# Early Lessons Learned Deploying a 100Gbps Network

Steve Cotter

Dept Head, Energy Sciences Network

May 4, 2011

Enterprise Innovation Symposium

Atlanta, GA

# Energy Sciences Network Overview

Since 1986, directly supporting the DOE Office of Science's 27,000+ collaborators worldwide with advanced network services and collaboration tools

One of two largest research and education (R&E) networks in the US

- Transports massive quantities of scientific data (more than any other US R&E network) from DOE facilities to the associated community of science collaborators
- Rich connectivity with R&E and commercial networks
  - <15% is Lab-Lab traffic – vast majority terminates or originates off-net
  - More than 140 peerings with all major commercial and R&E networks

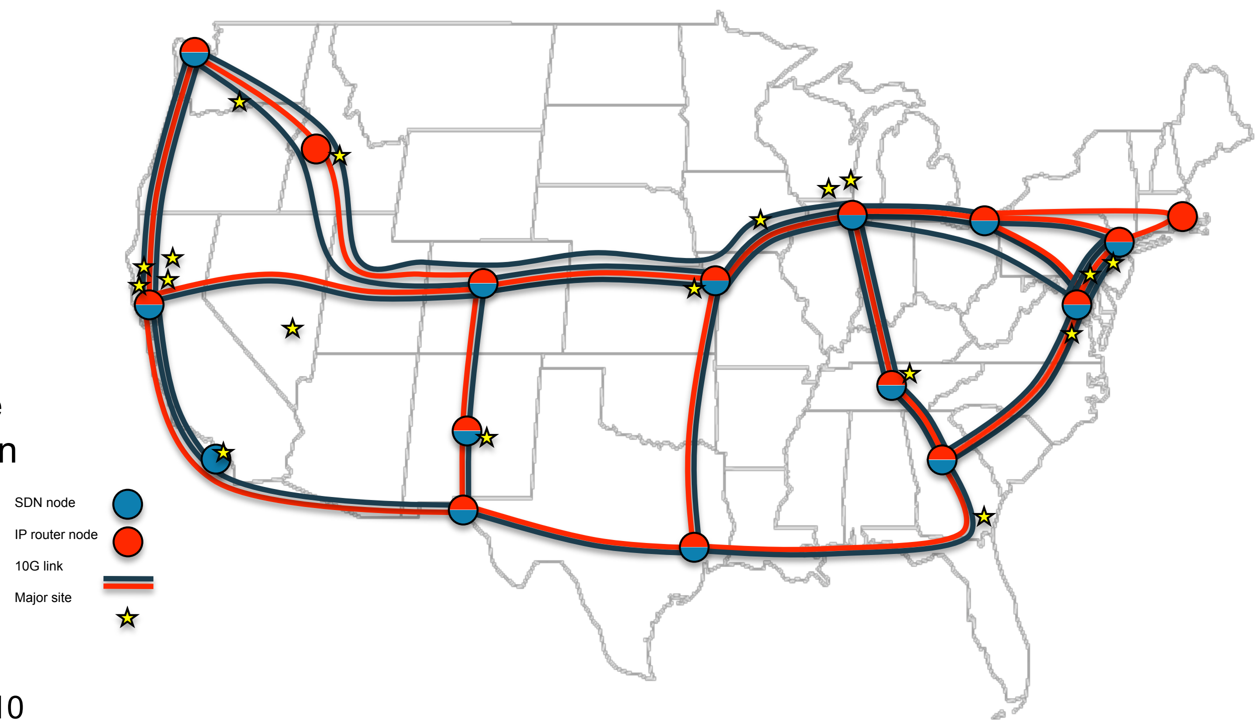
# ESnet4 Network

Two logical networks over shared optical infrastructure

- IP: day-to-day lab operational traffic, ‘mouse’ flows
- Science Data Network: ‘elephant’ flows

Advanced network services

- Heavily instrumented for network performance monitoring
- ESnet’s On-demand, Secure Circuit Advanced Reservation System (OSCARS)
  - 2009 Winner of the American Council for Technology & the Industry Advisory Council Excellence.gov Award
  - 2009 InformationWeek’s Top 10 Govt. IT Innovators Award



