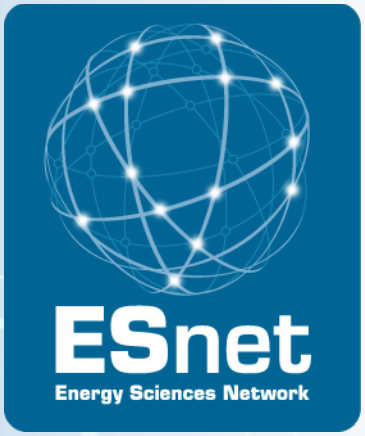


The perfSONAR Project at 10 Years: Status and Trajectory

Jason Zurawski - **ESnet Engineering & Outreach**

engage@es.net

GN3 (GÉANT) NA3, Task 2 - Campus Network Monitoring and Security Workshop

April 24th 2014

perfSONAR powered

With contributions from S. Balasubramanian, G. Bell, E. Dart, M. Hester, B. Johnston, A. Lake, E. Pouyoul, L. Rotman, B. Tierney and others @ ESnet

Overview

- Introduction
- The Ghost of perfSONAR Past
- perfSONAR Present
- Use Cases
- Future Directions & Unfinished Business

Overarching Motivation

Networks are an essential part of data-intensive science

- Connect data sources to data analysis
- Connect collaborators to each other
- Enable machine-consumable interfaces to data and analysis resources (e.g. portals), automation, scale

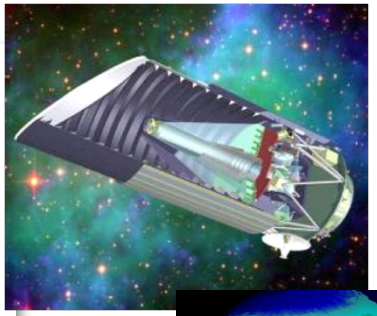
Performance is critical

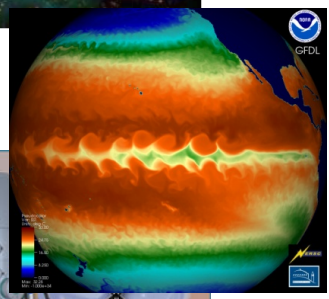
- Exponential data growth
- Constant human factors
- Data movement and data analysis must keep up

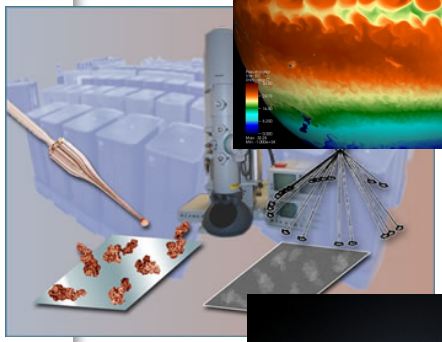
Effective use of wide area (long-haul) networks by scientists has historically been difficult

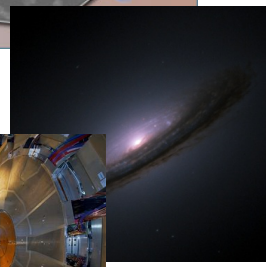
perfSONAR powered

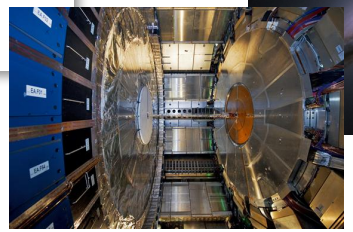
ESnet **Supports DOE Office of Science**











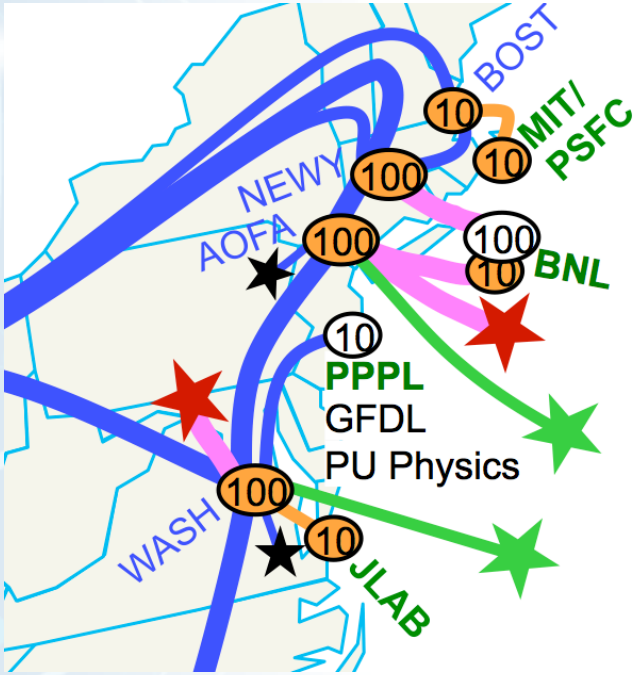
The Office of Science supports:

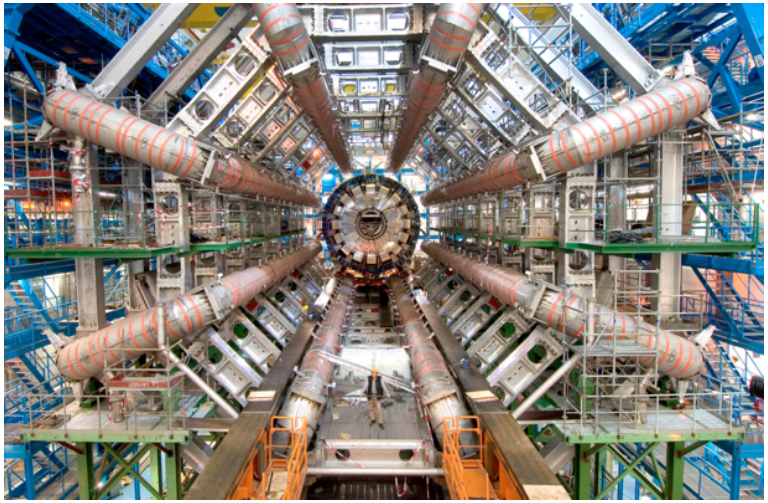
- 27,000 Ph.D.s, graduate students, undergraduates, engineers, and technicians
- 26,000 users of open-access facilities
- 300 leading academic institutions
- 17 DOE laboratories

perfSONAR powered

ESnet's Vision

Scientific progress is **completely unconstrained** by the physical location of instruments, people, computational resources, or data.





The Central Role of the Network

The very structure of modern science assumes science networks exist: high performance, feature rich, global scope

What is “The Network” anyway?

- “The Network” is the set of devices and applications involved in the use of a remote resource
  - This is not about supercomputer interconnects
  - This is about data flow from experiment to analysis, between facilities, etc.
- User interfaces for “The Network” – portal, data transfer tool, workflow engine
- Therefore, servers and applications must also be considered

What is important?

1. Correctness
2. Consistency
3. Performance

perfSONAR powered

Sample Data Transfer Rates

Data set size	1 Minute	5 Minutes	20 Minutes	1 Hour
	Time to transfer			
10PB	1,333.33 Tbps	266.67 Tbps	66.67 Tbps	22.22 Tbps
1PB	133.33 Tbps	26.67 Tbps	6.67 Tbps	2.22 Tbps
100TB	13.33 Tbps	2.67 Tbps	666.67 Gbps	222.22 Gbps
10TB > 100Gbps	1.33 Tbps	266.67 Gbps	66.67 Gbps	22.22 Gbps
1TB	133.33 Gbps	26.67 Gbps	6.67 Gbps	2.22 Gbps
100GB 100Gbps	13.33 Gbps	2.67 Gbps	666.67 Mbps	222.22 Mbps
10GB < 10Gbps	1.33 Gbps	266.67 Mbps	66.67 Mbps	22.22 Mbps
1GB	133.33 Mbps	26.67 Mbps	6.67 Mbps	2.22 Mbps
100MB < 100Mbps	13.33 Mbps	2.67 Mbps	0.67 Mbps	0.22 Mbps

This table available at:

http://fasterdata.es.net/fasterdata-home/requirements-and-expectations/

perfSONAR powered

TCP – Ubiquitous and Fragile

Networks provide connectivity between hosts – how do hosts see the network?

