What is perfSONAR?

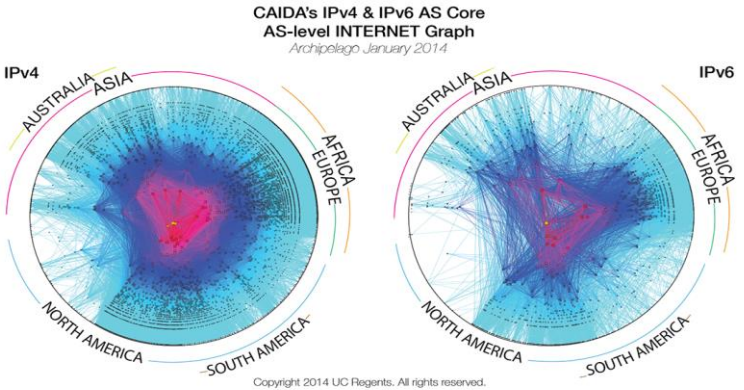
Doug Southworth, IN@IU, dojosout@iu.edu
Scott Chevalier, IN@IU, schevali@iu.edu

Outline

- Problem Statement on Network Connectivity
- Supporting Scientific Users
- Network Performance & TCP Behaviors w/ Packet Loss
- What is perfSONAR
- Deployment Overview
- Conclusions

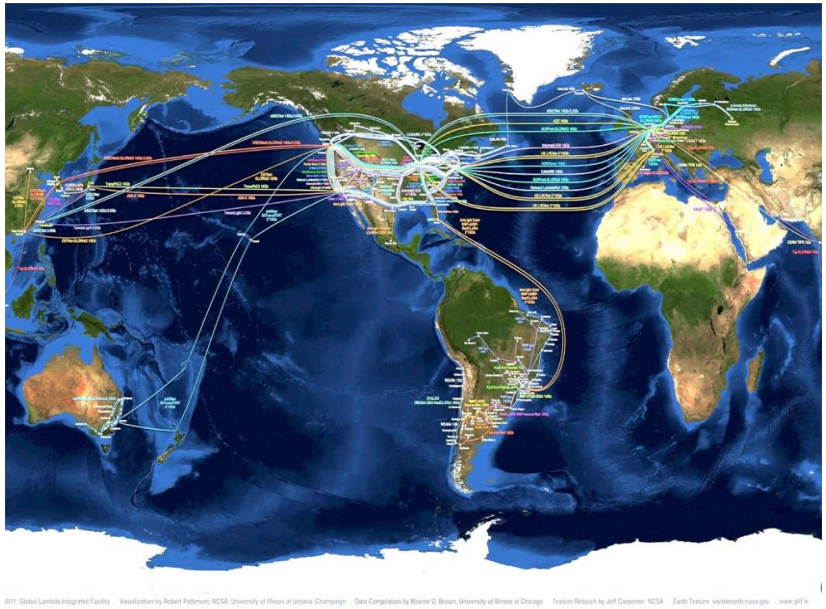
Problem Statement

- The global Research & Education network ecosystem is comprised of hundreds of international, national, regional and local-scale networks.



Problem Statement

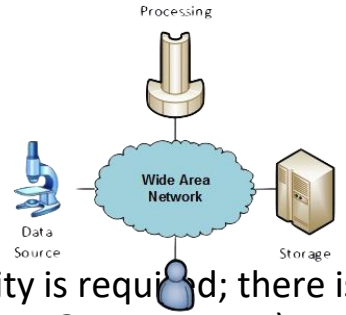
- While these networks all interconnect, each network is owned and operated by separate organizations (called domains) with different policies, customers, funding models, hardware, bandwidth and configurations.

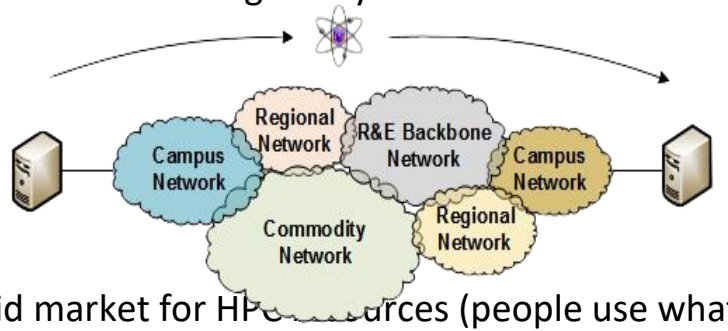


ESnet ENERGY SCIENCES NETWORK GÉANT INDIANA UNIVERSITY INTERNET2 UNIVERSITY OF MICHIGAN

The R&E Community

- The global Research & Education network ecosystem is comprised of hundreds of international, national, regional and local-scale resources – each independently owned and operated.
- This complex, heterogeneous set of networks ***must*** operate seamlessly from "end to end" to support science and research collaborations that are distributed globally.