# How Is ESnet Different?

Science network traffic is different than the commercial Internet

- Small number (1000's) of very large flows vs. billions of small flows
- Highly variable (peaks and troughs) vs. reasonably predictable
- Extremely low latency (real-time) applications vs. best effort
- As a result, science networking requires unique tools/services/ overlays vs. 'one-size fits all' Internet

Data transfer and data sharing are critical to scientific collaborations – in fact, scientific productivity is often determined by the ability to transfer/stream/share data

# Why 100Gbps now?

# Today’s Trends

Consistent growth since 1990

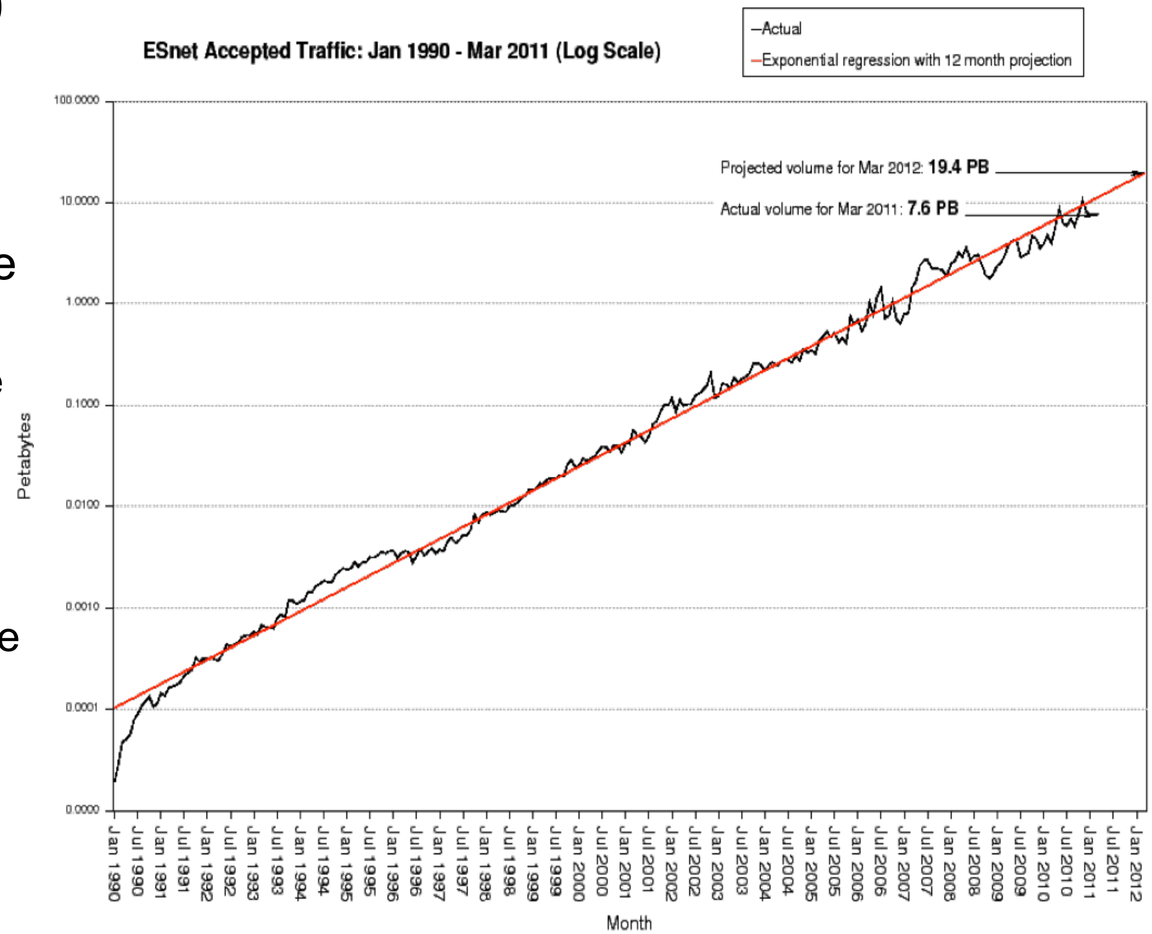
- 10-fold increase every 47 months, on average

