

perfSONAR Deployment on ESnet

Brian Tierney

ESnet

ISMA 2011 AIMS-3

Workshop on Active Internet Measurements

Feb 9, 2011

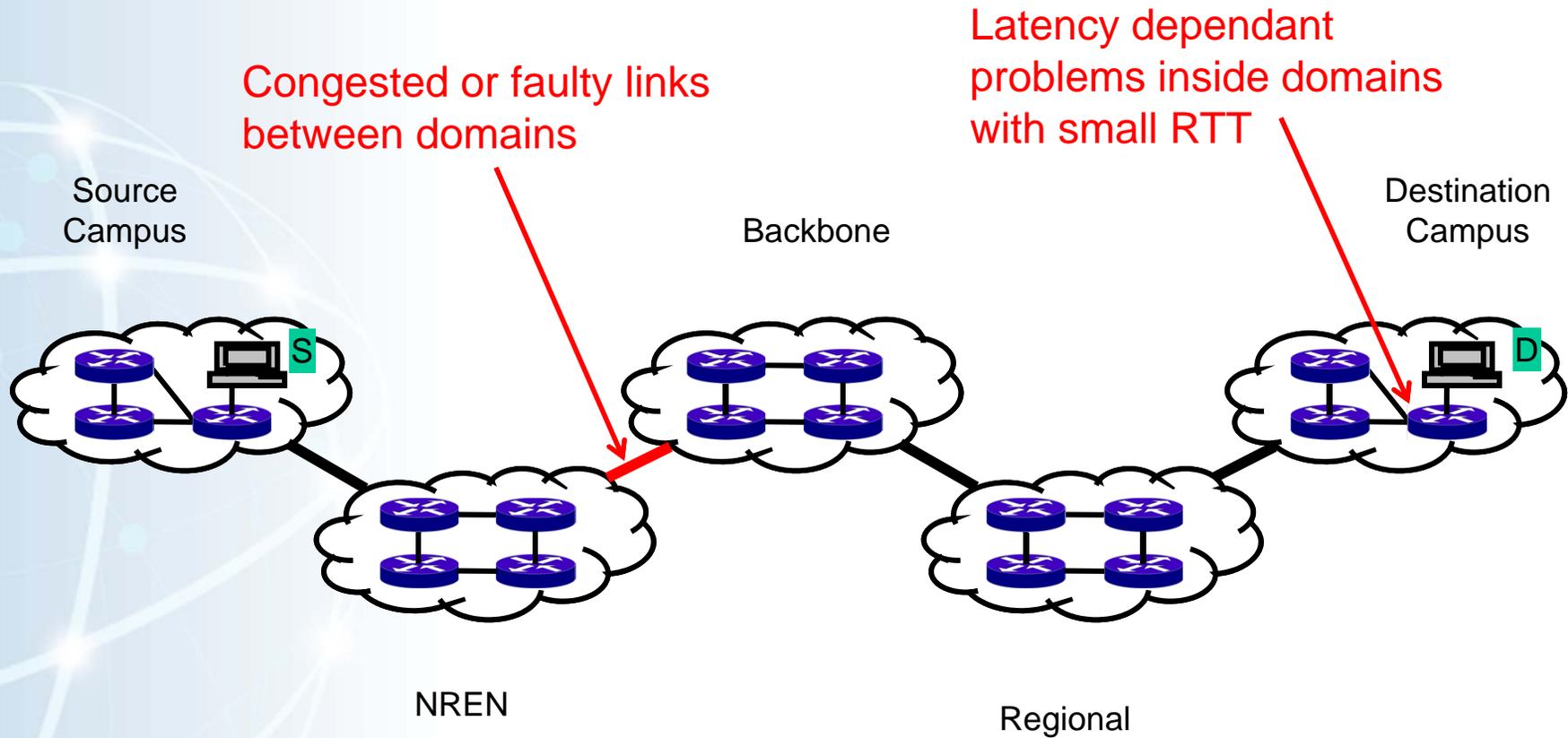


upload in www.top81.cn

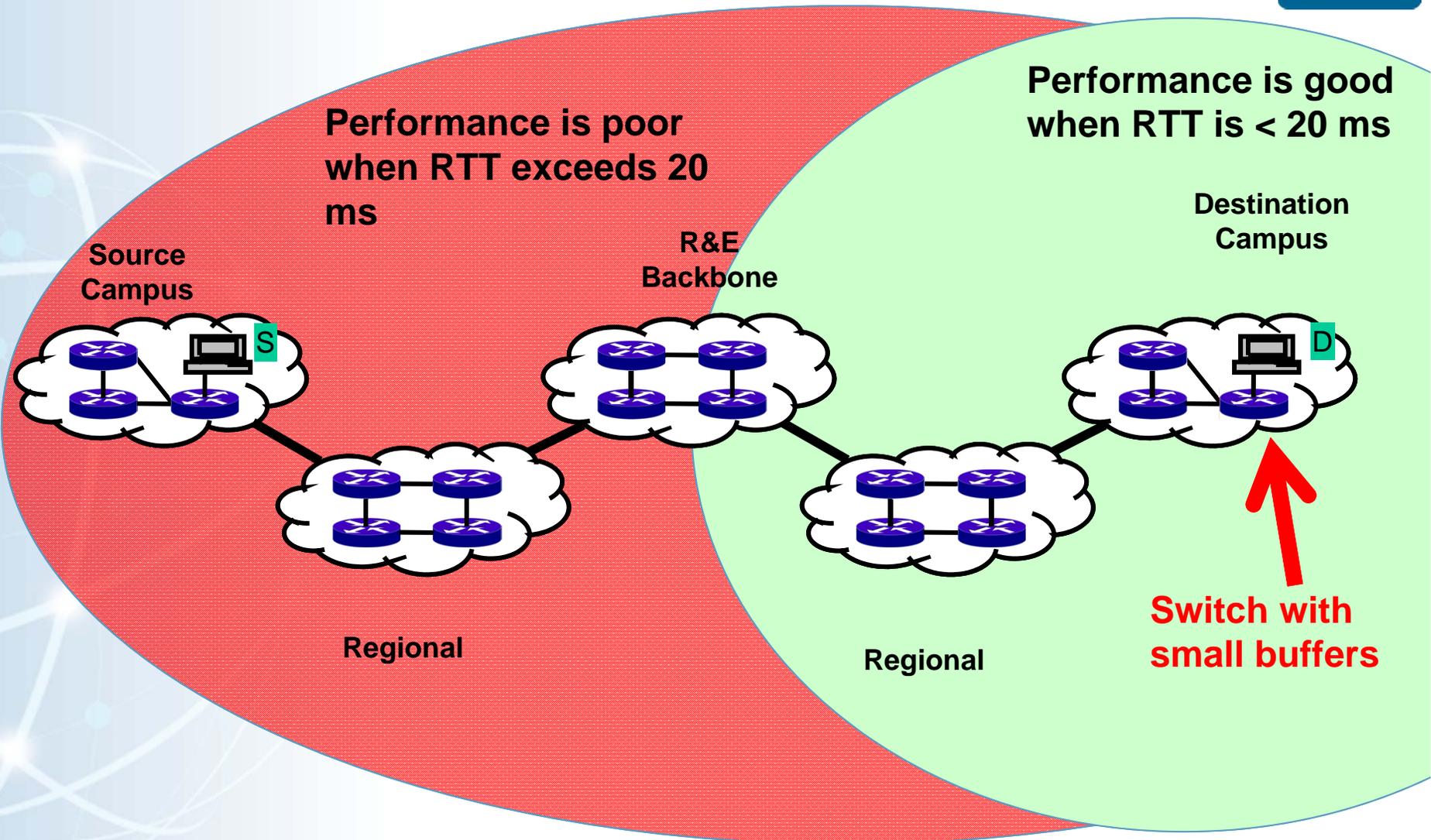
Why
does the
Network
seem so
slow?



Where are common problems?



Local testing will not find all problems





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



Common Soft Failures

Small Queue Tail Drop

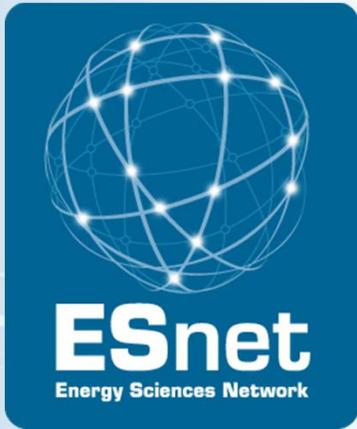
- Switches not able to handle the long packet trains prevalent in long RTT sessions and local cross traffic at the same time

Un-intentional Rate Limiting

- Processor-based switching on routers due to faults, acl's, or mis-configuration
- Security Devices
 - E.g.: 10X improvement by turning off Cisco Reflexive ACL

Random Packet Loss

- Bad fibers or connectors
- Low light levels due to amps/interfaces failing
- Duplex mismatch



Building a Global Network Diagnostic Framework



Addressing the Problem: perfSONAR

perfSONAR - an open, web-services-based framework for:

- running network tests
- collecting and publishing measurement results

ESnet is:

- Deploying the framework across the science community
- Encouraging people to deploy *'known good'* measurement points near domain boundaries
 - "known good" = hosts that are well configured, enough memory and CPU to drive the network, proper TCP tuning, clean path, etc.
- Using the framework to find and correct soft network failures.