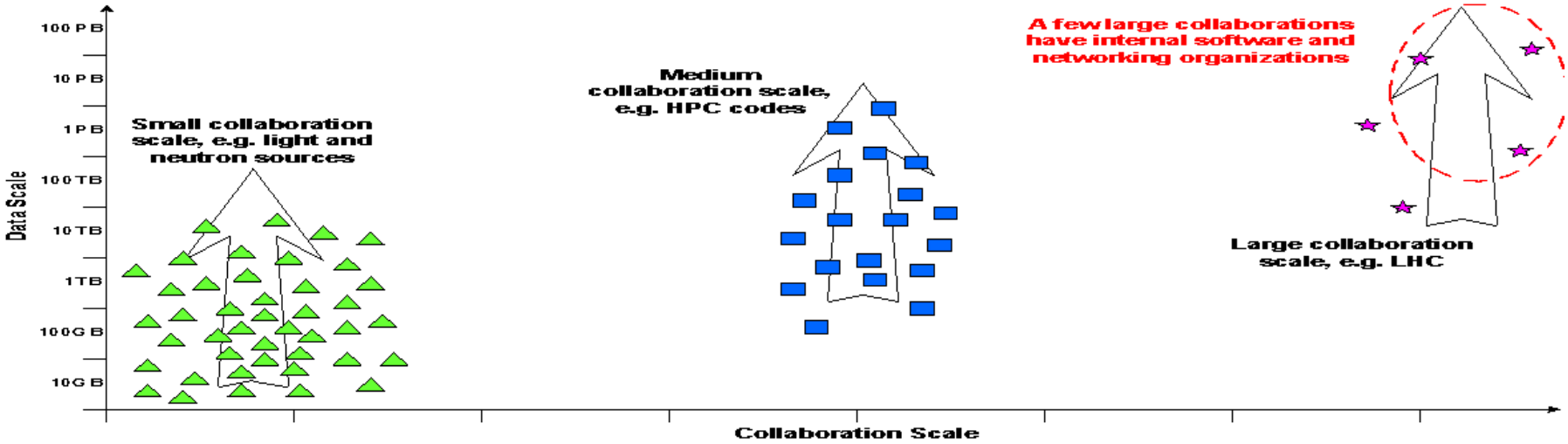
- Data mobility is required; there is no liquid market for HPC resources (people use what they can get – DOE, XSEDE, NOAA, etc. etc.)
  - To stay competitive, we must learn the use patterns, and support them
  - This may mean making sure your network, and the networks of others, are functional

ESnet ENERGY SCIENCES NETWORK GÉANT INDIANA UNIVERSITY INTERNET2 UNIVERSITY OF MICHIGAN

Outline

- Problem Statement on Network Connectivity
- Supporting Scientific Users
- Network Performance & TCP Behaviors w/ Packet Loss
- What is perfSONAR
- Deployment Overview
- Conclusions

Understanding Data Trends



http://www.es.net/science-engagement/science-requirements-reviews/

ESnet ENERGY SCIENCES NETWORK GÉANT INDIANA UNIVERSITY INTERNET2 UNIVERSITY OF MICHIGAN

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	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	13.33 Gbps	2.67 Gbps	666.67 Mbps	222.22 Mbps
10GB	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	13.33 Mbps	2.67 Mbps	0.67 Mbps	0.22 Mbps

This table available at:

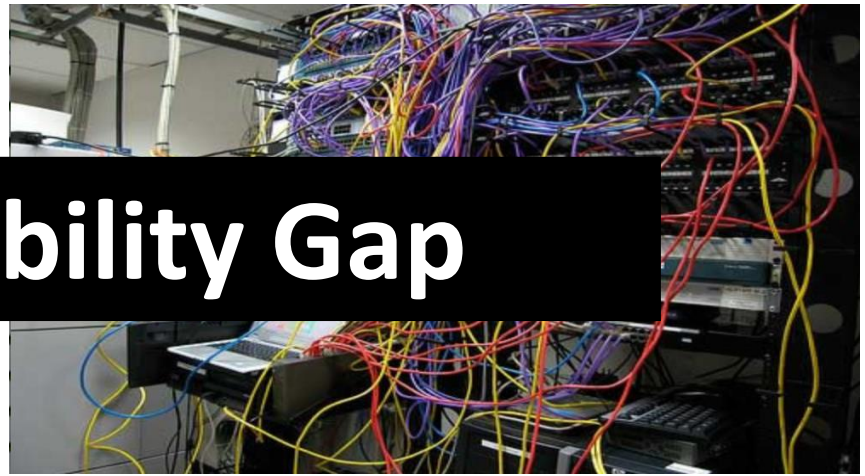
http://fasterdata.es.net/home/requirements-and-expectations

ESnet ENERGY SCIENCES NETWORK GÉANT INDIANA UNIVERSITY INTERNET2 UNIVERSITY OF MICHIGAN

Challenges to Network Adoption

- Causes of performance issues are complicated for users.
- Lack of communication and collaboration between the CIO's office and researchers on campus.
- Lack of IT expertise within a science collaboration
- User's perfor network is to work").
- Cultural change is hard ("we've always shipped disks!").
- Scientists want to do science not IT support

The Capability Gap



Outline

- Problem Statement on Network Connectivity
- Supporting Scientific Users
- Network Performance & TCP Behaviors w/ Packet Loss
- What is perfSONAR
- Deployment Overview
- Conclusions