More large scientific instruments coming on line

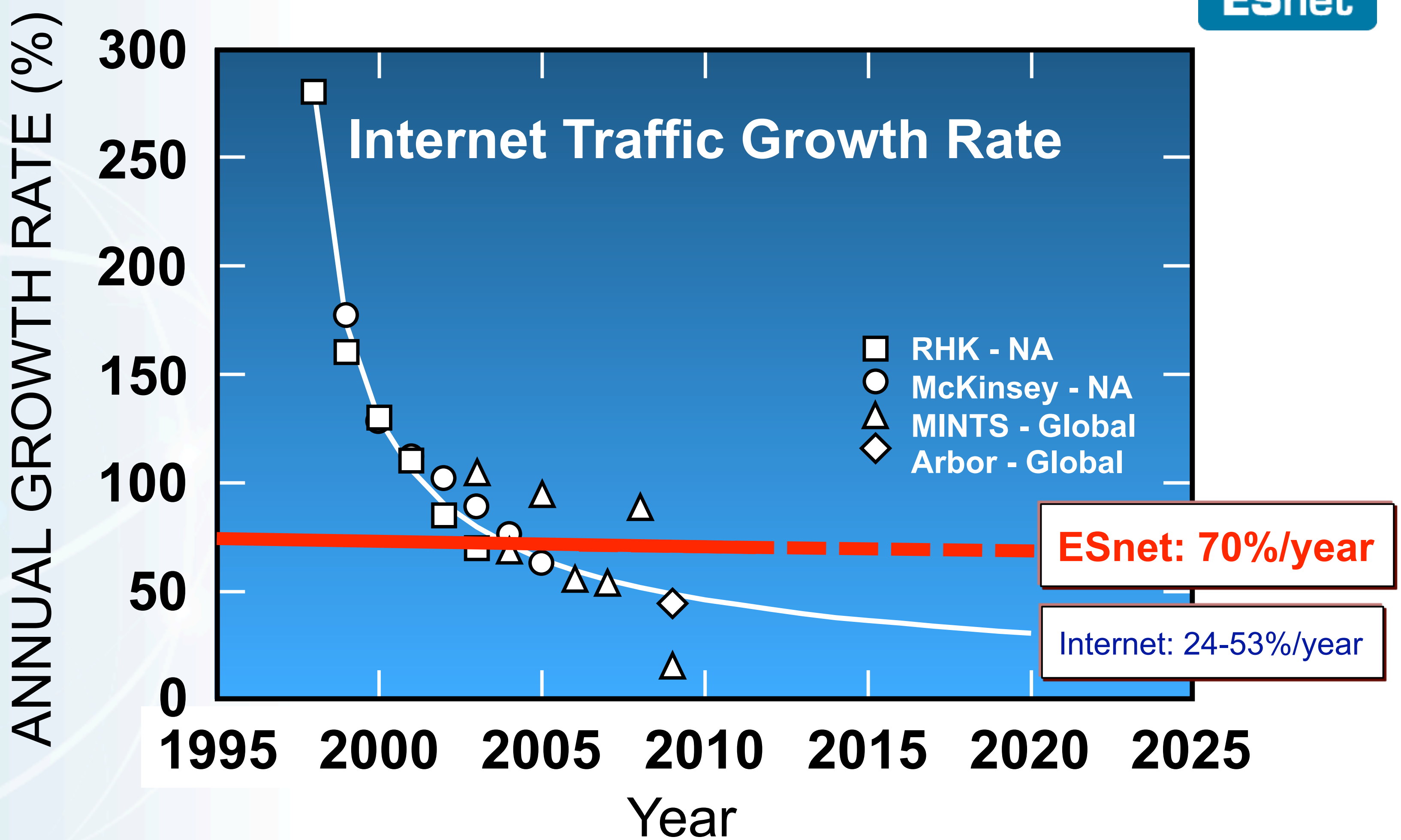
- Large Hadron Collider in CERN, telescopes, advance light sources, etc.

