perfSONAR-based Network Research

Brian Tierney, ESnet, bltierney@es.net

Febuary 10, 2016

perfSONAR

What is perfSONAR?

- perfSONAR is a tool to:
  - Set (hopefully raise) 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
  - Focus on Research and Education (R&E) Networking, 1Gbps links or higher
- 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

ESnet ENERGY SCIENCES NETWORK GÉANT INDIANA UNIVERSITY INTERNET2

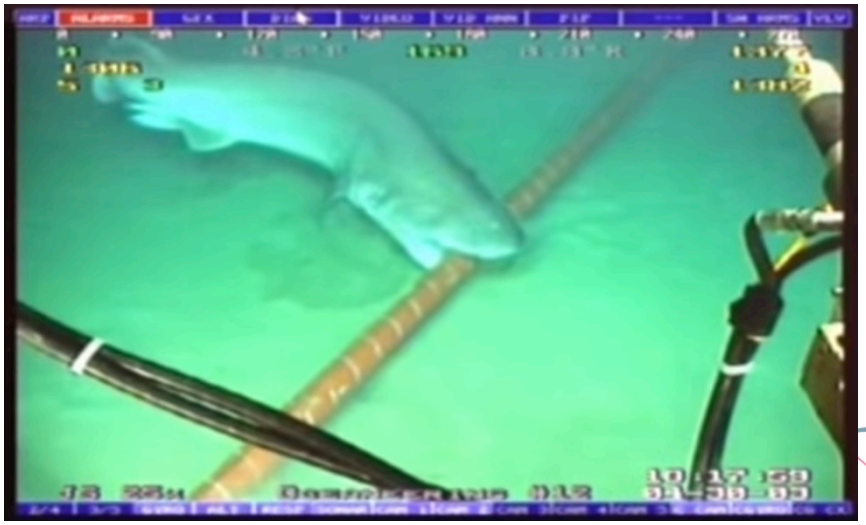
Current perfSONAR components

- **Measurement tools**
  - iperf3, bwctl, owamp, traceroute, paris-traceroute, etc.
- **Measurement archive**
- Central test mesh management tools
- **Host management tools**
  - Configure tests, configure NTP, etc.
- **Data analysis tools**
  - Plot data from the archive
  - Dashboard tools
- **Lookup Service**

perfSONAR

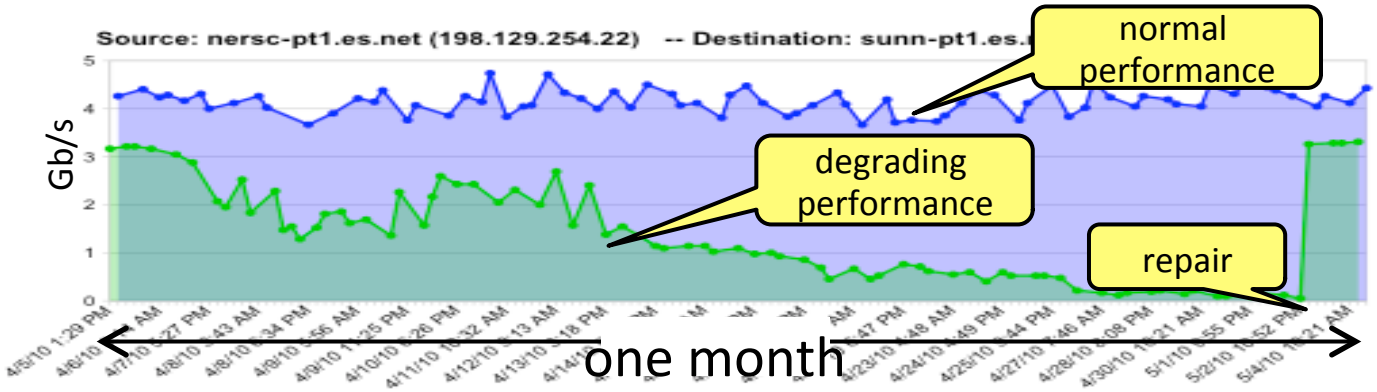
Hard vs. Soft Failures

- "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?