Lets Talk Performance …

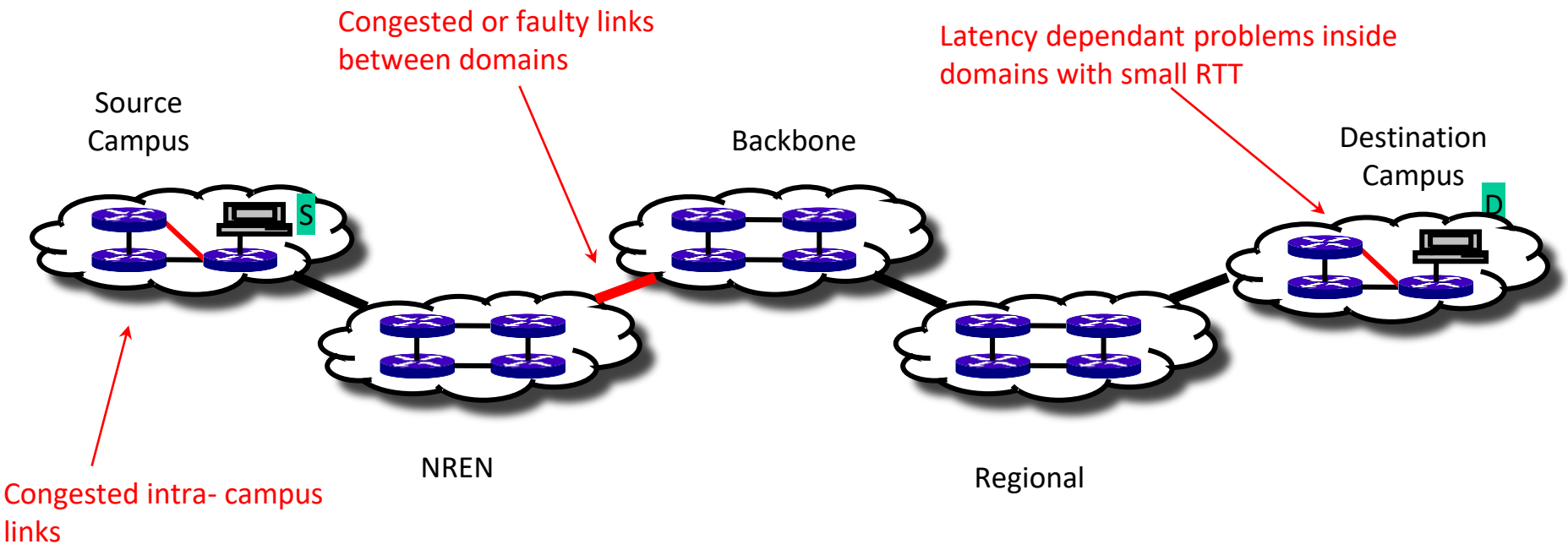
"In any large system, there is always something broken."
- ***Jon Postel***

- Modern networks are occasionally designed to be *one-size-fits-most*
  - e.g. if you have ever heard the phrase "converged network", the design is to facilitate CIA (Confidentiality, Integrity, Availability)

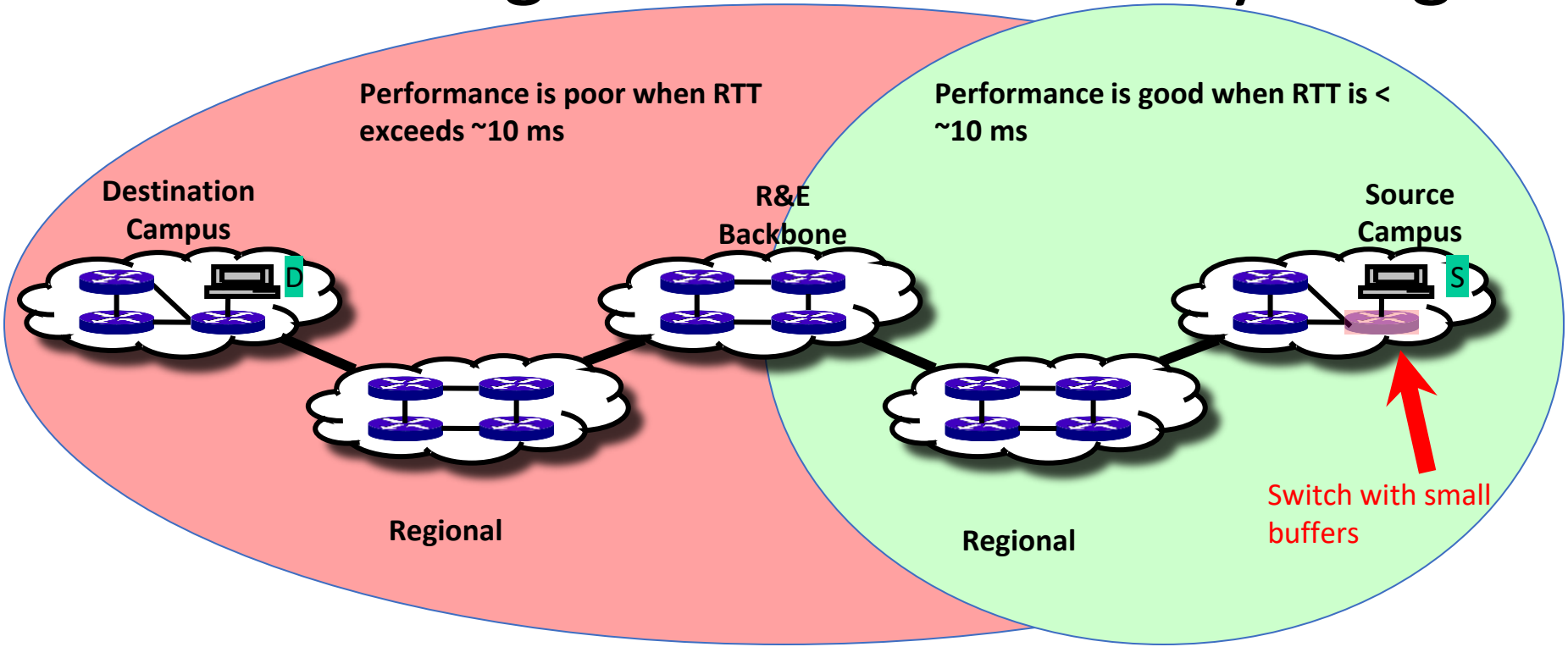
- It's all TCP
  - Bulk data movement is a common thread (move the data from the microscope, to the storage, to the processing, to the people – and they are all sitting in different facilities)
  - This fails when TCP suffers due to path problems (***ANYWHERE*** in the path)
  - It's easier to work with TCP than to fix it (20+ years of trying…)
- TCP suffers the most from unpredictability; Packet loss/delays are the enemy
  - Small buffers on the network gear and hosts
  - Incorrect application choice
  - Packet disruption caused by overzealous security
  - Congestion from herds of mice
- It all starts with knowing your users, and knowing your network