Science becoming more data intensive

- Seeing an acceleration in the generation of data at, if not exceeding, Moore's Law (2x every 18 months)



# Traffic Growth: ESnet vs. Internet




**SKK, 2010 (Sources: RHK, 2004; McKinsey, JPMorgan, AT&T, 2001; MINTS, 2009; Arbor, 2009).** Courtesy of Steve Korotky

# ARRA Advanced Networking Initiative (ANI)

ESnet received $62.4M in ARRA funds from DOE for an Advanced Networking Initiative to:

- Build an end-to-end 100Gbps prototype network
- Handle proliferating data needs between the three DOE supercomputing facilities and NYC international exchange point
- Build a network testbed facility for researchers and industry

DOE is also funding $5M in network research using the testbed facility: goal of near-term technology transfer to the ESnet production network

Separately funded $33 million for Magellan, an associated DOE cloud computing project that will utilize the 100Gbps network infrastructure

# Looking Beyond 100Gbps

DOE Terabits Workshop, Feb 16-17, 2011

http://science.energy.gov/~/media/ascr/pdf/program-documents/docs/Terabit_networks_workshop_report.pdf

DOE has a need to support 'exascale' applications, e.g. simulation, experiment data management, data analysis and remote visualization with terabit networking

- Fueled by the deluge of data being generated by researchers working across areas such as high energy physics, climate change, genomics, fusion, synchrotron light sources, etc.

Many future projections calling for exabyte bulk transfers by the end of the decade among DOE supercomputing facilities, large scientific instrument facilities, and university and international partner institutions.

# Looking Beyond 100Gbps

Findings:

- Exascale architectures will be under very tight energy budgets in 2018-2022 timeframe
  - Design breakthroughs necessary to keep exascale computers within 20MW range
- Imperative for end-to-end networks to provide terabit scalability w/o dramatically increasing energy profile
  - Need to leverage silicon & photonic integration and seamless integration of low-power computing architectures with campus LAN and WAN terabit networks

# Green House Gas Executive Order

Oct 5, 2009; President Obama signs an Executive Order Focused on Federal Leadership in Environmental, Energy, and Economic Performance
http://www.whitehouse.gov/assets/documents/2009fedleader_eo_rel.pdf

'The Executive Order requires Federal agencies to set a 2020 greenhouse gas emissions reduction target within 90 days; increase energy efficiency; ...relative to a fiscal year 2008 baseline…'

'The Executive Order builds on and expands the energy reduction and environmental requirements of Executive Order 13423 by making reductions of greenhouse gas emissions a priority of the Federal government, and by requiring agencies to develop sustainability plans focused on cost-effective projects and programs.'

# Looking at the Problem

# ICT Carbon Footprint

Worldwide ICT carbon footprint:

2007:

- 2% = 830 m tons $CO_2$
- Comparable to the global aviation industry

2020:

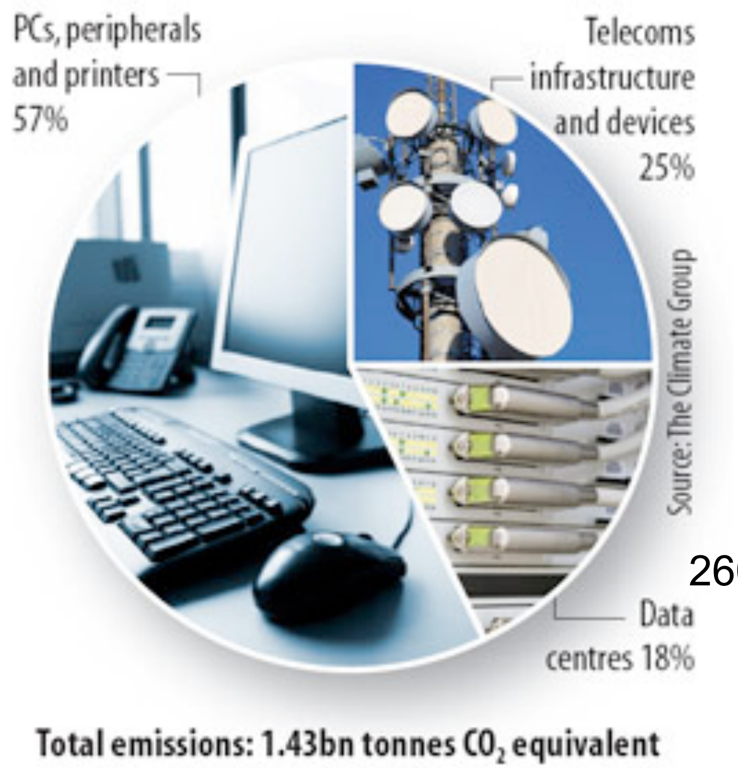
- Expected to grow to 4%

The Climate Group, GeSI report “Smart 2020”, 2008

## 2020 ICT Carbon Footprint

820m tons $CO_2$

360m tons $CO_2$




260m tons $CO_2$

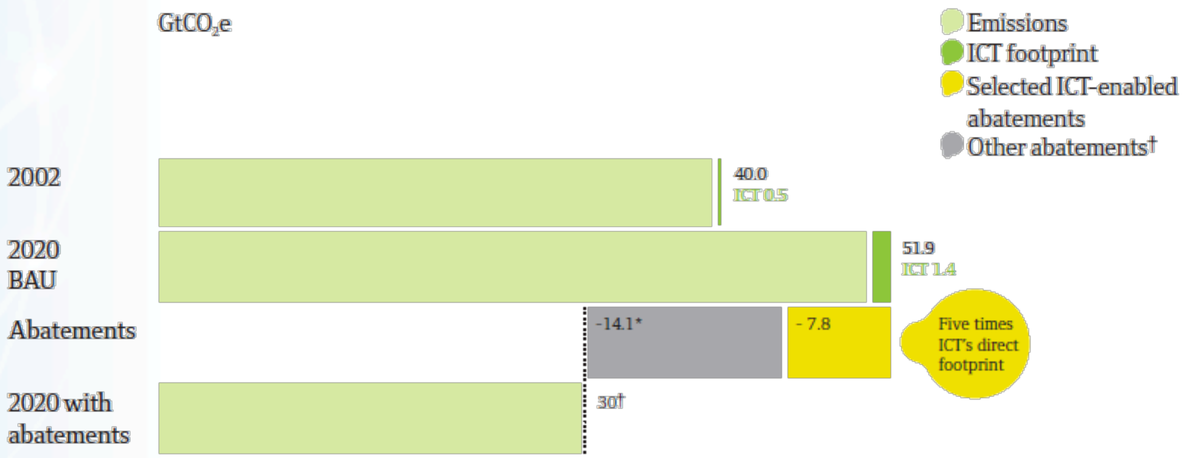
Slide courtesy of Dan Kilper, Bell Labs

# ICT’s Impact Goes Beyond its Footprint

ICT including networks are expected to play a critical role in reducing global GHG emissions

- Some predictions are that ICT can achieve global emissions reductions of 7.8 $GtCO_2e$ in 2020, 5x its est. carbon footprint

Fig. 1 ICT impact: The global footprint and the enabling effect




* For example, avoided deforestation, wind power or biofuels.
† 21.9 $GtCO_2e$ abatements were identified in the McKinsey abatement cost curve and from estimates in this study. Source: Enkvist P., T. Naucler and J. Rosander (2007), *'A Cost Curve for Greenhouse Gas Reduction'*, The McKinsey Quarterly, Number 1.