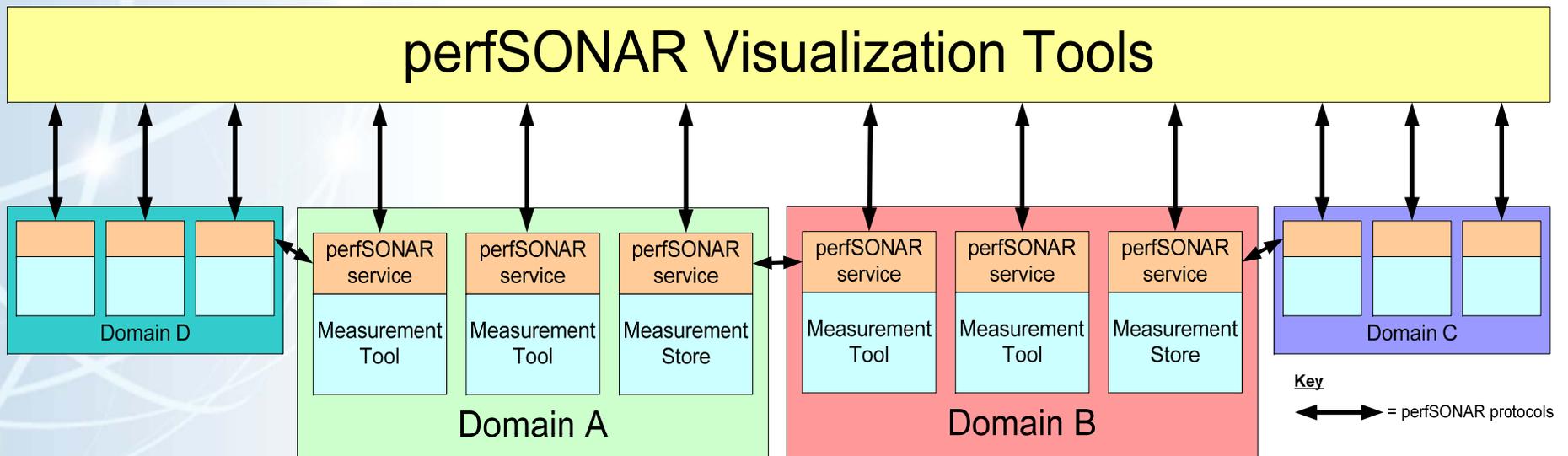
perfSONAR Architecture



The perfSONAR framework:

- Is middleware.
- Is distributed between domains.
- Facilitates inter-domain performance information sharing.

perfSONAR services 'wrap' existing measurement tools.



perfSONAR Services



Lookup Service

- gLS – Global lookup service used to find services
- hLS – Home lookup service for registering local perfSONAR metadata

Measurement Archives (data publication)

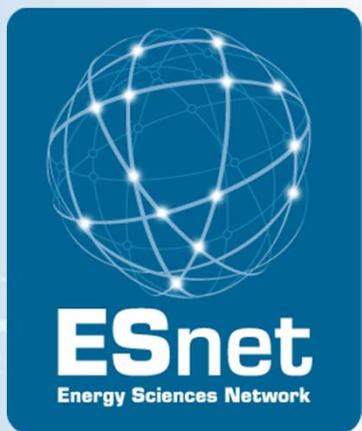
- SNMP MA – Interface Data
- pSB MA -- Scheduled bandwidth and latency data

PS-Toolkit includes these measurement tools:

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

PS-Toolkit includes these Troubleshooting Tools

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



ESNet PerfSONAR Deployment

ESnet Deployment Activities

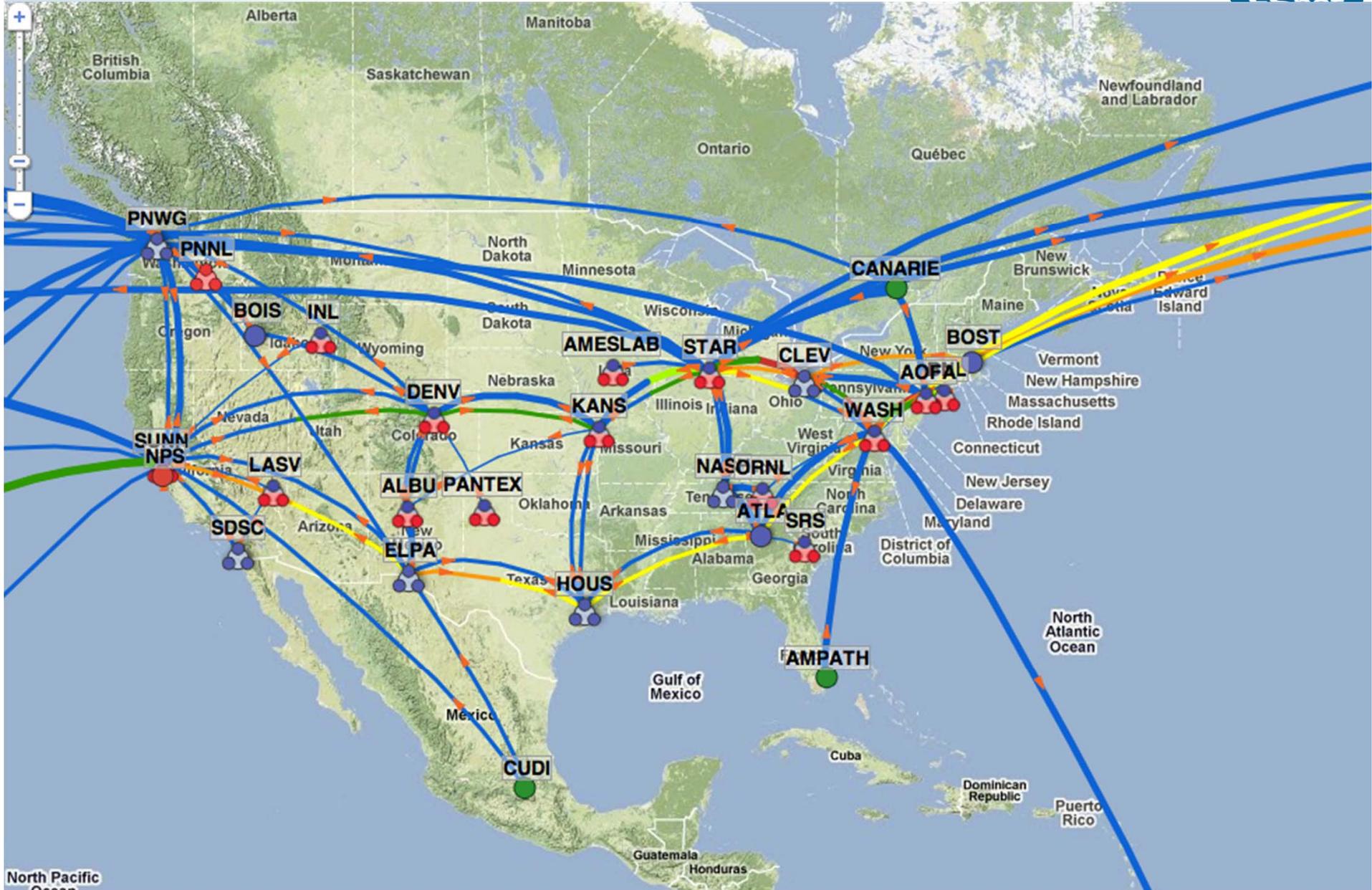


ESnet has deployed OWAMP and BWCTL servers next to all backbone routers, and at all 10Gb connected sites

- 30 locations deployed, ~20 more planned
- Full list of active services at:
 - <http://stats1.es.net/perfSONAR/directorySearch.html>
 - Instructions on using these services for network troubleshooting: <http://fasterdata.es.net>

These services have proven extremely useful to help debug a number of problems

<http://weathermap.es.net>





Global PerfSONAR-PS Deployments

Based on “global lookup service” (gLS) registration, Feb 2011: currently deployed in over 80 locations

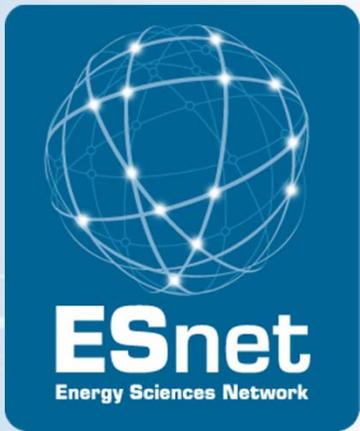
- ~ 80 bwctl and owamp servers
- ~ 125 active probe measurement archives
- ~ 20 SNMP measurement archives
- Countries include: USA, Australia, Hong Kong, Argentina, Brazil, Uruguay, Guatemala, Japan, China, Canada, Netherlands, Switzerland
- Many more deployments behind firewalls