ESnet ENERGY SCIENCES NETWORK GÉANT INDIANA UNIVERSITY INTERNET2 UNIVERSITY OF MICHIGAN

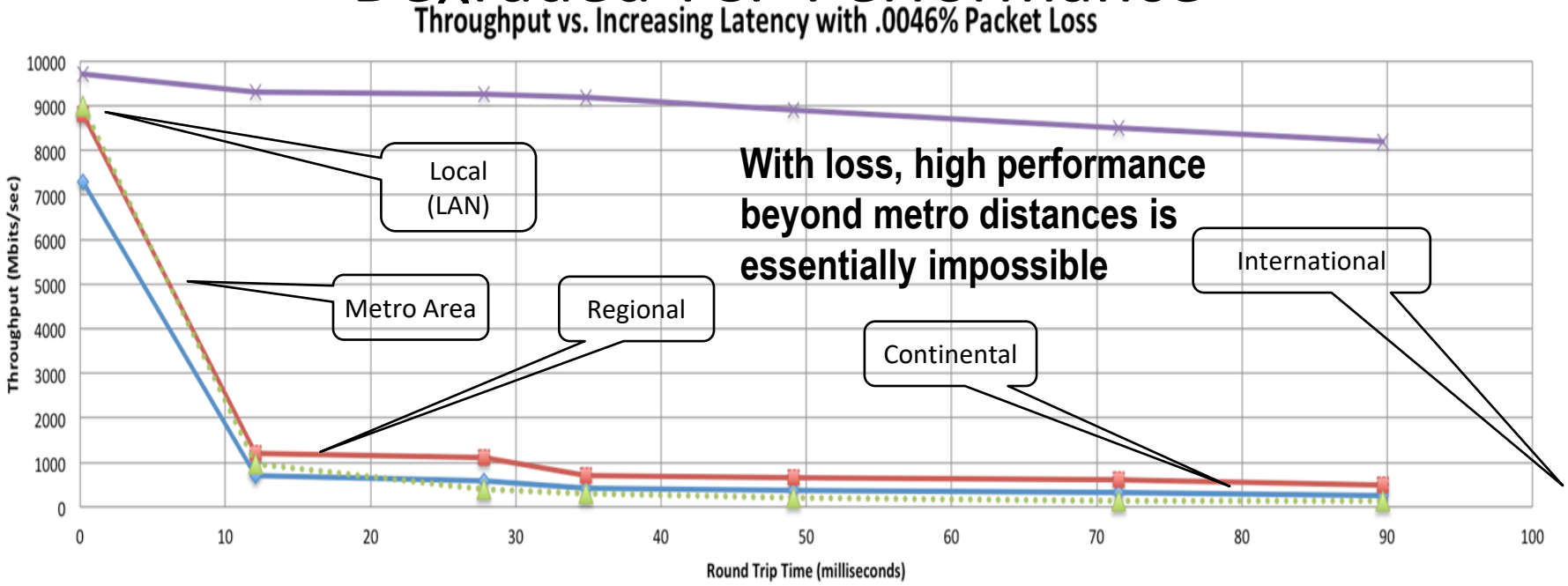
Where Are The Problems?