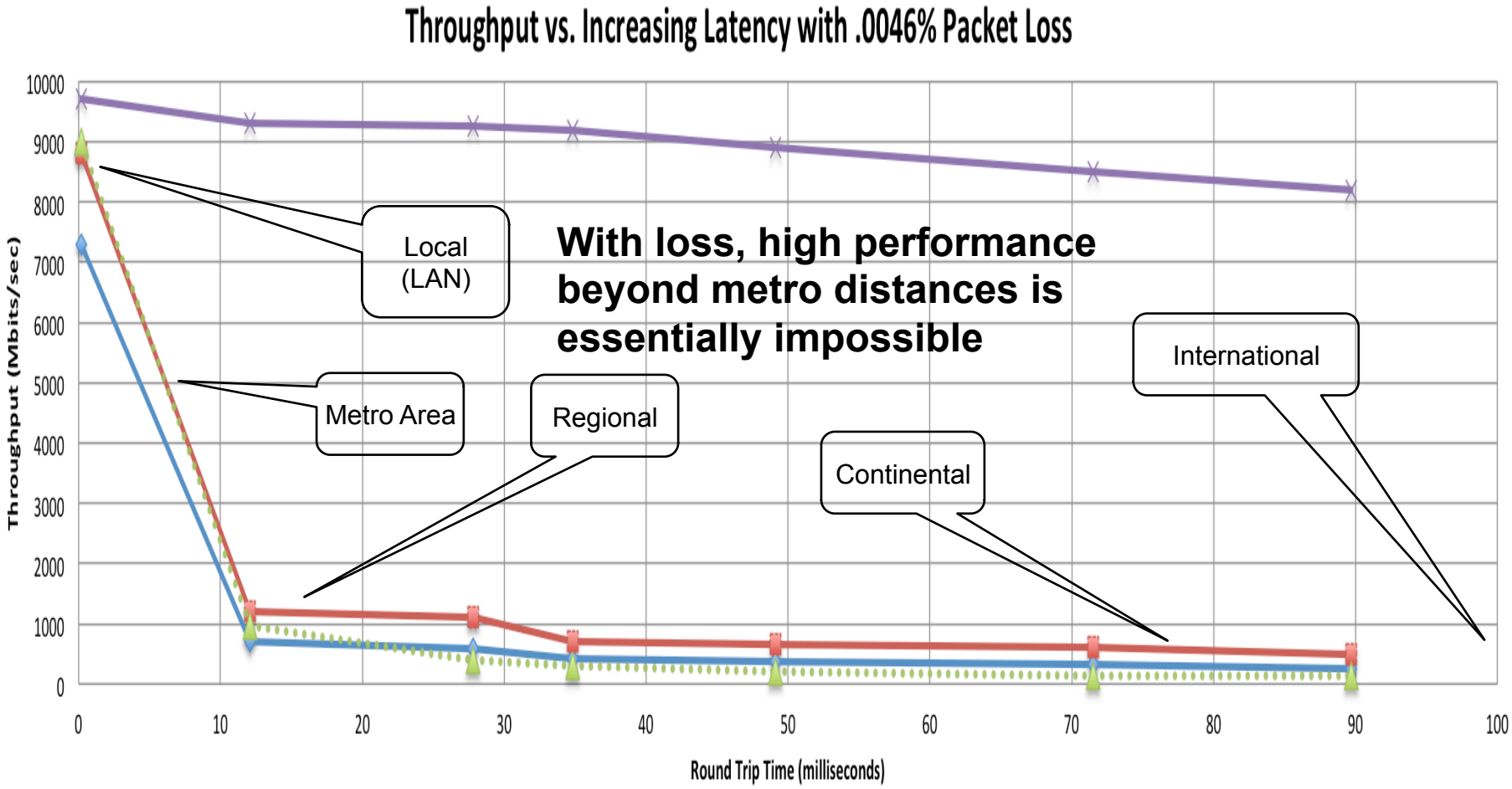
- From an application's perspective, the interface to "the other end" is a socket
- Communication is between applications – mostly over TCP

TCP – the fragile workhorse

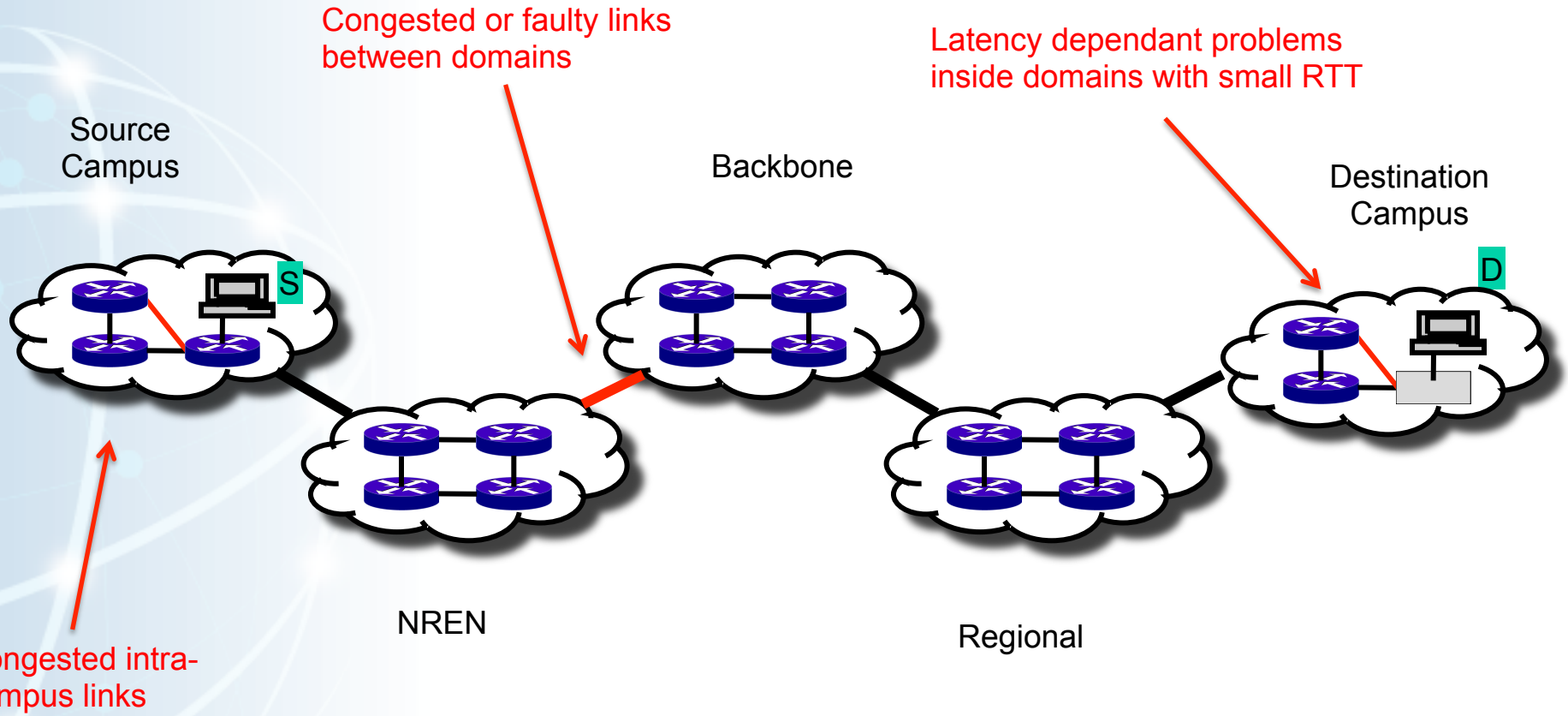
- TCP is (for very good reasons) timid – packet loss is interpreted as congestion
- Packet loss in conjunction with latency is a performance killer
- Like it or not, TCP is used for the vast majority of data transfer applications (more than 95% of ESnet traffic is TCP)

perfSONAR powered

A small amount of packet loss makes a huge difference in TCP performance

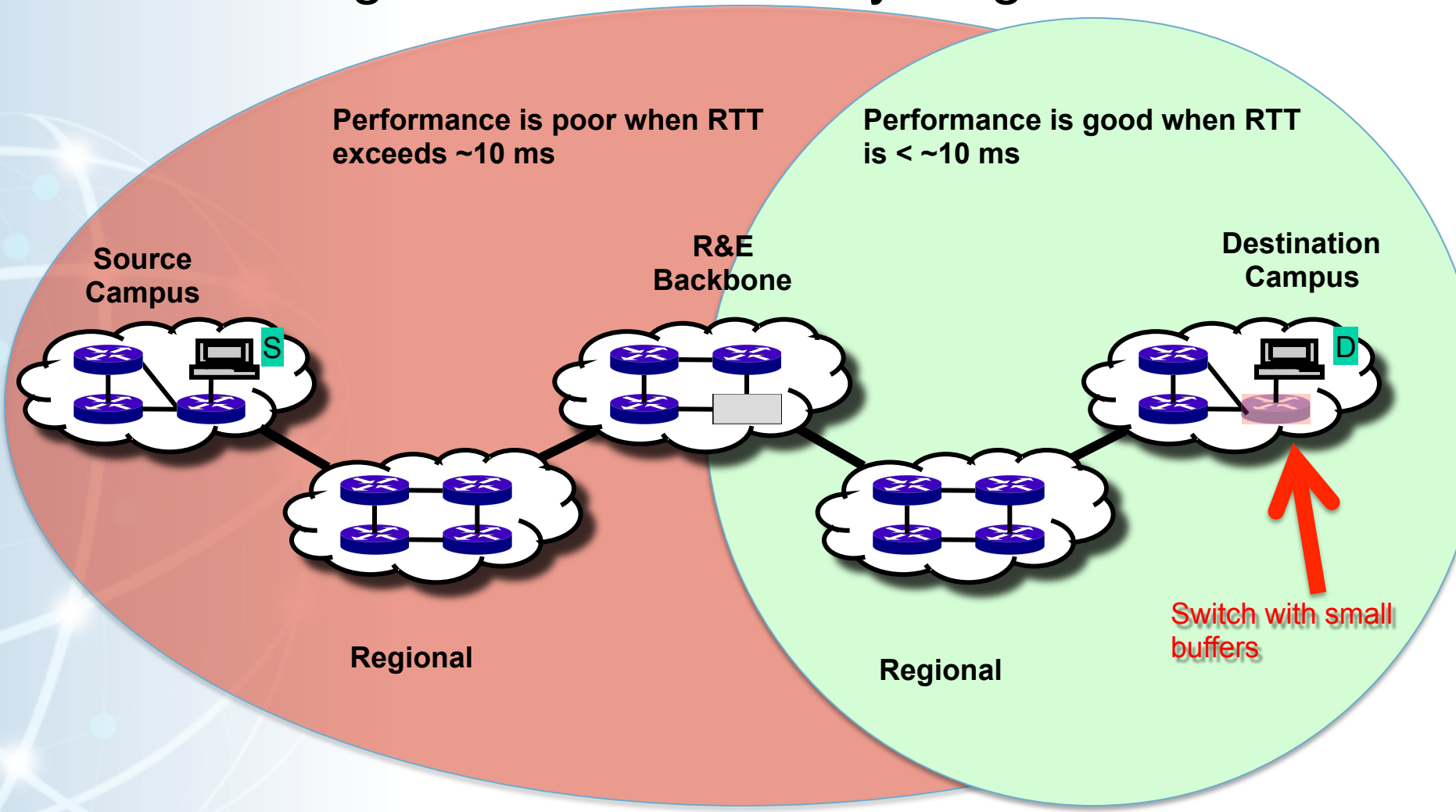


Where Are The Problems?



perfSONAR powered

Local Testing Will Not Find Everything



perfSONAR powered

Soft Network Failures

Soft failures are where basic connectivity functions, but high performance is not possible.

