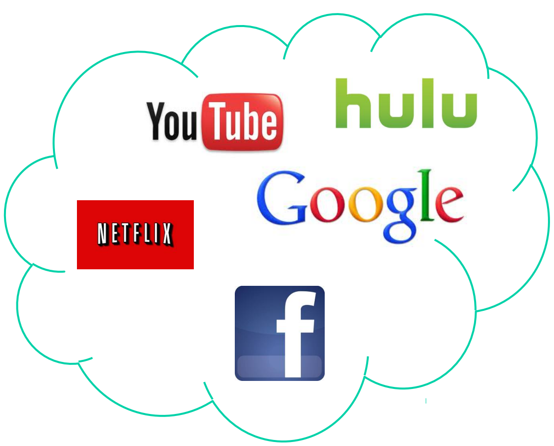
- Move very large scientific data sets (terabytes)
  - The 20 largest flow pairs on ESnet typically account for half of all traffic volume
  - "Big science" is the largest driver
    - LHC
    - Climate and environment research
    - Astrophysics and astronomy
  - Some (e.g. climate) are likely to require 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

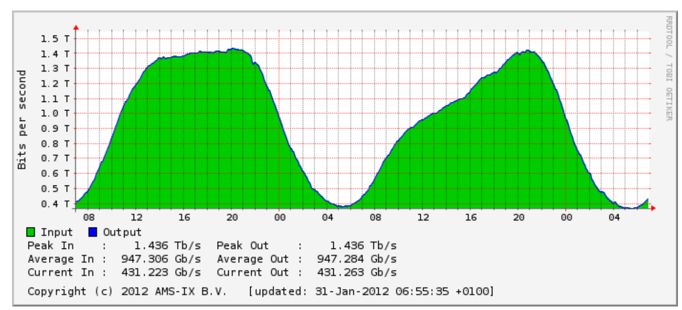
ESnet is not the Commercial Internet







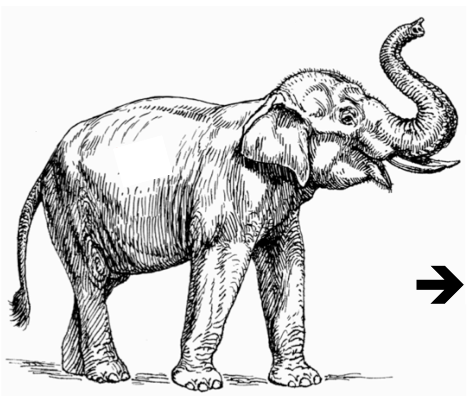




Elephant Flows Place Great Demands on Networks

Physical pipe that leaks water at rate of .0046% by volume.

Result
99.9954% of water transferred.



Network 'pipe' that drops packets at rate of .0046%.

Result
100% of data transferred, *slowly*, at <<5% optimal speed.

essentially fixed

determined by speed of light

$$\frac{\text{maximum segment size}}{\text{round-trip time}} \times \frac{1}{\sqrt{\text{packet-loss rate}}}$$

Through heroic engineering, we can minimize packet loss.

Assumptions: 10Gbps TCP flow, 80ms RTT.

See Eli Dart, Lauren Rotman, Brian Tierney, Mary Hester, and Jason Zurawski. The Science DMZ: A Network Design Pattern for Data-Intensive Science. In *Proceedings of the IEEE/ACM Annual SuperComputing Conference (SC13)*, Denver CO, 2013.

Differences between research and commodity networks

Research networks:

- Tend to be more oriented toward elephant flows (although some educational networks have a big share of mouse flows).
  - Lends itself to a lot of necessary headroom (not overprovisioning).
  - Large buffering not just good, but necessary.
  - Dropped packets are the ENEMY!
  - A lot of R&E engineering goes into preventing packet loss, *including* packet loss due to congestion and issues at ALL layers.
  - And a lot of effort also goes into monitoring networks for *performance* not just availability.
- Ownership of problems and outreach.