Local Testing Will Not Find Everything



Soft Failures Cause Packet Loss and Degraded TCP Performance



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**

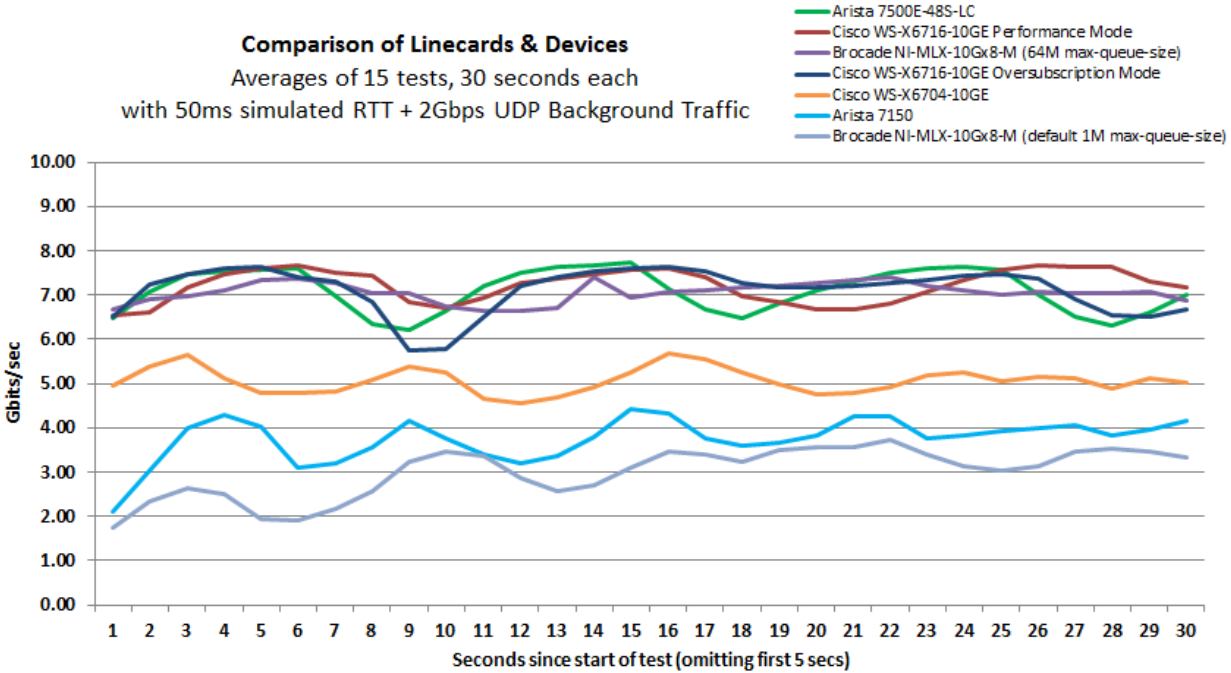
Problem Statement: Hard vs. Soft Failures

- **Hard failures are the kind of problems every organization understands**
  - **Fiber cut**
  - **Power failure takes down routers**
  - **Hardware ceases to function**
- **Classic monitoring systems are good at alerting hard failures**
  - **i.e., NOC sees something turn red on their screen**
  - **Engineers paged by monitoring systems**
- **Soft failures are different and often go undetected**
  - **Basic connectivity (ping, traceroute, web pages, email) works**
  - **Performance is just poor**
- **How much should we care about soft failures?**

Causes of Packet Loss

- Network Congestion
  - Easy to confirm via SNMP, easy to fix with $$
  - This is not a soft failure, but just a network capacity issue
  - Often people assume congestion is the issue when in fact it is not.
- Under-buffered switch dropping packets
  - Hard to confirm
- Under-powered firewall dropping packets
  - Hard to confirm
- Dirty fibers or connectors, failing optics/light levels
  - Sometimes easy to confirm by looking at error counters in the routers
- Overloaded or slow receive host dropping packets
  - Easy to confirm by looking at CPU load on the host

Under-buffered Switches are probably our biggest problem today



ESnet ENERGY SCIENCES NETWORK GÉANT INDIANA UNIVERSITY INTERNET2 UNIVERSITY OF MICHIGAN

Outline

- Problem Statement on Network Connectivity
- Supporting Scientific Users
- Network Performance & TCP Behaviors w/ Packet Loss
- What is perfSONAR
- Deployment Overview
- Conclusions

But … It's Not Just the Network

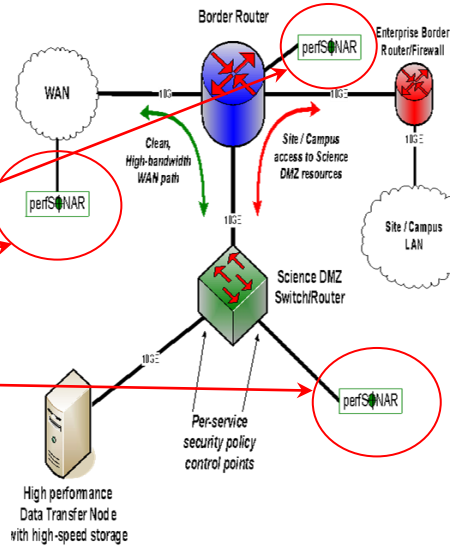
- **Perhaps you are saying to yourself "I have no control over parts of my campus, let alone the 5 networks that sit between me and my collaborators"**
  - Significant gains are possible in isolated areas of the OSI Stack
- **Things *you* control:**
  - Choice of data movement applications (say no to SCP and RSYNC)
  - Configuration of local gear (hosts, network devices)
  - Placement and configuration of diagnostic tools, e.g. *perfSONAR*
  - Use of the diagnostic tools
- **Things that require collaboration:**
  - Configuration of remote gear
  - Addressing issues when the diagnostic tools alarm
  - Getting someone to "care"

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.**
  - 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?
- **perfSONAR = Measurement Middleware**

perfSONAR

- All the previous Science DMZ network diagrams have little perfSONAR boxes everywhere
  - Consistent behavior requires correctness
  - Correctness requires the ability to find and fix problems
    - ***You can't fix what you can't find***
    - ***You can't find what you can't see***
    - ***perfSONAR lets you see***
- Especially important when deploying high performance services
  - If there is a problem with the infrastructure, need to fix it
  - If the problem is not with your stuff, need to prove it
    - Many players in an end to end path
    - Ability to show correct behavior aids in problem localization



What is perfSONAR?