US Atlas Deployment

- Monitoring all “Tier 1 to Tier 2” connections

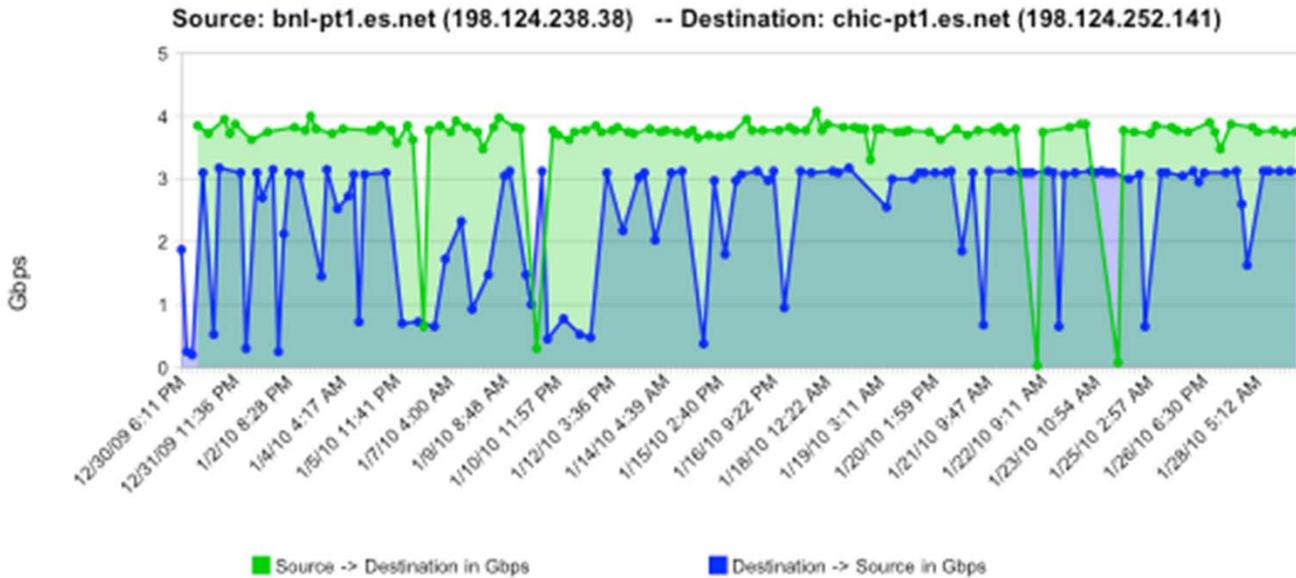
For current list of *public* services, see:

- <http://stats1.es.net/perfSONAR/directorySearch.html>



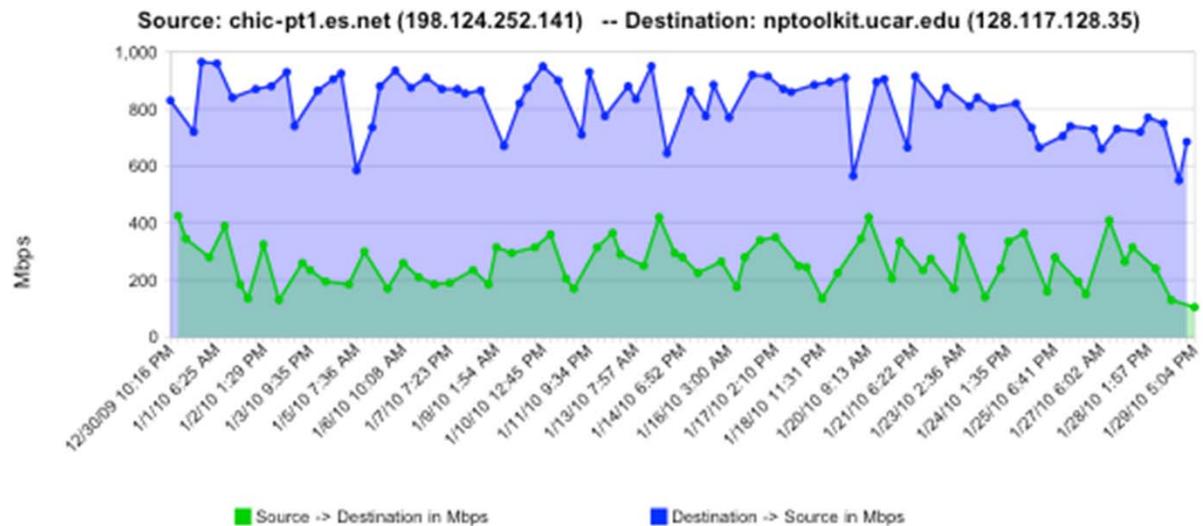
SAMPLE results

Sample Results



Heavily used path:
probe traffic is
“scavenger service”

Asymmetric
Results: different
TCP stacks?