Differences between research and commodity networks

Commodity networks:

- Tend to be more oriented toward mouse flows (although there may be exceptions for commercial big data applications and scientific uses of commodity cloud* providers.
  - Allows much less headroom to be provisioned.
  - Congestion okay; TCP will sort things out.
  - As long as buffers stay small!
  - In other words, commodity networks HATE bufferbloat (and users of commodity networks hate bufferbloat even more).
- R&E and commodity networks share the same underlying technology, but it's tuned and adjusted for different applications.

Differences between research and commodity networks

Analogy: Automotive technology

- Think of different types of automobiles. The all use the same underlying technology (internal combustion engine, rubber tires, etc.), but think of how different types of automobiles are tuned for different uses:
  - 18-wheel big rig
  - Formula-1 race car
  - Subaru wagon

Differences between research and commodity networks

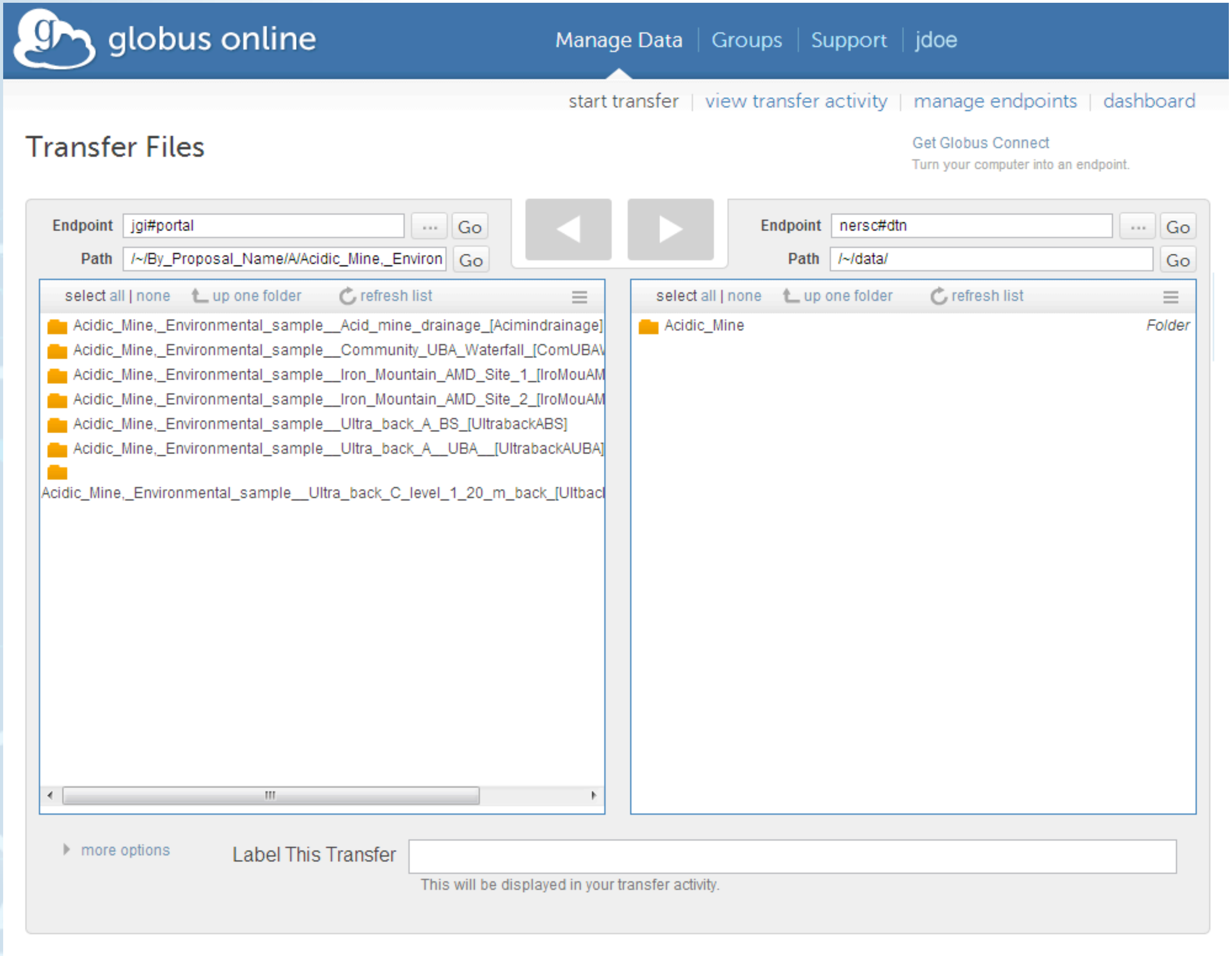
Analogy: Automotive technology

- Now, think of different roads. How do they match the car/truck types?
  - Dirt road (+rain==mud road)
  - Freeway
  - Race track
- Some cars and roads are more all-purpose than others, e.g. Subaru Outback and Freeway.
- Research networks are more like long-distance racetracks (or maybe the autobahn) and commodity networks are more like normal freeways.