TCP was intentionally designed to hide all transmission errors from the user:

- "As long as the TCPs continue to function properly and the internet system does not become completely partitioned, no transmission errors will affect the users." (From IEN 129, RFC 716)

Some soft failures only affect high bandwidth long RTT flows.

Hard failures are easy to detect & fix

- soft failures can lie hidden for years!

One network problem can often mask others

perfSONAR powered

The Metrics We Care about

Use the correct tool for the Job

- To determine the correct tool, maybe we need to start with what we want to accomplish …

What do we care about measuring?

- Packet Loss, Duplication, out-of-orderness (transport layer)
- Achievable Bandwidth (e.g. “Throughput”)
- Latency (Round Trip and One Way)
- Jitter (Delay variation)
- Interface Utilization/Discards/Errors (network layer)
- Traveled Route
- MTU Feedback

perfSONAR powered

Network Monitoring

- All networks do some form monitoring.
  - Addresses needs of local staff for understanding state of the network
    - Would this information be useful to external users?
    - Can these tools function on a multi-domain basis?
- Beyond passive methods, there are active tools.
  - E.g. often we want a ‘throughput’ number. Can we automate that idea?
  - Wouldn’t it be nice to get some sort of plot of performance over the course of a day? Week? Year? Multiple endpoints?

Where is the “Measurement Middleware”? Something to allow for the easy exchange of metrics that are collected locally, on a global scale?

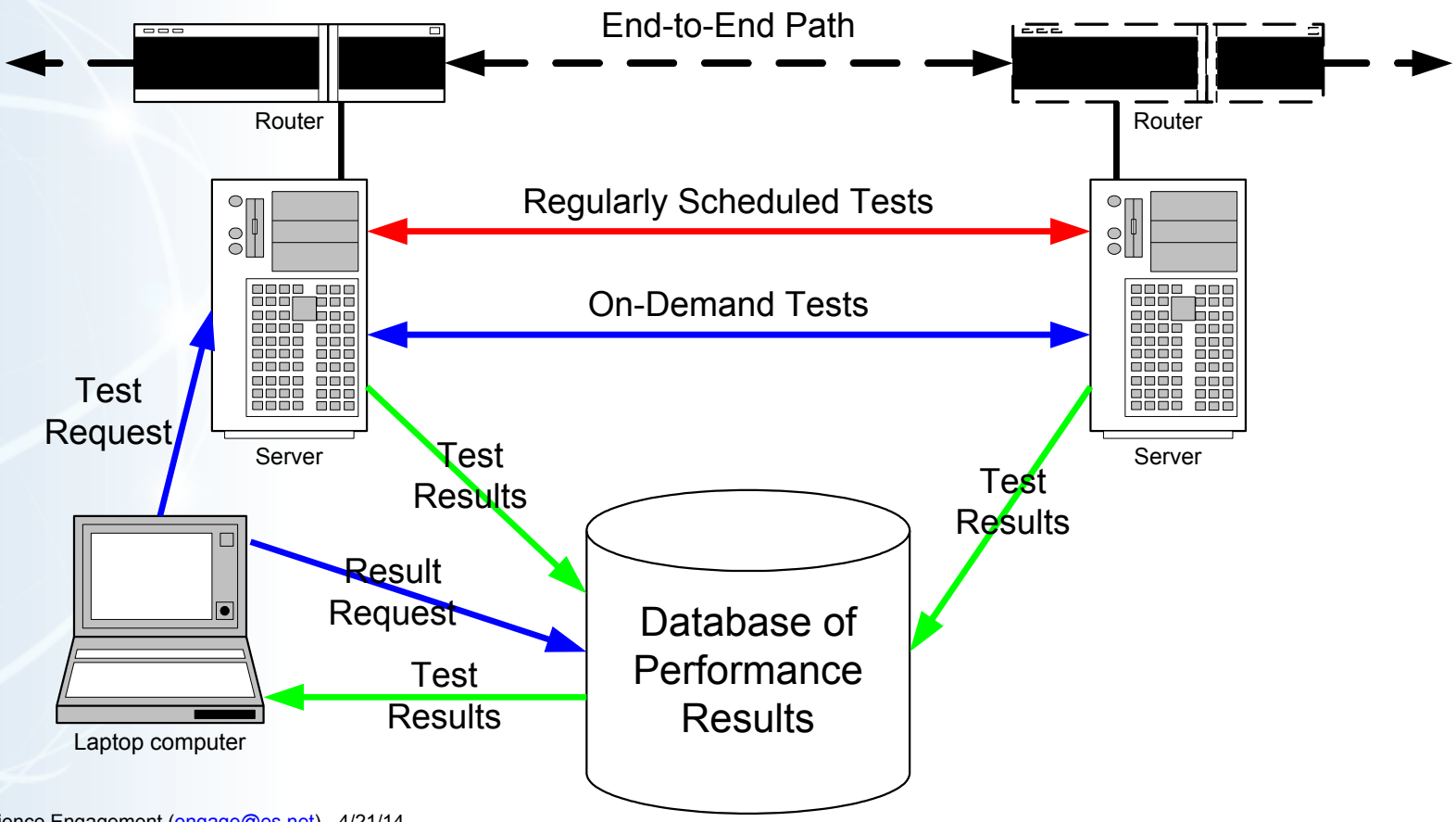
perfSONAR powered

Overview

- Introduction
- The Ghost of perfSONAR Past
- perfSONAR Present
- Use Cases
- Future Directions & Unfinished Business

The Ghost of perfSONAR Past

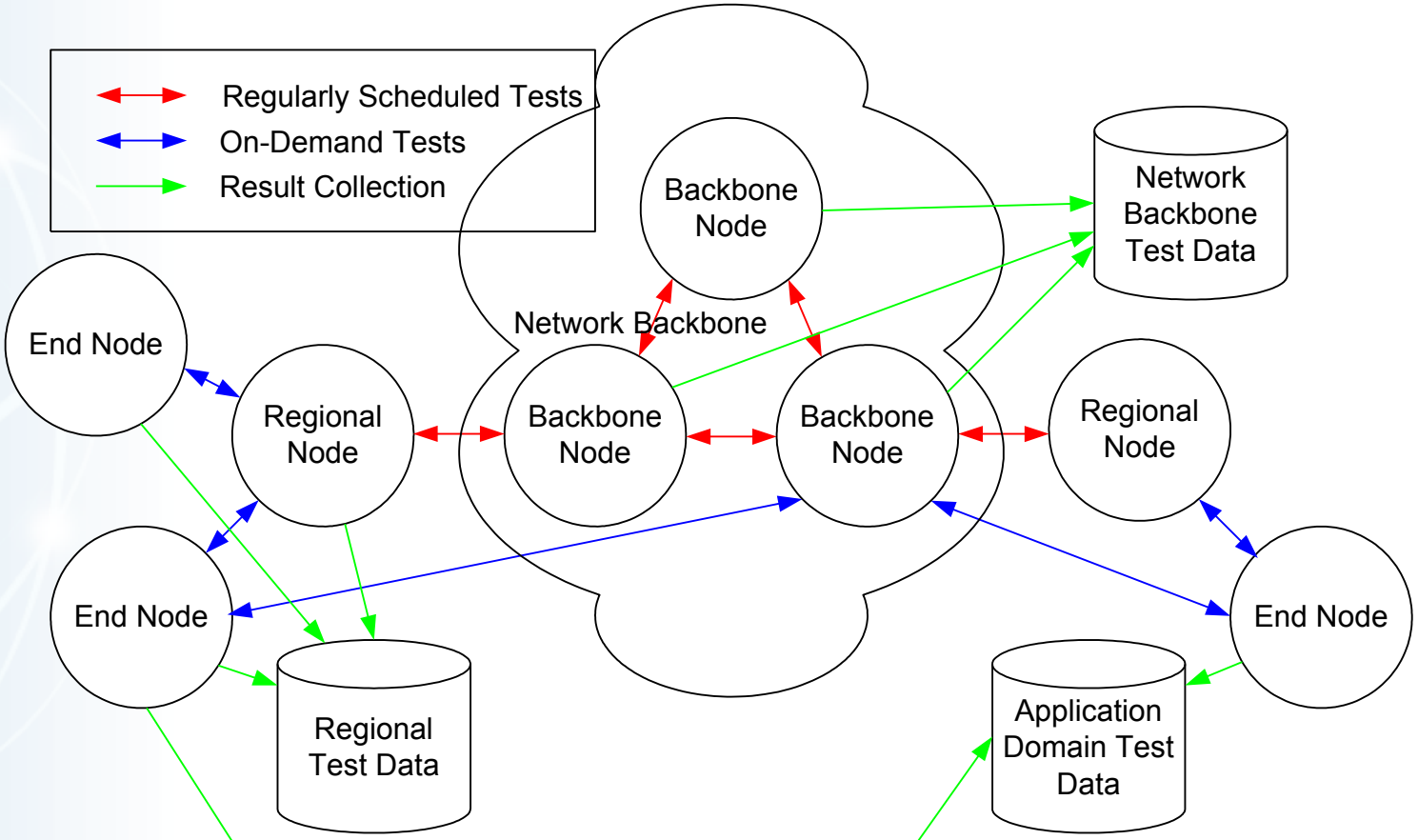
- Internet2 Performance Evaluation and Review Framework (PERF) - ~2002