Sample Results: Latency/Loss Data

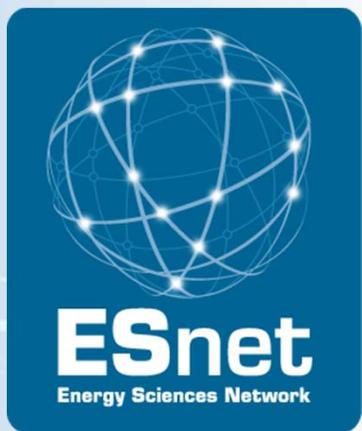


Source: ps-lat.es.net (198.129.254.187) -- Destination: bost-owamp.es.net (198.124.238.58)

One Way Delay



Timezone: PST



Network Research Using perfSONAR data

perfSONAR workshop series



ESnet and Internet2 are actively encouraging researcher use of the data we are collecting

NSF, DOE, and LSN sponsored a workshop to discuss the research uses of perfSONAR in Washington DC last summer.

- 90 attendees!

“The goal of the workshop is to use perfSONAR as a focus to cross-fertilize ideas from the network research community and the needs of the research and education networks around the world, documenting open areas and best practices.”

Workshop Website:

- <http://www.internet2.edu/workshops/perfSONAR/>

Workshop Report:

- <http://www.internet2.edu/workshops/perfSONAR/201007perfSONAR-Workshop-Report.pdf>



Accessing Archived Results

All results are stored in the perfSONAR “Measurement Archive” (MA)

- Periodic bwctl tests (throughput)
- Ongoing owamp tests (latency, loss, jitter)
- Periodic traceroute tests
- SNMP results for all router interface, including virtual interfaces
- ESnet topology

All results are publically accessible

Simple Web-service model

Easy to use Perl API to query for results

See: <http://fasterdata.es.net/fasterdata/perfSONAR/client-api/>

Sample Project: Malathi Veeraraghavan, Univ of Virginia



One-way Active Measurement Protocol(OWAMP)

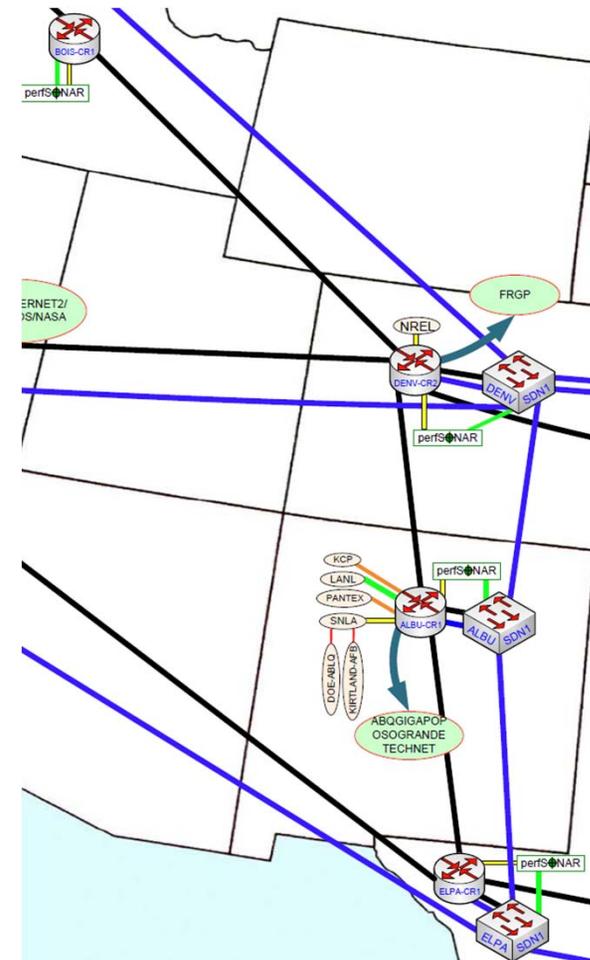
Packet interval: 0.1 sec

- 10 packets per sec
- 600 packets per minute

Use perl programs provided by
perfSONAR

Sample columns of the OWAMP
data file:

- endTime, loss, maxError, max_delay
min_delay, sent startTime
- one report per minute