Differences between research and commodity networks

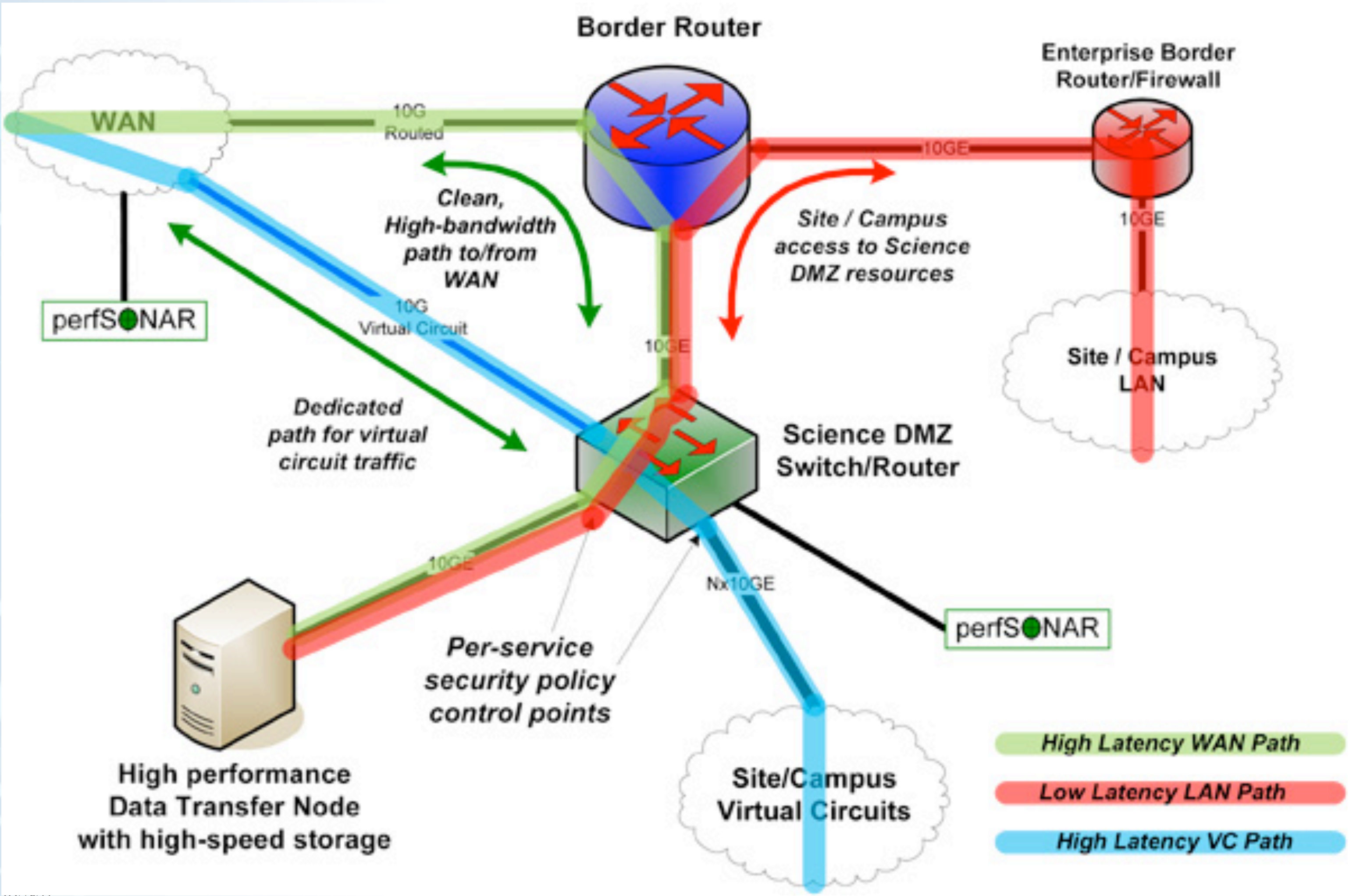
Tools and killer apps

- Good commodity networks will try to anticipate the next killer app and plan for it.
- Good research networks are actively *developing* killer apps for use on their networks.
  - Globus Online and GridFTP
- Establish and/or encourage use of Data Transfer Nodes (DTNs)—specially tuned computers that than transfer lots of data at very high speed (i.e. fill the pipe).
- Promulgate the Science DMZ concept.
- Develop PerfSONAR toolkit for continuous *active testing* of the network.


globus online
Manage Data
Groups
Support
jdoe
start transfer
view transfer activity
manage endpoints
dashboard
Transfer Files
Get Globus Connect
Turn your computer into an endpoint.
Endpoint jgi#portal
Path /~/By_Proposal_Name/A/Acidic_Mine,_Environ
Go
Endpoint nersc#dtn
Path /~/data/
Go
select all | none
up one folder
refresh list
Acidic_Mine,_Environmental_sample__Acid_mine_drainage_[Acimindrainage]
Acidic_Mine,_Environmental_sample__Community_UBA_Waterfall_[ComUBAV
Acidic_Mine,_Environmental_sample__Iron_Mountain_AMD_Site_1_[IroMouAM
Acidic_Mine,_Environmental_sample__Iron_Mountain_AMD_Site_2_[IroMouAM
Acidic_Mine,_Environmental_sample__Ultra_back_A_BS_[UltrabackABS]
Acidic_Mine,_Environmental_sample__Ultra_back_A__UBA__[UltrabackAUBA]
Acidic_Mine,_Environmental_sample__Ultra_back_C_level_1_20_m_back_[Ultbacl
Acidic_Mine
Folder
more options
Label This Transfer
This will be displayed in your transfer activity.