perfSONAR powered

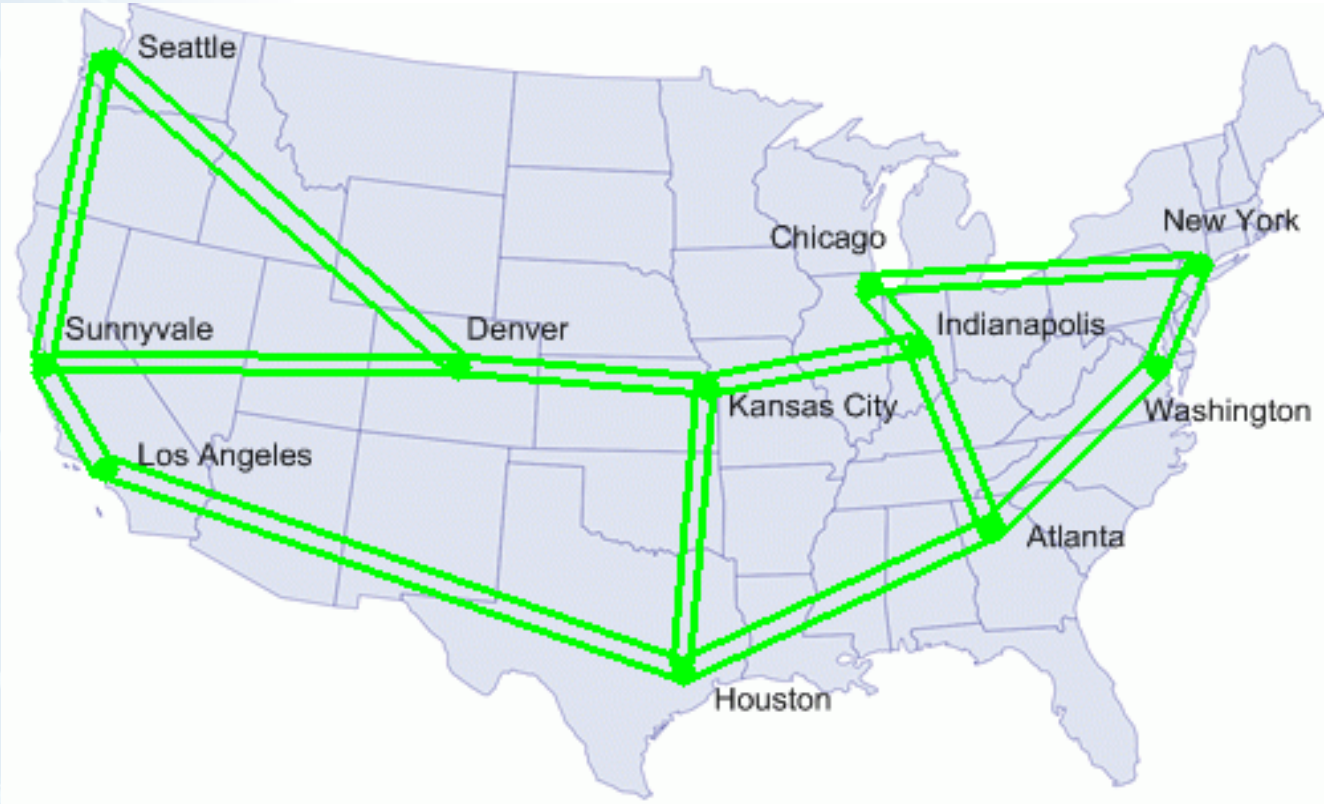
The Ghost of perfSONAR Past

- Internet2 Performance Evaluation and Review Framework (PERF) - ~2002



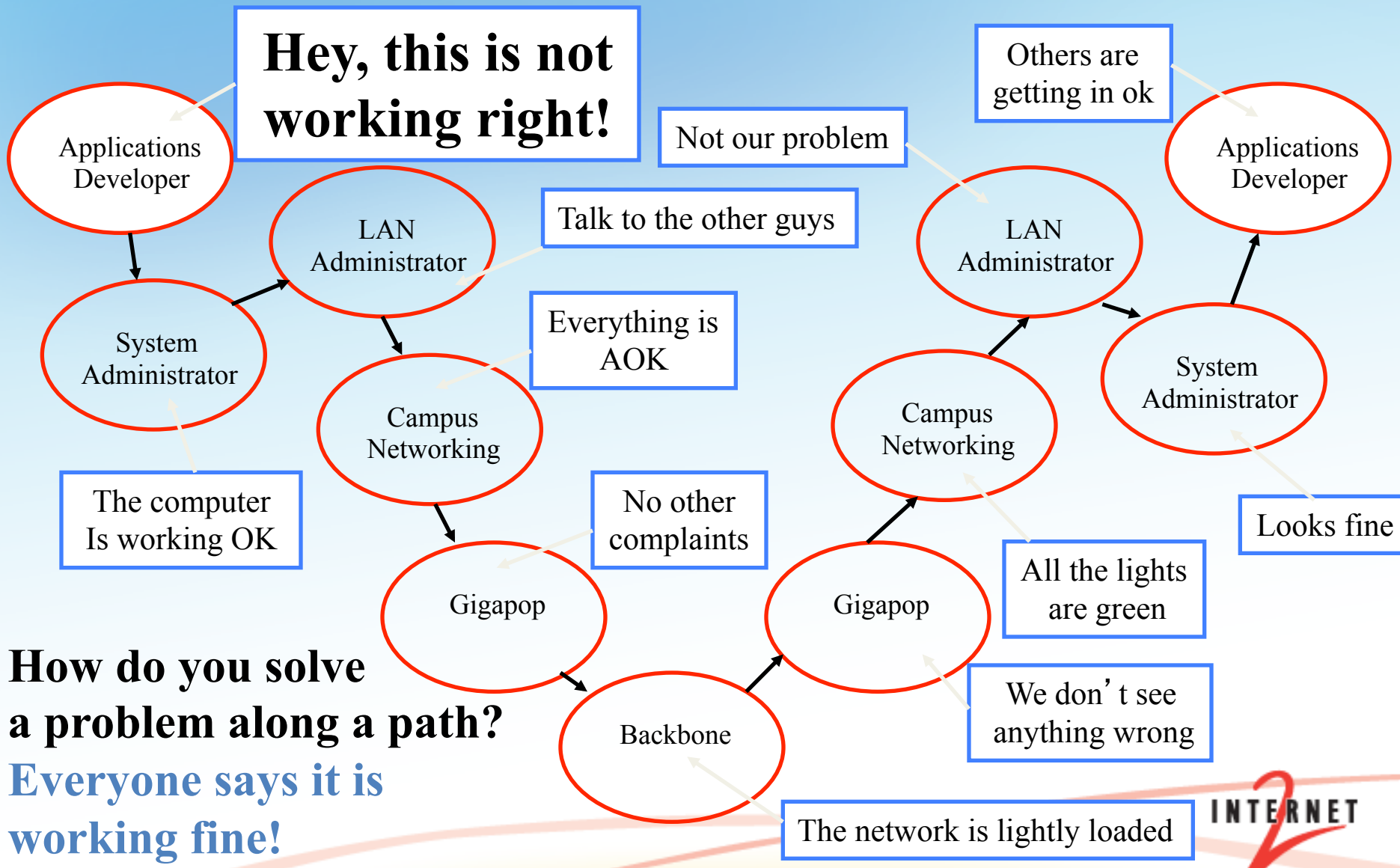
The Ghost of perfSONAR Past

- Internet2 E2E piPEs (End-to-End Performance Initiatives Performance Environment System) ~2003/2004