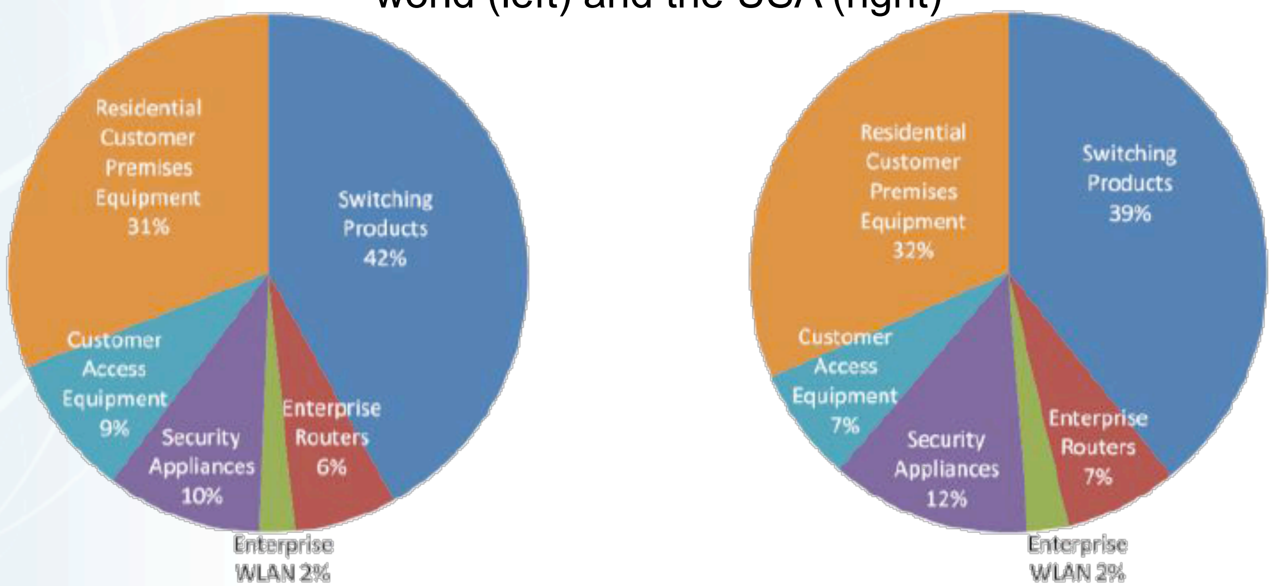
The Climate Group, ‘Smart 2020’, 2008

# Networking Equipment

2009 LBNL Report* determined network equip in buildings (excl. edge devices or core networks) consumed 18TWh in US, 51TWh globally in 2008

- 1% of building energy consumption
- Growing at **6%** per year

Breakdown of Energy Use by Major Product Category in 2008
world (left) and the USA (right)




*'Data Network Equipment Energy Use and Savings Potential in Buildings'; Lanzisera, Nordman, Brown

# The Opportunity

Latest generation of data centers quite efficient

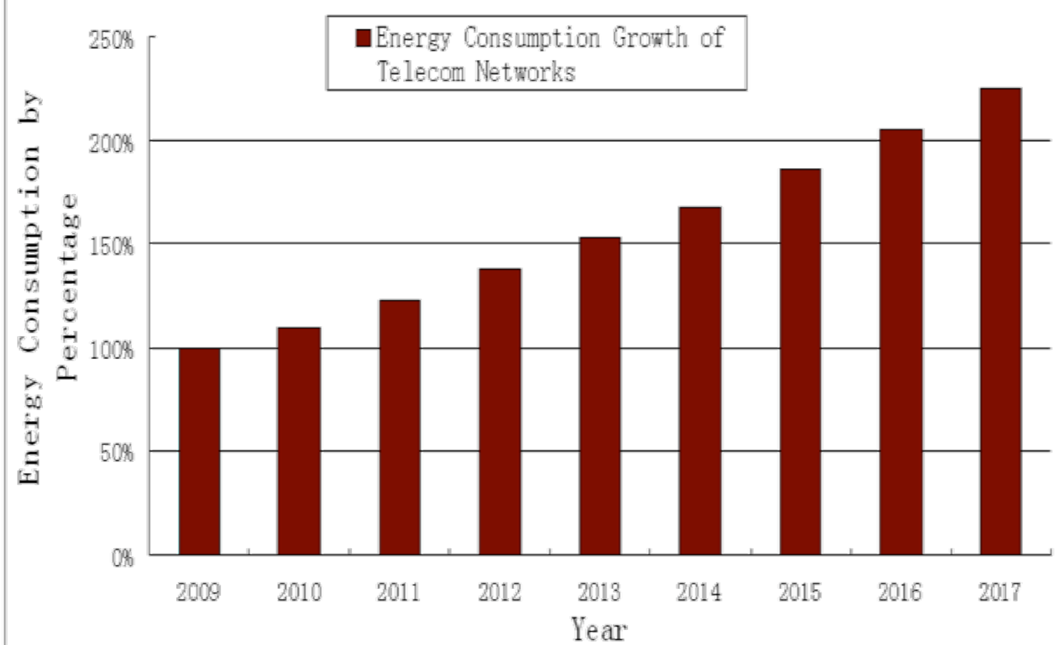
- Future efficiencies will come from more-efficient hardware and cleaner sources of energy http://gigaom.com/cleantech/google-green-czar-no-moores-law-for-data-center-efficiency/