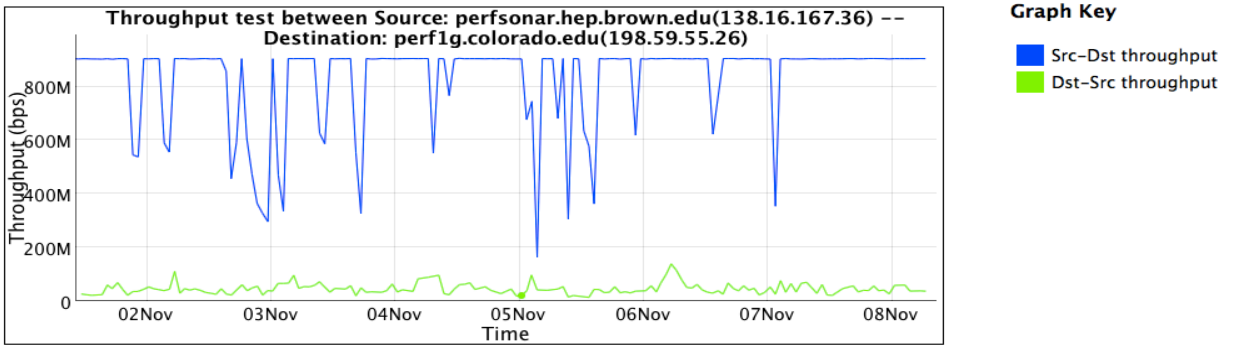


Main perfSONAR role: Find "Soft Failures"

perfSONAR



Gradually failing optics



Under-Powered firewall device

ESnet ENERGY SCIENCES NETWORK GÉANT INDIANA UNIVERSITY INTERNET2

perfSONAR History

- perfSONAR can trace its origin to the Internet2 “End 2 End performance Initiative” from the year 2000.
- What has changed since 2000?
  - 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
    - *more good news in latest Linux kernel, but it will take 3-4 years before this is widely deployed*
  - The Bad News:
    - The wizard gap is still large
    - Under-buffered and switches and routers are still common
    - Under-powered/misconfigured firewalls are common
    - Soft failures still go undetected for months
    - User performance expectations are still too low

The perfSONAR collaboration

perfSONAR

- The perfSONAR collaboration is a Open Source project lead by ESnet, Internet2, Indiana University, and GEANT.
  - Each organization has committed 1.5 FTE effort to the project
  - Plus additional help from many others in the community (OSG, RNP, SLAC, and more)
- The perfSONAR Roadmap is influence by
  - requests on the project issue tracker
  - annual user surveys sent to everyone on the user list
  - regular meetings with VO using perfSONAR such as the WLCG and OSG
  - discussions at various perfSONAR related workshops
- Based on the above, every 6-12 months the perfSONAR governance group meets to prioritize features based on:
  - impact to the community
  - level of effort required to implement and support
  - availability of someone with the right skill set for the task

ESnet ENERGY SCIENCES NETWORK GÉANT INDIANA UNIVERSITY INTERNET2

perfSONAR

public perfSONAR Servers (Jan 2016)

- Total of around 1700 publicly registered servers
  - Equal number of non-registered servers?
- ESnet: 70
  - mostly 10G, includes a 40G host in Boston
- GEANT: 22
- Internet2: 3
- Some other top deployments:
  - Onenet (24), AMPATH (8), bc.net (10), RNP (8), Canarie (13), kreonet (14), NERO(12), AARnet (19), JGN (17), CENIC (5), KANREN (5)

perfSONAR Hardware

- These days you can get a good 1U host capable of pushing 10Gbps TCP for around $500 (+10G NIC cost, $750?).
  - See perfSONAR user list
- And you can get a host capable of 1G for around $150!
  - Get a multi-core Intel Celeron-based host
    - ARM is not fast enough
  - e.g.: ZBOX by ZOTAC: https://www.zotac.com/us/product/mini_pcs/zbox-ci323-nano