- perfSONAR is a tool to:
  - Set network performance expectations
  - Find network problems (soft failures)
  - Track network performance history

  …All in multi-domain environments
- These problems are all harder when multiple networks are involved
- perfSONAR provides a standard way to publish active and passive monitoring data
  - This data is interesting to network analysts as well as network operators

perfSONAR History

- perfSONAR can trace its origin to the Internet2 “End 2 End performance Initiative” from the year 2000. What has changed since then?
  - The Good News:
    - TCP is much less fragile; Cubic is the default CC alg, autotuning is and larger TCP buffers are everywhere
    - Reliable parallel transfers via tools like Globus Online
    - High-performance UDP-based commercial tools like Aspera
  - The Bad News:
    - The capability gap between end users and network engineers is still large
    - Frame size mismatches are still common
    - Under-buffered and switches and routers are still in use
    - Under-powered/misconfigured firewalls still exist
    - Soft failures can go undetected for months
    - Users, in general, don’t know what to expect

Simulating Performance

- It's infeasible to perform at-scale data movement all the time – as we see in other forms of science, we need to rely on simulations
- Network performance comes down to a couple of key metrics:
  - Throughput (how much can I get out of the network?)
  - Latency (time it takes to get to/from a destination)
  - Packet loss/duplication/ordering (for some sampling of packets, do they all make it to the other side without serious abnormalities occurring?)
  - Network utilization (bandwidth in use vs total capacity)
- We can get many of these from a selection of active and passive measurement tools – enter the perfSONAR Toolkit

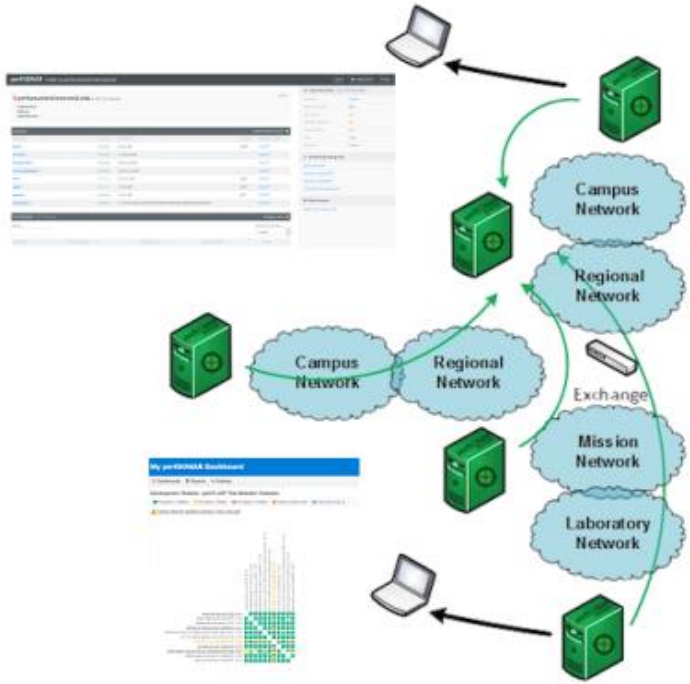
perfSONAR Toolkit

- The perfSONAR Toolkit is an open source implementation and packaging of the perfSONAR measurement infrastructure and protocols
  - http://docs.perfsonar.net/install_getting.html
- All components are available as RPMs, DEBs, and bundled as CentOS 7, Debian 9, or Ubuntu 16/18 - based packages (as for perfSONAR v4.2.x)
  - perfSONAR tools are much more accurate if run on a dedicated perfSONAR host
- Very easy to install and configure
  - Usually takes less than 30 minutes

Outline

- Problem Statement on Network Connectivity
- Supporting Scientific Users
- Network Performance & TCP Behaviors w/ Packet Loss
- What is perfSONAR
- Deployment Overview
- Conclusions

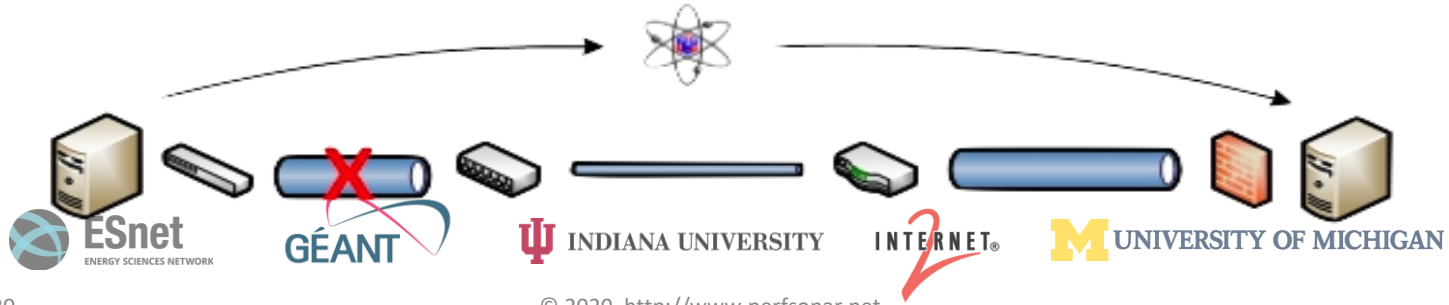
Toolkit “Beacon” Use Case



- The general use case is to establish some set of tests to other locations/facilities
- To answer the what/why questions:
  - Regular testing with select tools helps to establish patterns – how much bandwidth is available during the course of the day – or when packet loss appears
  - We do this to points of interest to see how well a real activity (e.g. Globus transfer) would do.
- If performance is ‘bad’, don’t expect much from the data movement tool