perfSONAR powered

Problem: Nobody's Fault



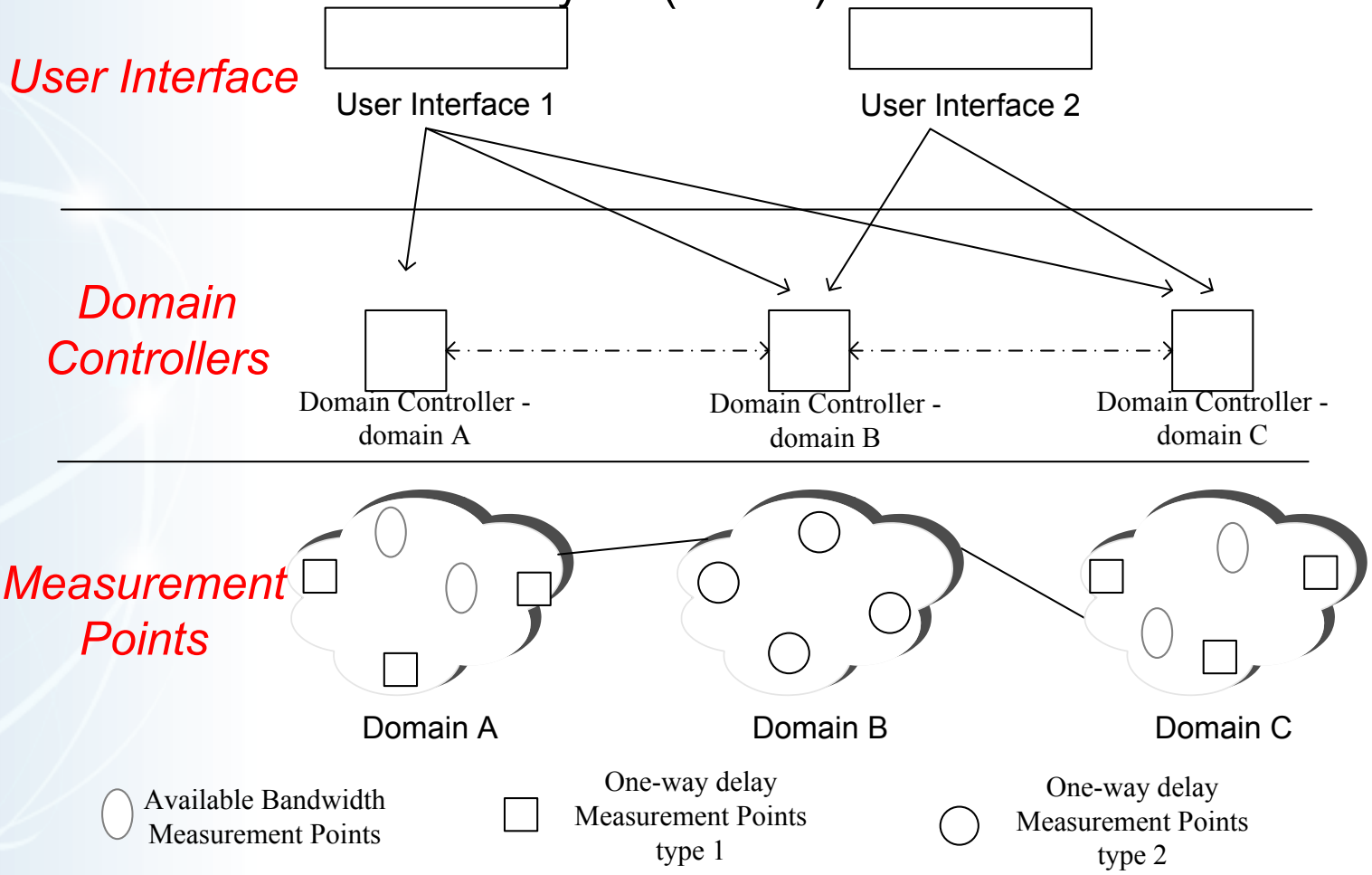
How do you solve a problem along a path?

Everyone says it is working fine!

INTERNET 2

The Ghost of perfSONAR Past

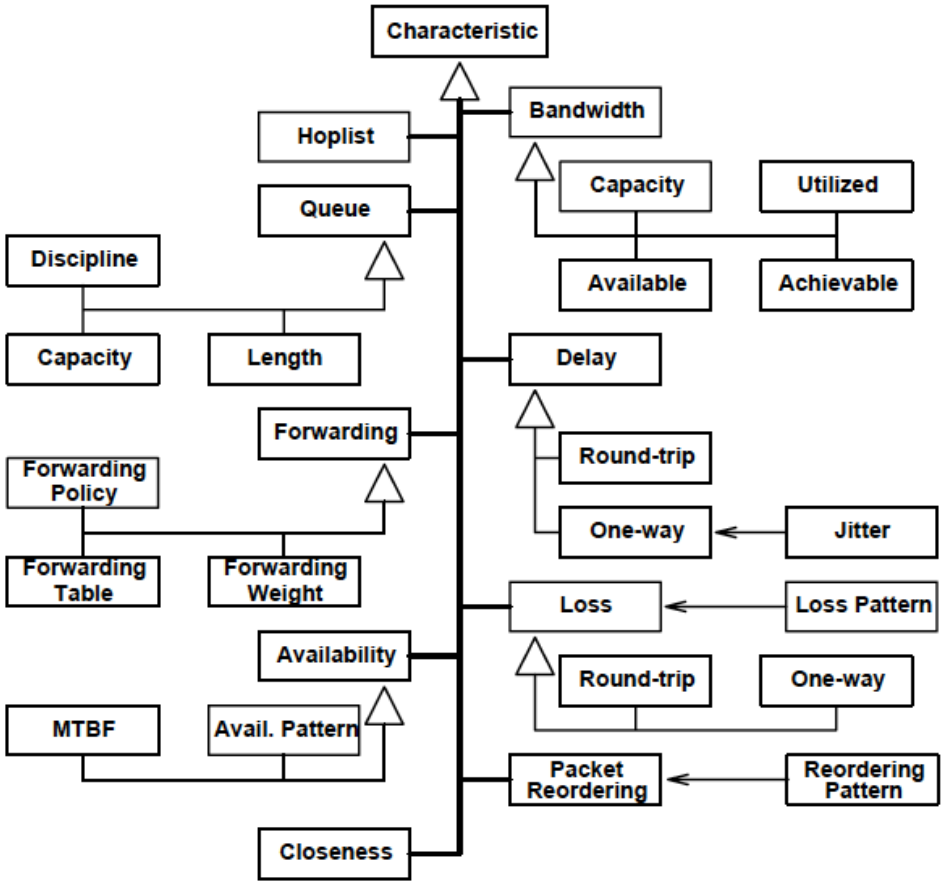
- GEANT2/JRA1 Framework Layers (~2004)



perfSONAR powered

The Ghost of perfSONAR Past

- Global Grid Forum (~2003)
  - Hierarchy of Network Performance Characteristics
    - http://www.ogf.org/documents/GFD.23.pdf
  - Request Schema Requirements and Sample Implementation
  - Report Schema Requirements and Sample Implementation



The Ghost of perfSONAR Past

- ~2004/2005
  - Merging the efforts of GGF/OGF, Internet2 E2E piPES, GN2/JRA1
    - "SONAR": You could call it piPEs v2.0 (if you're American) or GFD (if you're European)
    - A Services-Based Measurement Framework for Building Dynamic, Self-Organizing Performance Communities
  - Added additional partners from ESnet and RNP
  - First release in ~2005
  - pS/MDM Offerings in ~2007
  - pS Performance Toolkit ~2009
  - ~1000 Deployed Instances ~2013

Overview

- Introduction
- The Ghost of perfSONAR Past
- perfSONAR Present
- Use Cases
- Future Directions & Unfinished Business

What is perfSONAR?

perfSONAR is a tool to:

- Set network performance expectations
- Find network problems (“soft failures”)
- Help fix these problems

All in multi-domain environments

- These problems are all harder when multiple networks are involved

perfSONAR is provides a standard way to publish active and passive monitoring data

- This data is interesting to network researchers as well as network operators

perfSONAR powered

perfSONAR Toolkit

The “perfSONAR Toolkit” is an open source implementation and packaging of the perfSONAR measurement infrastructure and protocols from ESnet and Internet2

http://psps.perfsonar.net

All components are available as RPMs, and bundled into a CentOS 6-based “netinstall” and a “Live CD”

- perfSONAR tools are much more accurate if run on a dedicated perfSONAR host, not on the DTN

Very easy to install and configure

- Usually takes less than 30 minutes

perfSONAR Toolkit Services

PS-Toolkit includes these measurement tools:

- BWCTL: network throughput
- OWAMP: network loss, delay, and jitter
- traceroute

Test scheduler:

- runs bwctl, traceroute, and owamp tests on a regular interval

Measurement Archives (data publication)