Zhenzhen Yan and M. Veeraraghavan, University of Virginia

Sample Results: PerfSONAR OWAMP data analysis

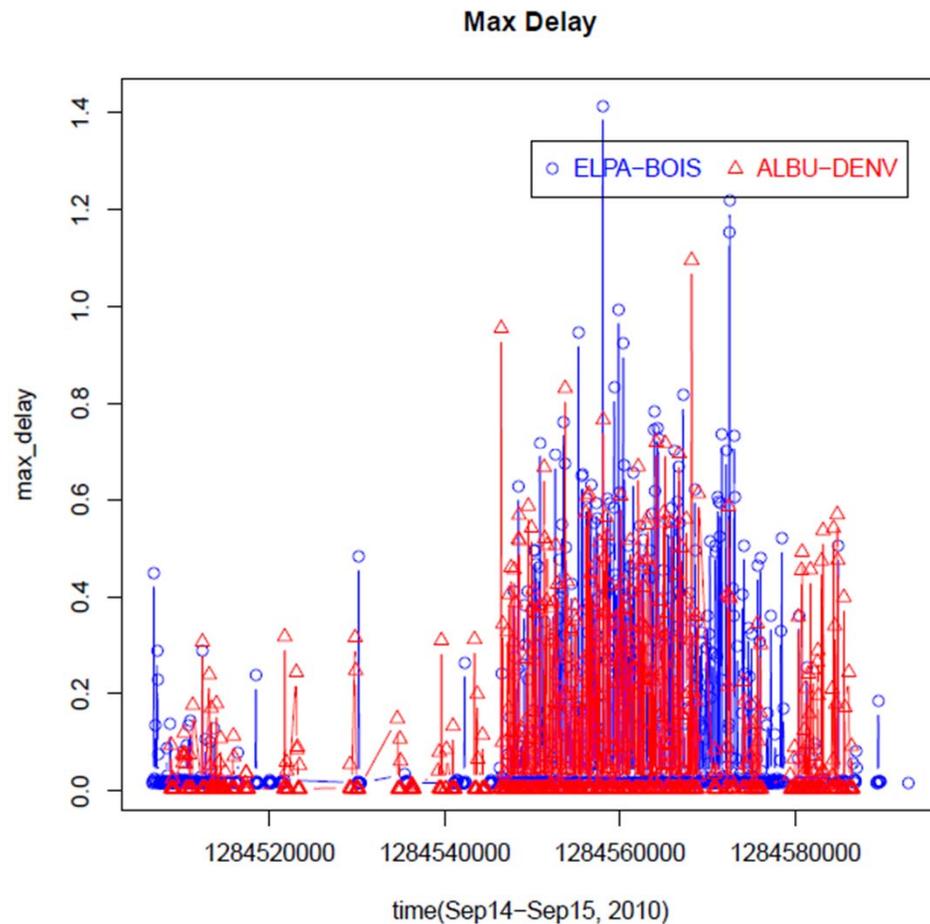


Max delay plot:

- ELPA-BOIS
- ALBU-DENV

Overlapping paths

Data traffic not host issues?



Zhenzhen Yan and M. Veeraraghavan, University of Virginia

Sample Results: Dependence on day of week



	IQR (max-delay) in sec	
Day of week	SUNN-BOST (min-delay = 0.036)	KANS-CHIC (min-delay = 0.005)
Sunday	0.08876	0.077011
Monday	0.12059	0.136785
Tuesday	0.10407	0.128747
Wednesday	0.11138	0.091315
Thursday	0.12504	0.231436
Friday	0.13171	0.128005
Saturday	0.10733	0.198049

Another Sample Project: Constantine Dovrolis, Georgia Tech



Pythia: Detection, Localization and Diagnosis of Performance Problems using perfSONAR (DOE-funded)

Pythia will be a data-analysis tool

- Processing data collected from PerfSONAR (owamp)
- Focusing on performance problems

Detection:

- “noticeable lossrate between ORNL and AARNet at 10:54:02 GMT”

Localization:

- “it happened at PNW-AARnet link”

Diagnosis:

- “it was a high-loss event due to insufficient router buffering”



How to Participate

Deploy perfSONAR!

- Using the “NP Toolkit” takes < 15 minutes to configure

Use perfSONAR to find & correct the hidden performance problems in your networks.

Help write analysis and visualization tools

- There is a huge amount of data publicly available ready to be mined
- E.g.: look for correlations between active probes and passively collected SNMP data



More Information

Information on downloading/installing perfSONAR

- <http://psps.perfsonar.net/>
- <http://fasterdata.es.net/fasterdata/perfSONAR/>

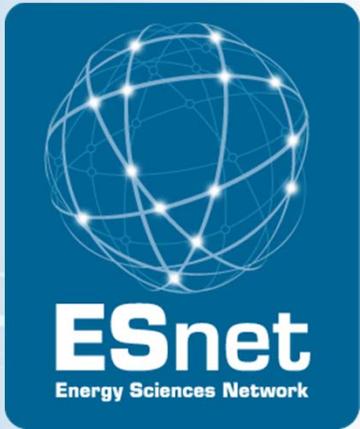
Plot ESnet perfSONAR data:

- <http://stats1.es.net/>

perfSONAR Client API:

- <http://fasterdata.es.net/fasterdata/perfSONAR/client-api/>

email: BLTierney@es.net



Extra Slides



Components of a Global Diagnostic Service

Globally accessible measurement services

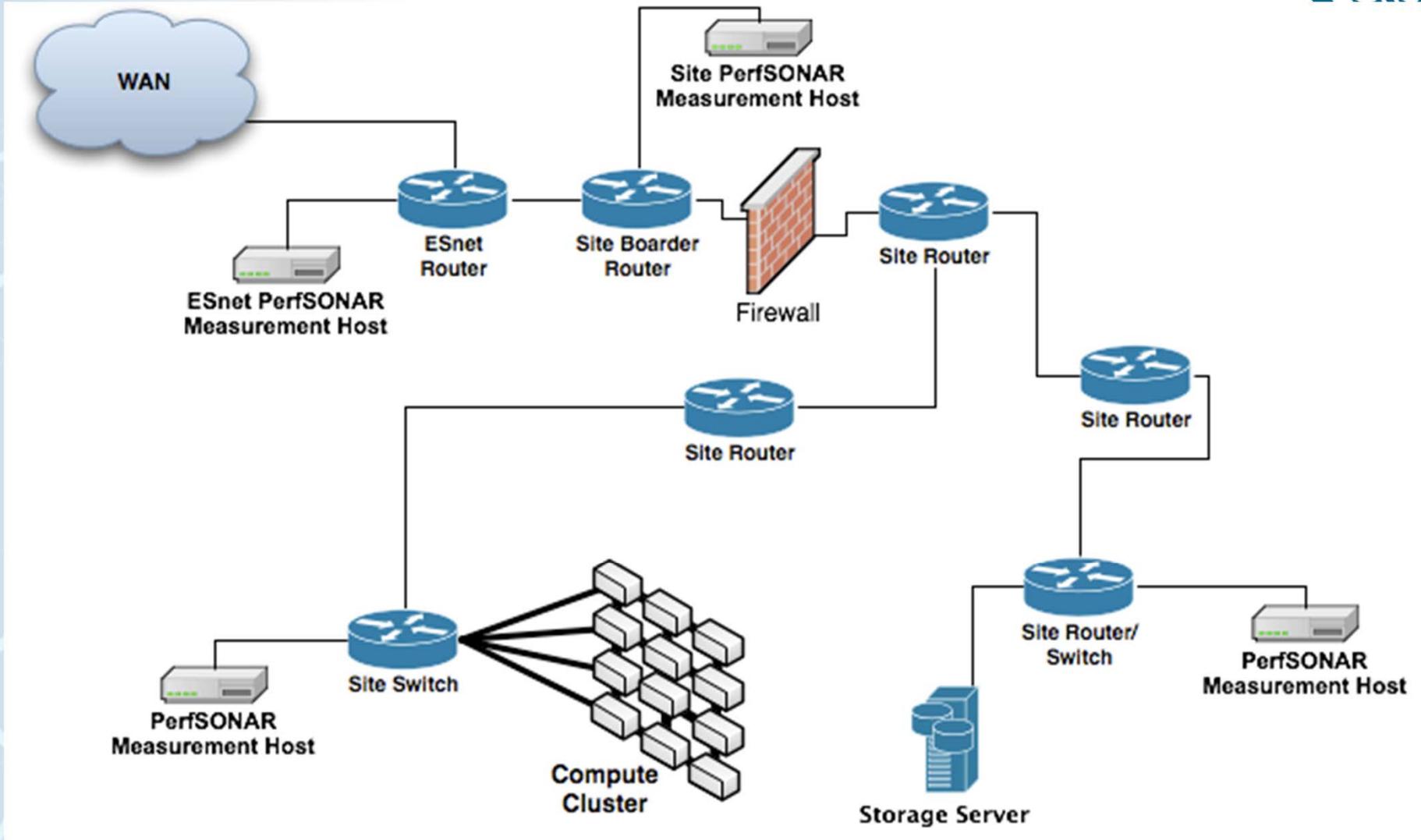
- Support for both active probes and passive results (SNMP)
- In particular: throughput testing servers
 - Recommended tool for this is bwctl
 - <http://www.internet2.edu/performance/bwctl/>
 - Includes controls to prevent DDOS attacks

Services must be registered in a globally accessible lookup service

Open to the entire R&E network community

- Ideally using light-weight authentication and authorization

Typical Campus Deployment





Providing Diagnostic Services to ESnet Users

Support adhoc network measurements for troubleshooting and infrastructure verification.

- By ESnet Staff within the backbone
- By ESnet sites to the backbone
- By ESnet **customers** across the backbone

Support regularly scheduled tests for diagnosing problems, demonstrating capabilities and monitoring.

- Between ESnet POPS
- Between ESnet sites & peers to ESnet POPS
- From sites across ESnet.

Visualization

- Allow ESnet user community to better understand our network & it's capabilities.
- Allow ESnet users to understand how their use impacts the backbone.

Alarming

- Automated analysis of regularly scheduled measurements to raise alerts.