03/29/11

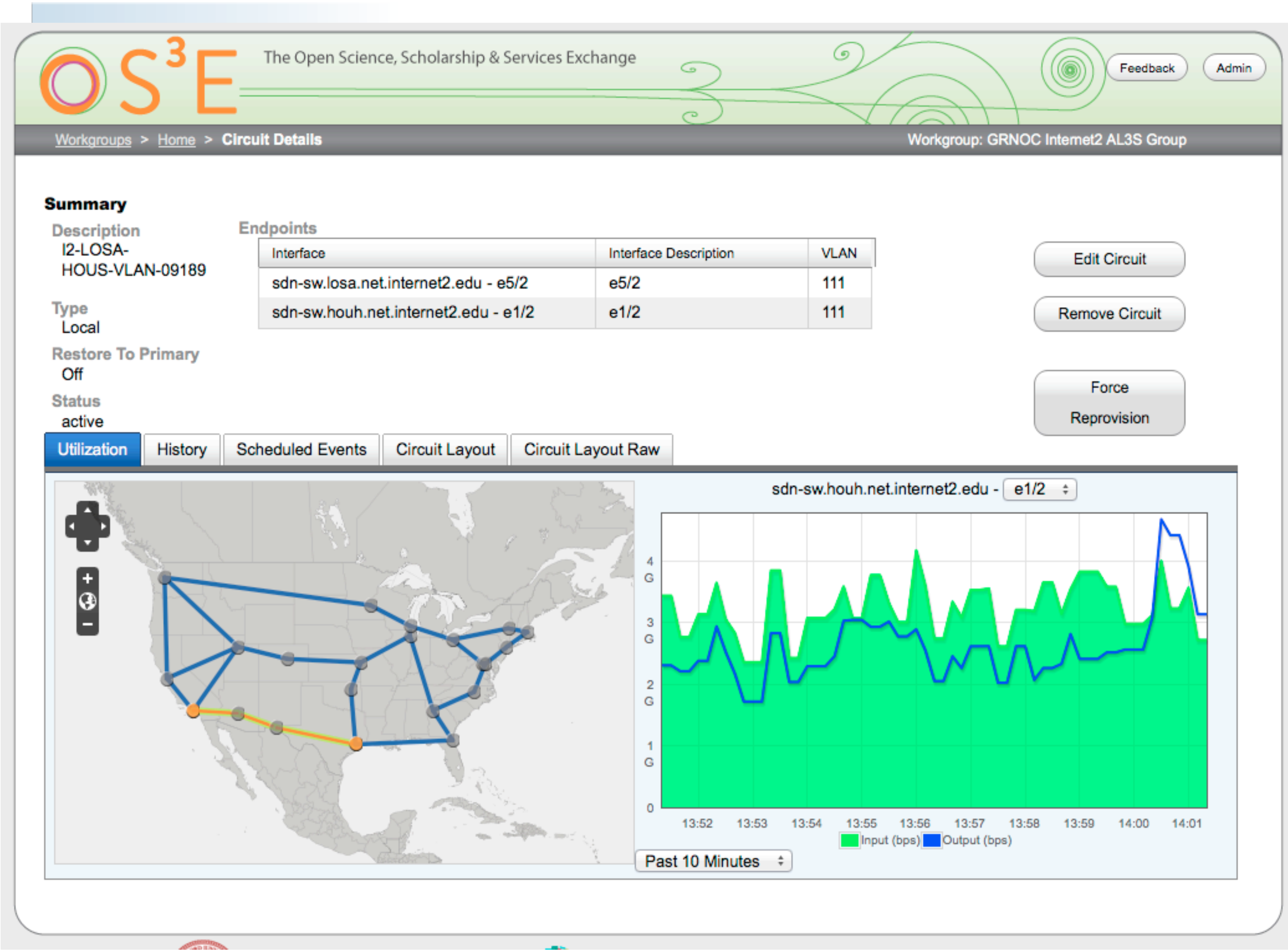
Differences between research and commodity networks



Differences between research and commodity networks

Network manipulation and SDN

- Good commodity networks make good use of MPLS and are actively working on applications of OpenFlow for internal optimization and management of network resources.
- Research networks provide *programmatic interfaces* to allow customers to manipulate the network (without disrupting other customers).
  - ESnet OSCARS: Allows for L2 and L3 VCs to be established using MPLS. Can be used to establish *interdomain* VCs.
  - Internet2 AL2S and OESS: Allows for VCs using OpenFlow—one of the first user-driven, at-scale uses of OpenFlow.


OS3E
The Open Science, Scholarship & Services Exchange
Feedback
Admin
Workgroups > Home > Circuit Details
Workgroup: GRNOC Internet2 AL3S Group
Summary
Description
I2-LOSA-HOUS-VLAN-09189
Type
Local
Restore To Primary
Off
Status
active
Endpoints
Interface | Interface Description | VLAN
sdn-sw.losa.net.internet2.edu - e5/2 | e5/2 | 111
sdn-sw.houh.net.internet2.edu - e1/2 | e1/2 | 111
Edit Circuit
Remove Circuit
Force Reprovision
Utilization
History
Scheduled Events
Circuit Layout
Circuit Layout Raw
sdn-sw.houh.net.internet2.edu - e1/2
4 G
3 G
2 G
1 G
0
13:52
13:53
13:54
13:55
13:56
13:57
13:58
13:59
14:00
14:01
Input (bps)
Output (bps)
Past 10 Minutes

On-demand Secure Circuits and Advance Reservation System

The OSCARS service is provided by ESnet, the Energy Sciences Network.
OSCARS software developed by ESnet, in collaboration with Internet2 and ISI East

May 17, 2014 11:22

ACTIVE

Reservations | Reservation Details | Create Reservation | User Profile | User List | Login/Logout