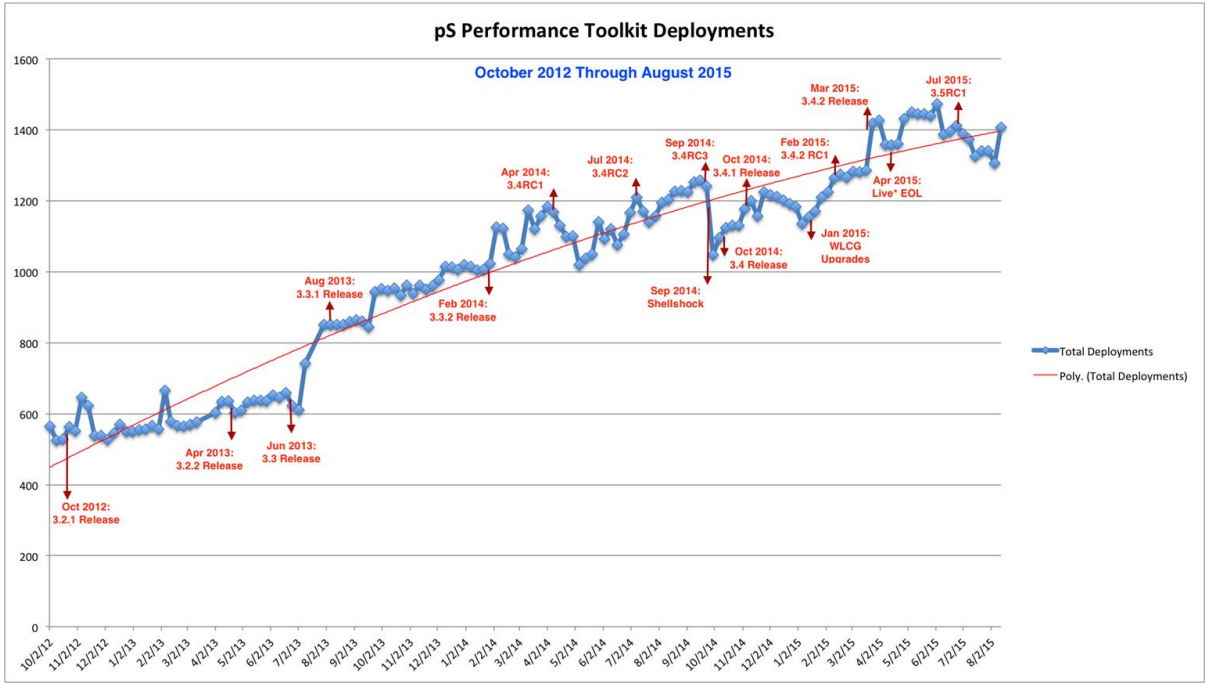
Benefits: Finding the needle in the haystack

- Above all, perfSONAR allows you to maintain a healthy, high-performing network because it helps identify the soft failures in the network path.
  - Classical monitoring systems have limitations
    - Performance problems are often only visible at the ends
    - Individual network components (e.g. routers) have no knowledge of end host state
  - perfSONAR tests the network in ways that classical monitoring systems do not
- More perfSONAR distributions equal better network visibility.

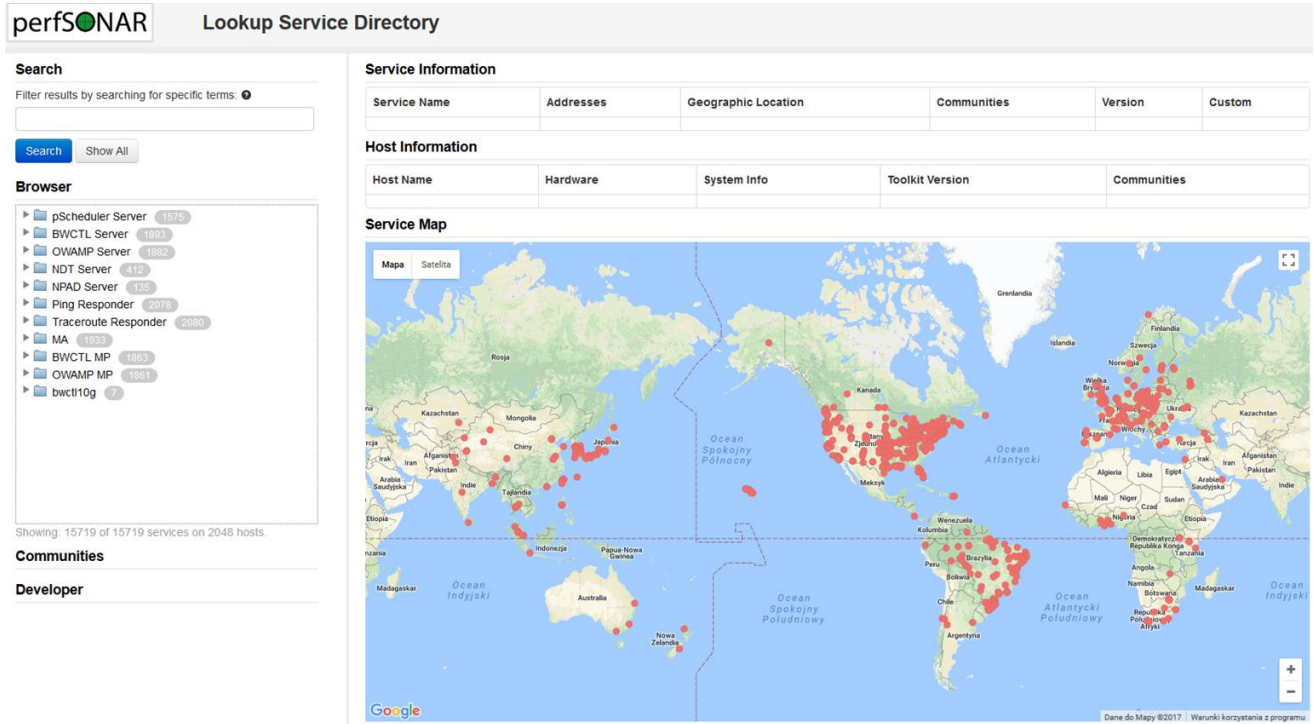


Deployment By The Numbers

- Last updated August 2015. Adoption trend increases with each release. CC-NIE and innovation platform helped as well.