LBNL's Lanzisera, Nordman, Brown report:

- Reductions of ~20% are easily achievable, potential savings >50%

Other research indicates big gains possible

Percentage increase of energy consumption in Telecom networks – double by 2017 of existing 2009 levels




Source: Yi Zhang; Chowdhury, P.; Tornatore, M.; Mukherjee, B.; , "Energy Efficiency in Telecom Optical Networks," Communications Surveys & Tutorials, IEEE , vol.12, no.4, pp.441-458, Fourth Quarter 2010

# ESnet's Energy Consumption

# Putting It in Perspective

Power usage:

- CERN supercollider: 180 MW
- DOE supercomputer: 6 MW
- ESnet: 400 KW

Although small in relative terms, end-to-end understanding of the energy consumption will lead to architectural insights with impact on network energy-efficiency

# ESnet's 2008 Baseline

No incentive to track energy consumption of network

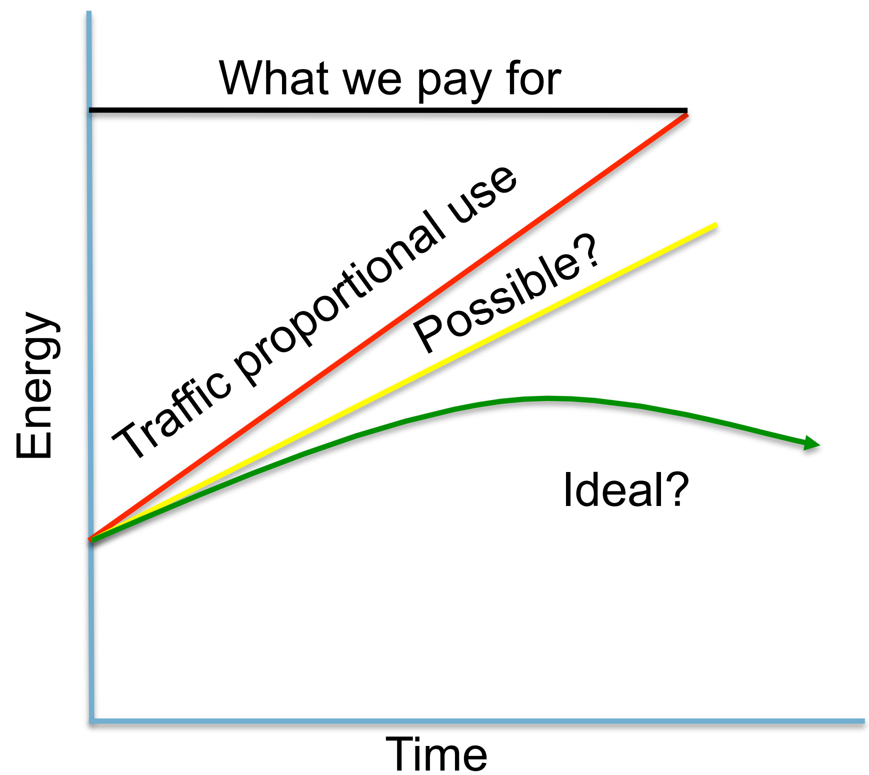
- Focused on meeting demand
- Pay for near-max power day network is commissioned

No good way to track 'real' network energy consumption

- Breakered vs. metered power
- Lack of visibility into equipment

What is the energy curve of my new 100 Gbps network - other than "More"?