NEW GRI QUERY

REFRESH MODIFY CANCEL CLONE

GRI	es.net-4599
Status	ACTIVE
User	bmah@es.net
Description	OF TB for ezkissel@indiana.edu, nersc-tb-of-1 to bnl-tb-of-1
Start date	5/16/2014
Start time	10:54
End date	5/26/2014
End time	17:00
Created time	2014/05/16 10:54
Bandwidth (Mbps)	10
Source	urn:ogf:network:domain=es.net:node=nersc-mr2:port=xe-2/1/0:link=*
Destination	urn:ogf:network:domain=es.net:node=bnl-mr3:port=xe-1/1/0:link=*

Path

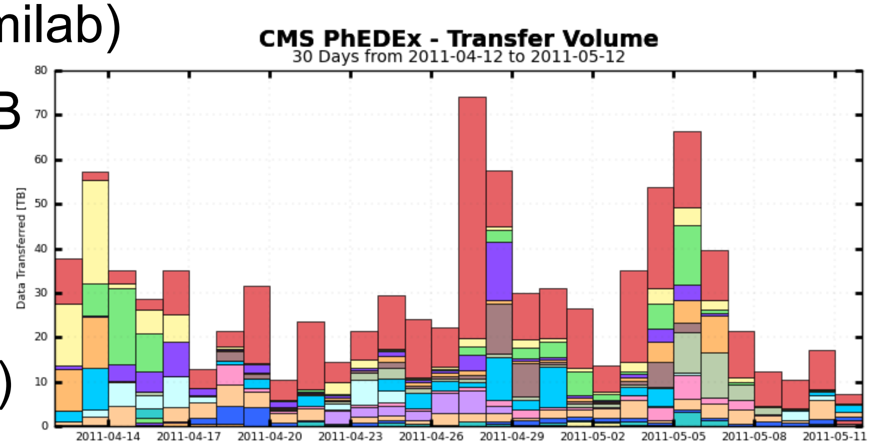
VLAN	Hop
3289	urn:ogf:network:domain=es.net:node=nersc-mr2:port=xe-2/1/0:link=*
n/a*	urn:ogf:network:domain=es.net:node=nersc-mr2:port=xe-8/1/0:link=5
n/a*	urn:ogf:network:domain=es.net:node=sunn-cr5:port=to_nersc-mr2_ip-a:link=0
n/a*	urn:ogf:network:domain=es.net:node=sunn-cr5:port=to_sacr-cr5_ip-a:link=0
n/a*	urn:ogf:network:domain=es.net:node=sacr-cr5:port=to_sunn-cr5_ip-a:link=0
n/a*	urn:ogf:network:domain=es.net:node=sacr-cr5:port=to_denv-cr5_ip-a:link=0
n/a*	urn:ogf:network:domain=es.net:node=denv-cr5:port=to_sacr-cr5_ip-a:link=0
n/a*	urn:ogf:network:domain=es.net:node=denv-cr5:port=to_kans-cr5_ip-a:link=0
n/a*	urn:ogf:network:domain=es.net:node=kans-cr5:port=to_denv-cr5_ip-a:link=0
n/a*	urn:ogf:network:domain=es.net:node=kans-cr5:port=to_chic-cr5_ip-a:link=0
n/a*	urn:ogf:network:domain=es.net:node=chic-cr5:port=to_kans-cr5_ip-a:link=0
n/a*	urn:ogf:network:domain=es.net:node=chic-cr5:port=to_star-cr5_ip-a:link=0
n/a*	urn:ogf:network:domain=es.net:node=star-cr5:port=to_chic-cr5_ip-a:link=0
n/a*	urn:ogf:network:domain=es.net:node=star-cr5:port=to_newy-cr5_ip-a:link=0
n/a*	urn:ogf:network:domain=es.net:node=newy-cr5:port=to_star-cr5_ip-a:link=0
n/a*	urn:ogf:network:domain=es.net:node=newy-cr5:port=to_bnl-mr2_ip-a:link=0
n/a*	urn:ogf:network:domain=es.net:node=bnl-mr2:port=xe-1/1/0:link=5
n/a*	urn:ogf:network:domain=es.net:node=bnl-mr2:port=xe-0/0/0:link=0
n/a*	urn:ogf:network:domain=es.net:node=bnl-mr3:port=xe-0/0/0:link=0
3289	urn:ogf:network:domain=es.net:node=bnl-mr3:port=xe-1/1/0:link=*