http://stats.es.net/ServicesDirectory/

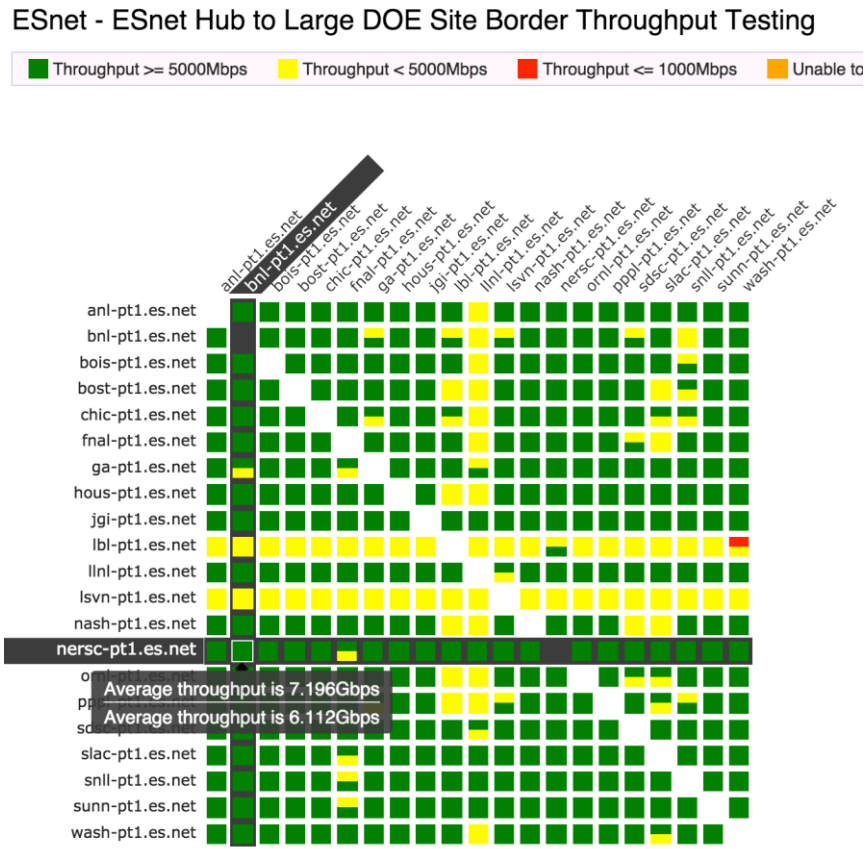


perfSONAR Dashboard: Raising Expectations and improving network visibility

Status at-a-glance
- Packet loss
- Throughput
- Correctness

Current live instances at
http://pas.net.internet2.edu/
http://ps-dashboard.es.net/

Drill-down capabilities:
- Test history between hosts
- Ability to correlate with other events
- Very valuable for fault localization and isolation



Outline

- Problem Statement on Network Connectivity
- Supporting Scientific Users
- Network Performance & TCP Behaviors w/ Packet Loss
- What is perfSONAR
- Deployment Overview
- Conclusions

Benefit: Active and Growing Community

- Active email lists provide:
  - Instant access to advice and expertise from the community.
  - Ability to share metrics, experience and findings with others to help debug issues on a global scale.
- Joining the community automatically increases the reach and power of perfSONAR
  - The more endpoints means exponentially more ways to test and discover issues, compare metrics

ESnet ENERGY SCIENCES NETWORK GÉANT INDIANA UNIVERSITY INTERNET2 UNIVERSITY OF MICHIGAN

perfSONAR Community

- The perfSONAR collaboration is working to build a strong user community to support the use and development of the software.
- perfSONAR Mailing Lists
  - Announcement Lists:
    - https://lists.internet2.edu/sympa/subscribe/perfsonar-announce
  - Users List:
    - https://lists.internet2.edu/sympa/subscribe/perfsonar-user

Resources

- perfSONAR website
  - http://www.perfsonar.net
- perfSONAR Documentation
  - http://docs.perfsonar.net
- perfSONAR mailing lists
  - https://www.perfsonar.net/about_contact.html
- perfSONAR directory
  - http://stats.es.net/ServicesDirectory
- perfSONAR YouTube Channel
  - https://www.youtube.com/perfSONARProject
- FasterData Knowledgebase
  - http://fasterdata.es.net

What is perfSONAR?

Meshbuilder Workshop

Scott Chevalier, IN@IU, schevali@iu.edu
Antoine Delvaux, GEANT Project, antoine.delvaux@man.poznan.pl
Doug Southworth, IN@IU, dojosout@iu.edu

ESnet ENERGY SCIENCES NETWORK GÉANT INDIANA UNIVERSITY INTERNET2 UNIVERSITY OF MICHIGAN