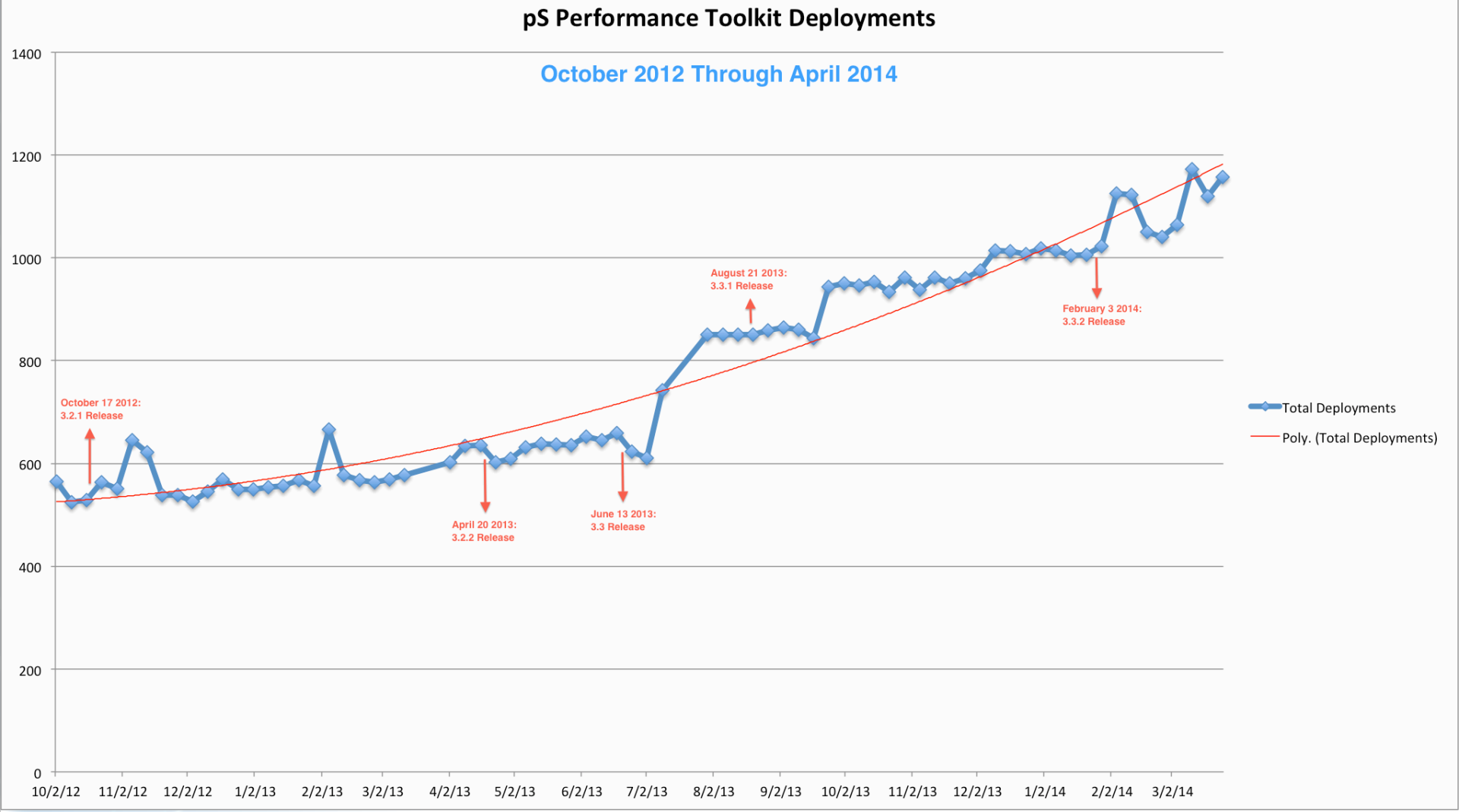
- SNMP MA – router interface Data
- pSB MA -- results of bwctl, owamp, and traceroute tests

Lookup Service: used to find services

PS-Toolkit includes these web100-based Troubleshooting Tools

- NDT (TCP analysis, duplex mismatch, etc.)
- NPAD (TCP analysis, router queuing analysis, etc)

perfSONAR Present



Lookup Service Directory Search: http://stats.es.net/ServicesDirectory/

perfSONAR

perfSONAR Global Service and Data View

Browser

Communities Filter:

Select one or more communities to refine results.

10G
AARNet
ACORN
ACORN-NS
AGLT2
ALICE

Text Filter:

Further refine results by text matching across multiple fields.

Filter

Showing: 4920 of 4920 services

- BWCTL Server 658
- MA 916
- NDT Server 710
- NPAD Server 569
- OWAMP Server 650
- phoebus 4
- Ping Responder 608
- Traceroute Responder 805

Service Information

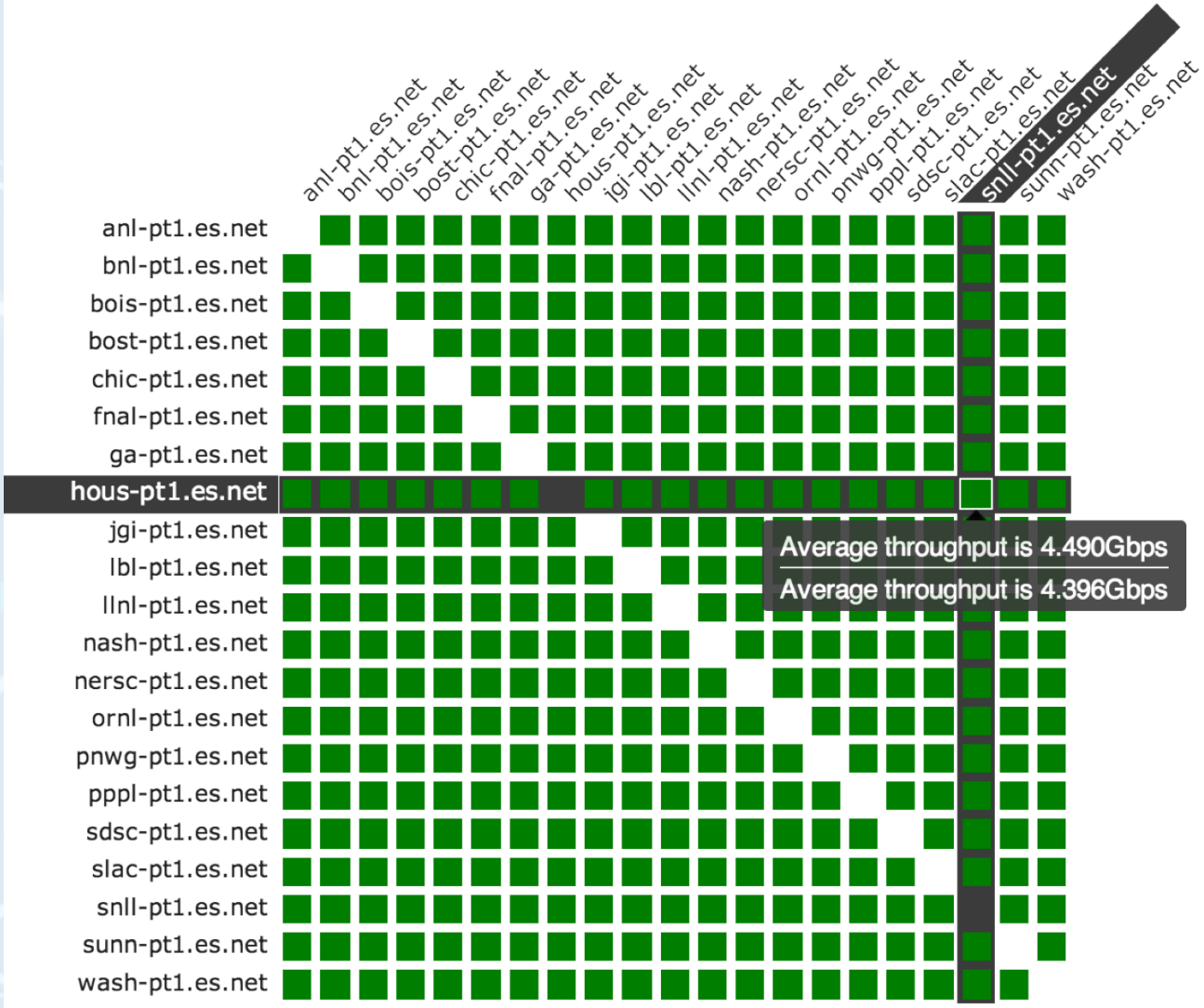
Service Name	Addresses	Geographic Location	Communities	Example Command-Line

Host Information

Host Name	Hardware	System Info	Toolkit Version	Communities

Service Map

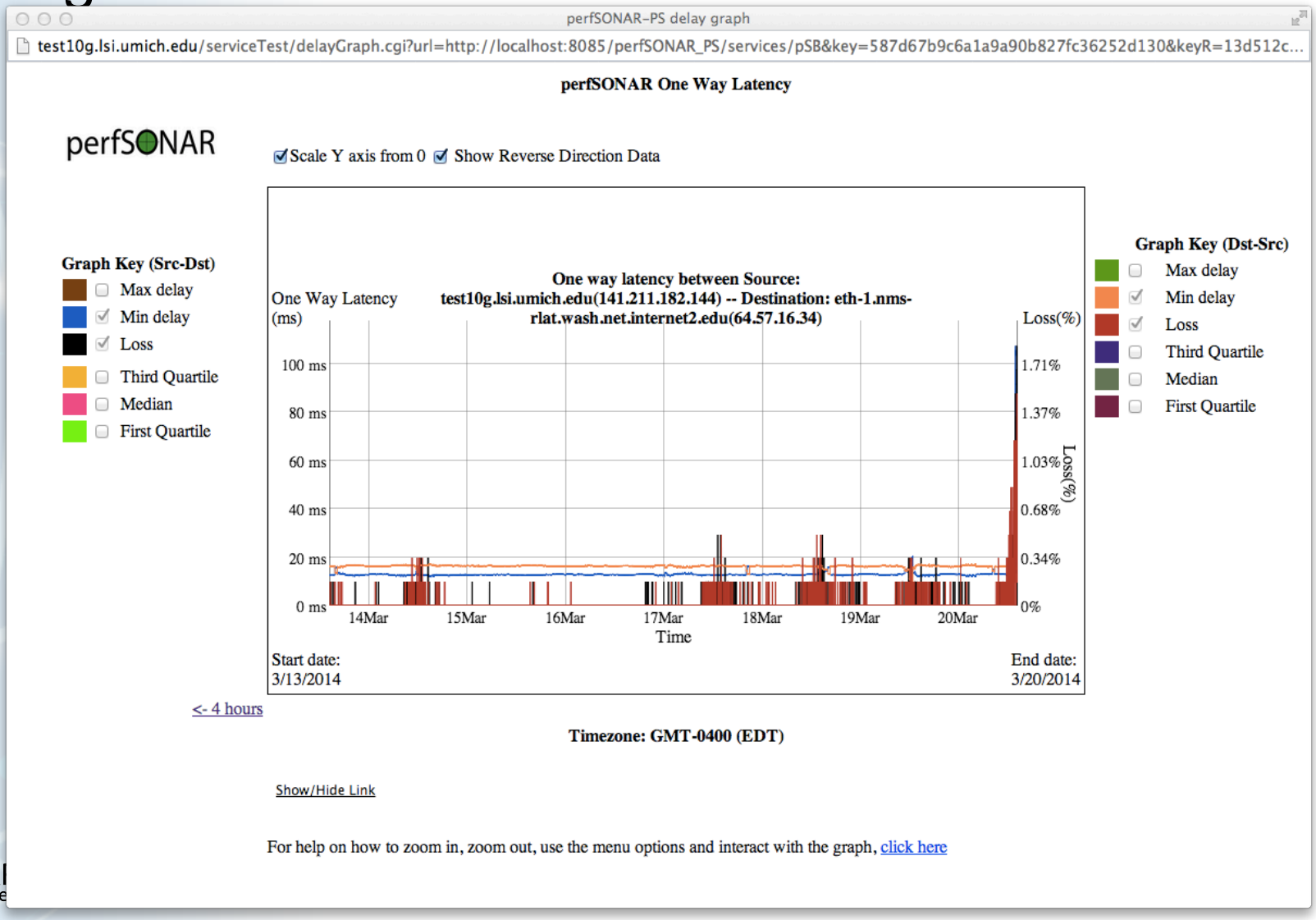
perfSONAR Dashboard: http://ps-dashboard.es.net