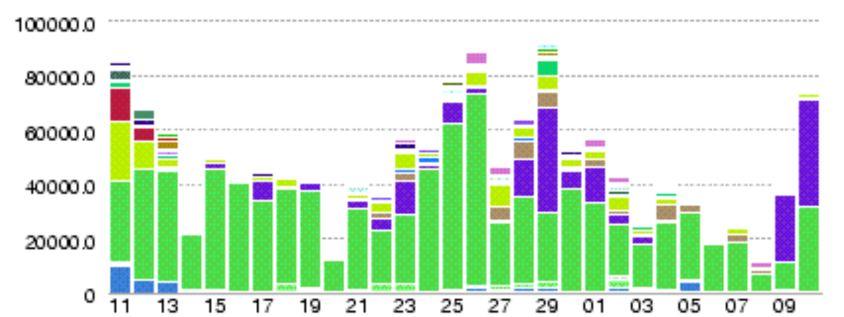
Policing	Soft
Protection	None
Apply QOS	Yes
Source VLAN	3289
Tagged	true

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



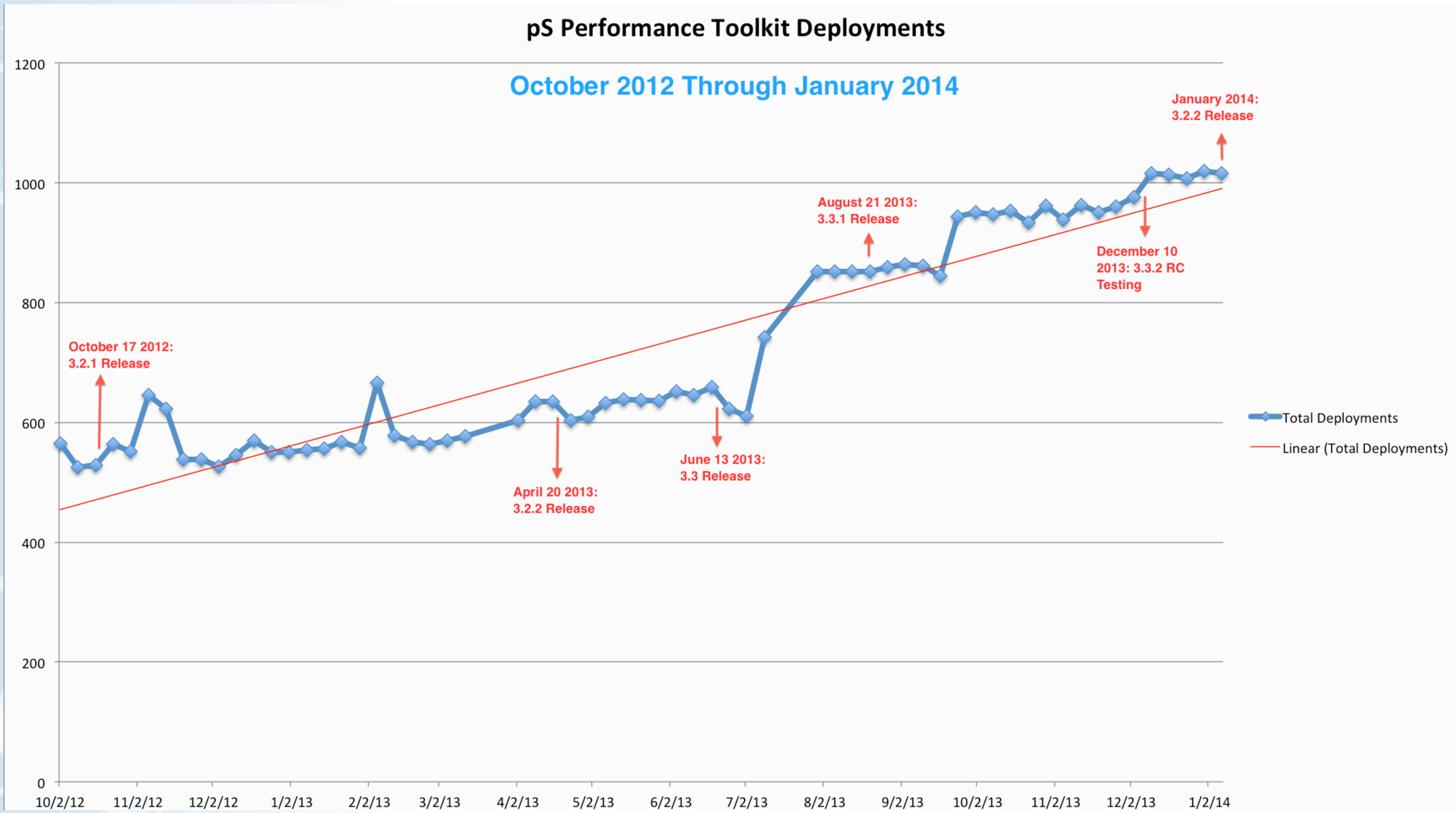


03/29/11

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 is the 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
    - We don’t worry about buffer bloat!

perfSONAR Deployments now at 1100 hosts



The future of research networks

- More bandwidth
  - Most have moved to 100G and 400G is already in trial (==more efficient use of spectrum)
- More agility to better utilize available bandwidth
  - New protocols: NDN, RDMA, enhancements to (or replacement of) TCP
  - SDN and programmability to dynamically redirect different types of flows to different circuits
  - Improvements to OSCARS to better schedule bandwidth
    - Includes automatic L3VPN creation
  - Tune end systems
    - Fasterdata provides a lot of data on system tuning

Commonalities

- Commodity networks peer with each other, usually at commercial exchange points (IXPs). These are operated by companies like Telx, Equinix, Any2, and others.
- Many research networks also peer—both with other research networks and with commodity networks. In addition to being at commodity exchange points, R&E networks may peer at R&E exchanges, e.g. Pacific Wave, WIX, Starlight, etc.
- Some research networks do not purchase transit, nor do they provide transit to their customers. However, such networks are typically not referred to as “Tier 1,” as that term generally applies only to commercial networks.

Previous Differences & Commonalities

Some features were pioneered or initially supported by research networks, but are now more generally supported.

- IPv6:
  - Initially supported by ESnet (first network in North America to have production IPv6) and Internet2.
  - Now widely supported on the Internet.