Network Energy Use Prediction



# More Challenges

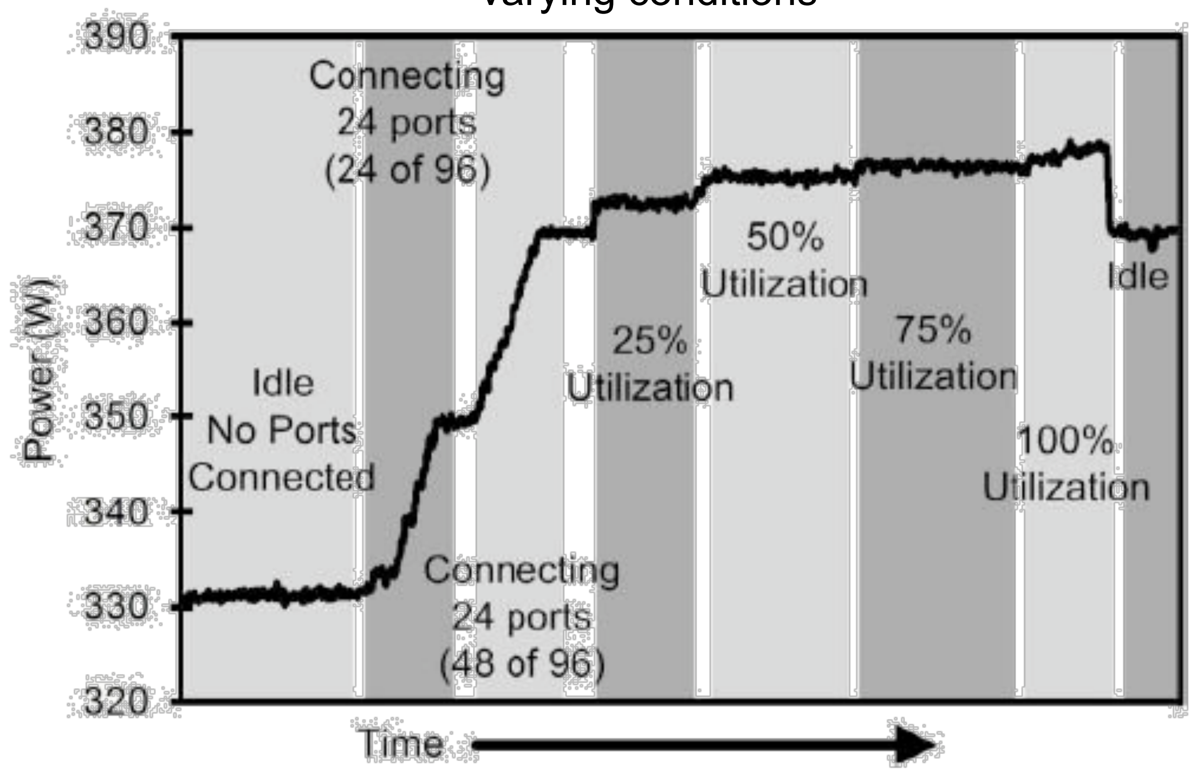
Energy efficiency research for networks is still nascent

- Theoretical models not backed by real live data, Top-down/ bottom-up different

What we do know: little correlation between traffic and power

- Power changes little with data utilization and moderately with port utilization

Power of a modular network switch under varying conditions




*'Data Network Equipment Energy Use and Savings Potential in Buildings'; Lanzisera, Nordman, Brown

# Leveraging the ANI Project

Instrument the 100 Gbps network for energy consumption and take live measurements

- Build the tools to collect and visualize live network energy consumption
- Correlate it with network traffic information
- Build tools to analyze the data on a per network service, per network layer basis
  - Improve theoretical models by comparing with real power consumption
- Make the data OPEN!

Develop models to predict more accurately the energy consumption of various application scenarios

- Data-intensive Science, Cloud Computing
- Solutions like Follow-the-Moon/Sun model, Data to compute/computation to data movement etc.

# Long-term Goals

DOE should stand for the Department of **(LESS)** Energy

Create an energy-aware network ecosystem - tools, equipment, PDUs, meters, business models to incorporate traffic proportionality

Create comprehensive GHG impacts through trade-off application studies based on empirically-verified network energy-efficiency models

- For ex. What is the impact of moving data to computation? Is it more energy efficient to move computation to data?
- Is Cloud more energy efficient than local computing - under what scenarios?

Twitter: ESnetUpdates

Blog: http://esnetupdates.wordpress.com/

# Thanks!