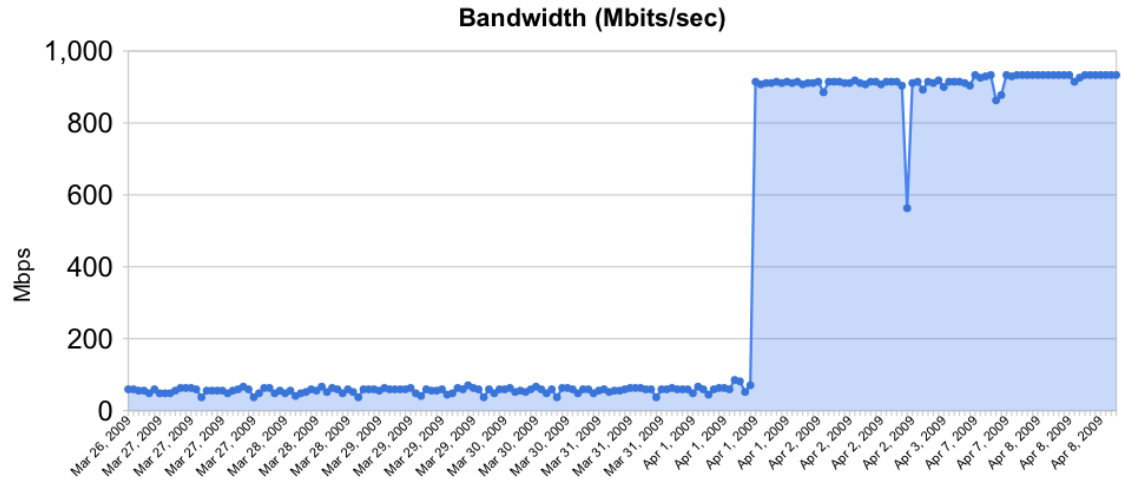
Overview

- Introduction
- The Ghost of perfSONAR Past
- perfSONAR Present
- Use Cases
- Future Directions & Unfinished Business