  - Some* might argue that commodity networks have surpassed research networks in promotion and adoption of IPv6.
- Multicast:
  - Research networks are still among the few that support interdomain multicast, although interest and demand appears to be waning.

*E.g. myself.

Commodity and Research Network Conferences

Or, how did NANOG come about?

- Remember, NSFNET was a backbone network, with a bunch of separate regional networks to provide connectivity to campuses. (Note also that Internet2 and CANARIE have adopted this model.)
- To get the regional network folks together to collaborate, share war stories and best practices, and keep up to date on the technology, NSFNET sponsored a conference called Regional Techs.
- When NSFNET shut down, Regional Techs became NANOG.
  - Initially (until ~2011-2012) managed by Merit.
  - Now a self-sustaining organization.
  - Mostly R&E people at first because that's what the Internet <u>was</u>.

Commodity and Research Network Conferences

- When Internet2 formed, they copied the Regional Techs model and created a meeting called "Joint Techs." This meeting existed until 2012.
- 2014 and beyond, Internet2 has a "technology exchange" meeting every fall.
- Other regions (outside North America) have a *NOG or similar group. These together are commonly called "the NOGs."
- Dedicated research networks often have "stakeholders" or "coordination committee" meetings to provide feedback on network strategy and operations and/or to exchange information.

Summary

- Research and commercial network requirements have diverged
- Still have a great deal in common
- Some lines are being blurred, while new lines are being drawn.
- Expect more collaboration between research and commodity networks (and not just peering)
- 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 and vice versa
    - OpenFlow is a good example…
    - NANOG is a great place for collaboration between both communities