Congestion – via OWAMP Loss

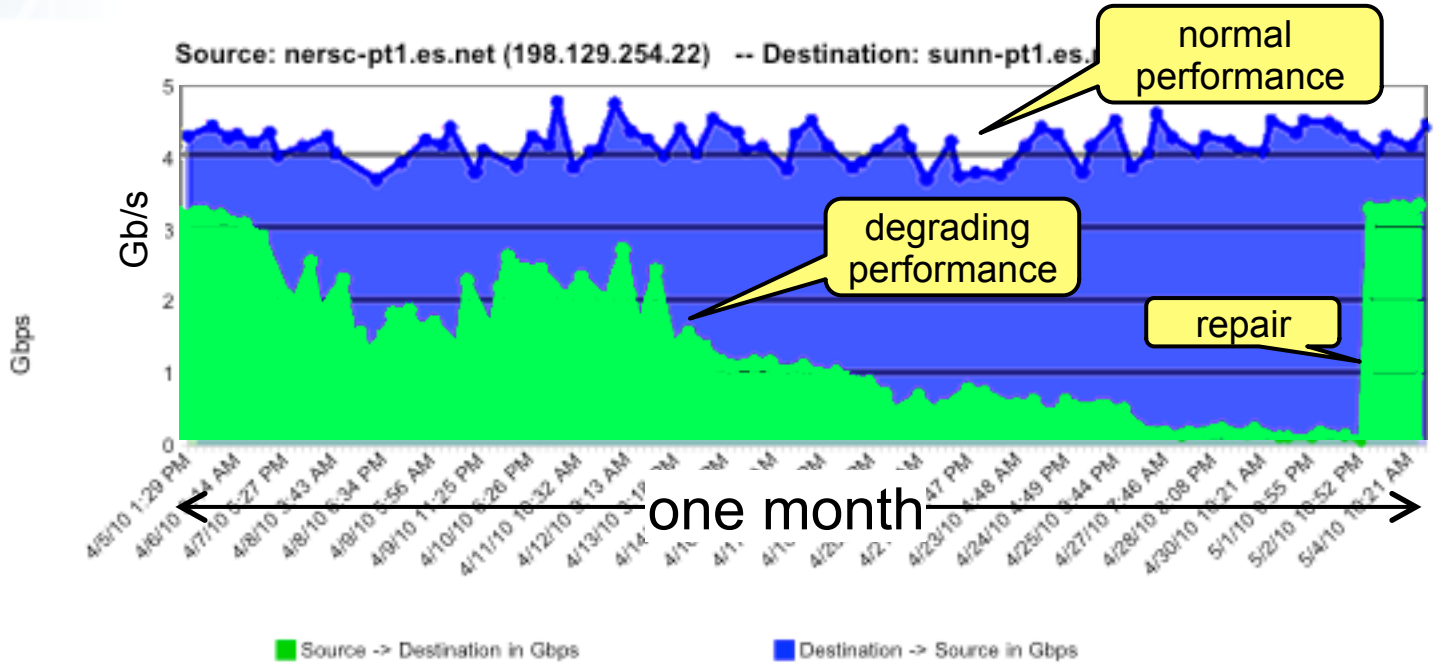


Soft Routing/Interface Failures



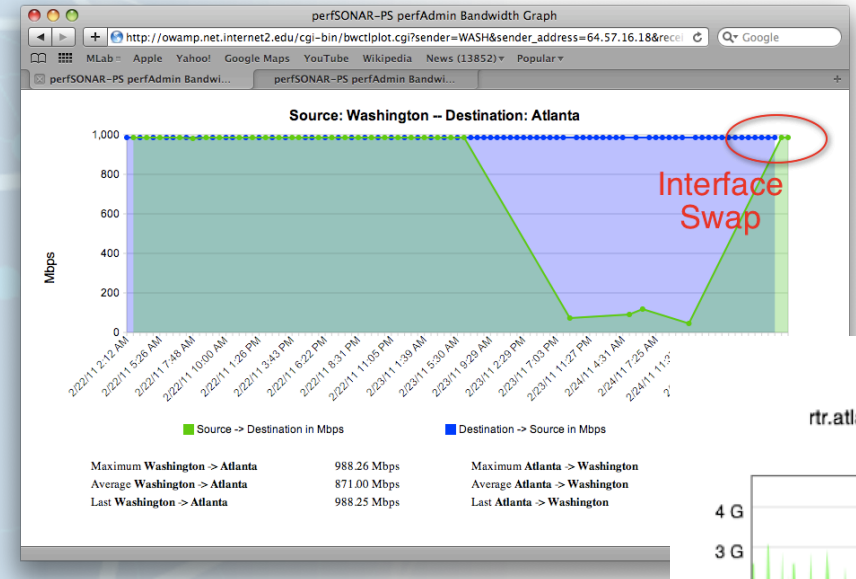
Rebooted router
with full route table

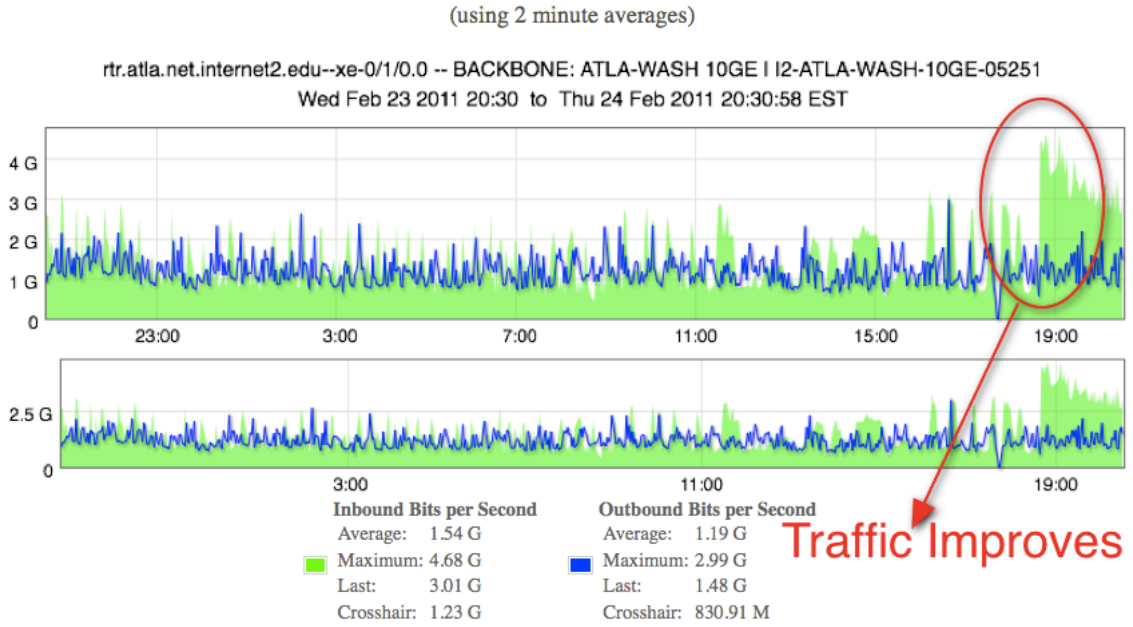
Gradual failure
of optical line
card



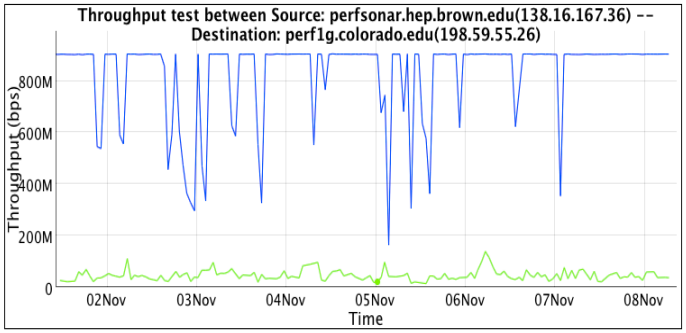
Failing Optics (BWCTL and Utilization)

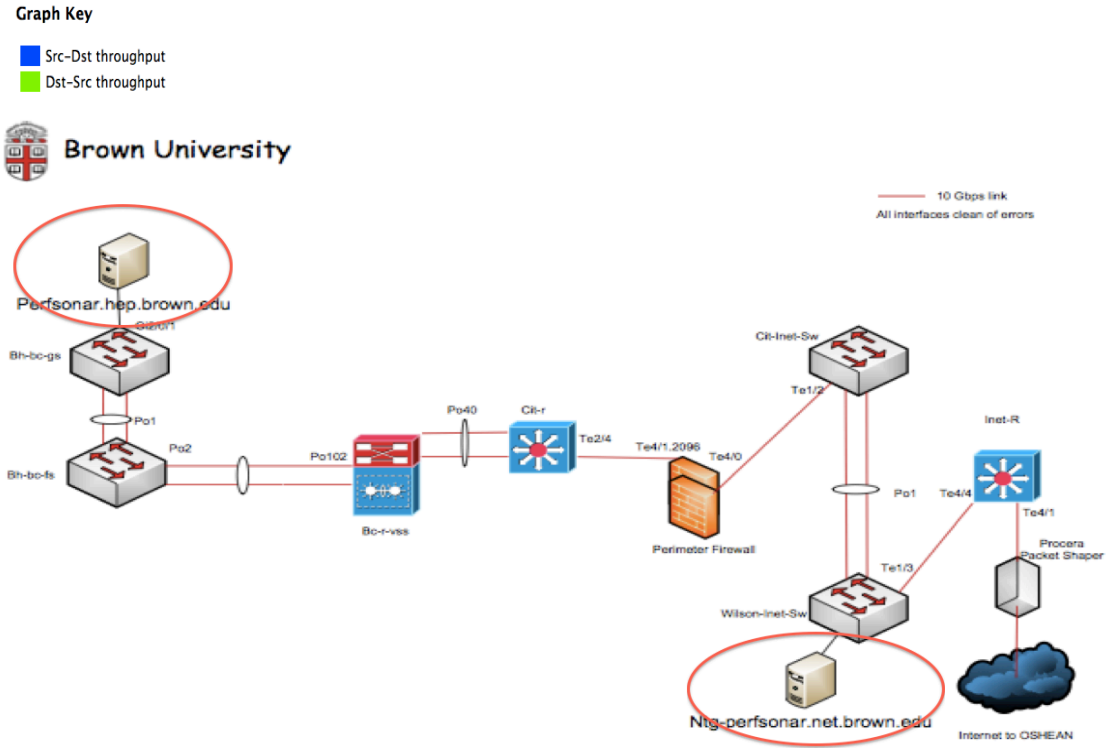
- Example taken from Internet2 backbone



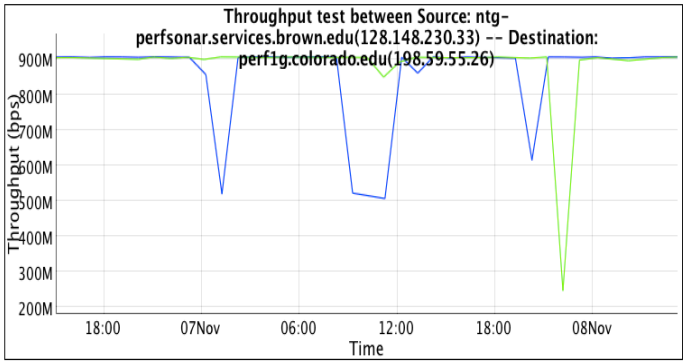


Firewalls





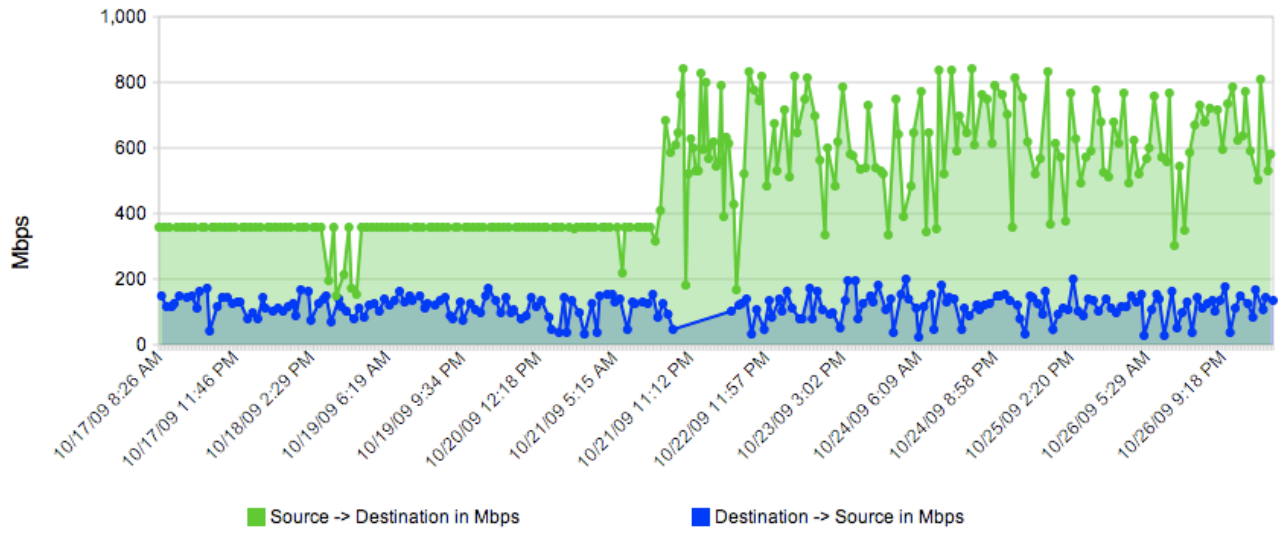
- When used as a comparison tool – we can see that security devices are impacting performance
- Powerful incentive to fix this in the scientific path



perfSONAR powered

Host Tuning

- Simple example – play with the settings in /etc/sysctl.conf when running some BWCTL tests.
- See if you can pick out when we raised the memory for the TCP window



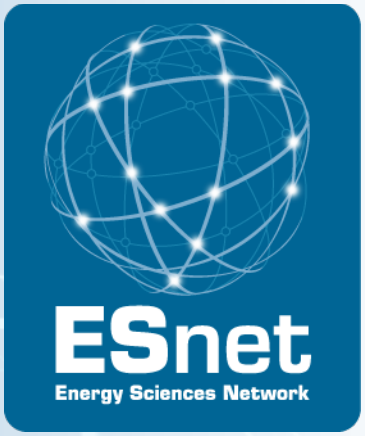
perfSONAR powered

Overview

- Introduction
- The Ghost of perfSONAR Past
- perfSONAR Present
- Use Cases
- Future Directions & Unfinished Business

Future Directions & Unfinished Business

- Deployment
  - Installation via ISO/Live CD was the greatest catalyst for deployment. This brought measurement tools to the unsophisticated user
  - What does perfSONAR on a small appliance (cubox, Intel NUC, Raspberry PI) look like, and what do we need to do to support it?
- Tools
  - There will always be new tools – the data abstraction supports sharing
  - SDN integration? How can monitoring data be used to fundamentally change networking protocols?
- Use Cases
  - Designed by engineers for engineers – except when the Scientists and CEOs discovered what it can do
  - Can a simple web page be made that allows a live tests between anything in the world?
  - Could we get the same software used on the server, integrated into the routing device?
- Adoption
  - Scaling to 1000+ instances was great – how can we scale to 10000? Greater emphasis on the project means more helpers needed



The perfSONAR Project at 10 Years: Status and Trajectory

Questions/Comments/Criticisms?

Jason Zurawski - zurawski@es.net

ESnet Science Engagement – engage@es.net

http://psps.perfsonar.net

perfSONAR powered

U.S. DEPARTMENT OF ENERGY
Office of Science

BERKELEY LAB