- VMs are not recommended
  - Tools more accurate if can guarantee NIC isolation

perfSONAR 3.5 Update

- perfSONAR 3.5 released October, 2015
  - Modernize the GUIs
  - Support for central host management and node auto-configuration
  - Support for Debian, VMs, and other installation options
  - Support for low cost ($150), 1 Gbps nodes

perfSONAR

Expanded perfSONAR Use Cases

- Previous Use Case: perfSONAR Toolkit
  - Includes CentOS 6 and all perfSONAR components
- New Use Cases
  - perfSONAR tools only
    - Support for both RHEL-based and Debian-based hosts
  - perfSONAR hosts that are centrally managed
    - Central manager package
    - Testpoint package

ESnet ENERGY SCIENCES NETWORK GÉANT INDIANA UNIVERSITY INTERNET2

perfSONAR for Network Researchers

- Active measurement interesting for network researchers
  - Traceroute data automatically collected along with bwctl/owamp results
  - TCP retransmits as measured by iperf3
- Data easy to download for analysis
  - esmond-ps-get-bulk
    - Output CSV or JSON
    - See: https://pypi.python.org/pypi/esmond_client
- Additional Information at:
  - http://docs.perfsonar.net/client_apis.html

perfSONAR on Low Cost Hardware

- Motivation: make perfSONAR affordable enough to deploy on all subnets
- Assumptions:
  - 1Gbps test nodes
  - Centralized measurement archive
  - Centralized configuration management
  - Debian Linux

Current perfSONAR development

perfSONAR

- One of the themes for v3.6 will be “Control and Scalability”
  - perfSONAR is successful because of the ‘default open’ model.
  - BUT, as the number of perfSONAR hosts worldwide grows, we need a way to control
    - Who is running tests
    - How often are they allowed to run tests
    - What hosts can I run tests to? How to I get my host added to someone else’s list of allowed hosts?
- Working on a new test scheduler (psScheduler):
  - Shared by all tests and aware of the resources each uses
  - Containing finer grained controls about who can run tests and what tests they are allowed to run.
  - Increased visibility and control as to when tests will be run

ESnet ENERGY SCIENCES NETWORK GÉANT INDIANA UNIVERSITY INTERNET2®

Roadmap for v3.6

perfSONAR

- A test scheduler (psScheduler):
  - Shared by all tests and aware of the resources each uses
  - Containing finer grained controls about who can run tests and what tests they are allowed to run.
  - Increased visibility and control as to when tests will be run
- New graphs that allow for easier comparison of multiple metrics
  - based on ESnet Tools team react-based plotting tools
- A web interface for creating test meshes
- Easier selection of endpoints based on topology location, geographic location, accessibility and/or custom searches
- Dashboards that support alerting based on patterns across an entire mesh
- Debian 8 support
- CentOS 7 versions of the tools, testpoint, core, and central management bundles
  - Full CentOS 7 Toolkit will be in the next release
- Pre-packaged perfSONAR VM images

ESnet ENERGY SCIENCES NETWORK GÉANT INDIANA UNIVERSITY INTERNET2

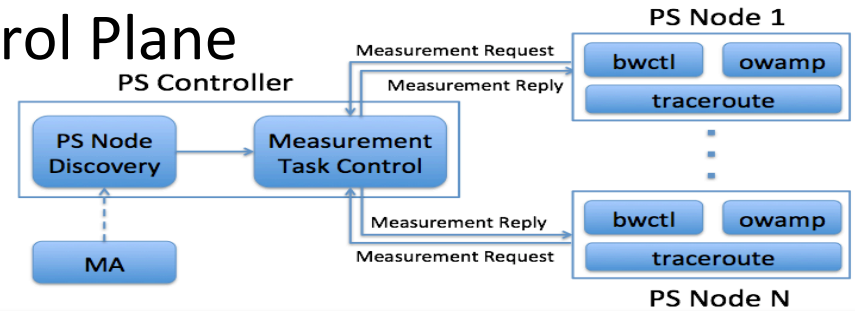
Example perfSONAR Research Projects

perfSONAR Control Plane (PSCP) Project

Prof. Yan Luo (yan_luo@uml.edu)

perfSONAR Control Plane (PSCP)

- Objectives
  - Measurement Archive Data Analysis
    - What are the measurement results? What can we learn?
  - Automatic perfSONAR Peer Selection
    - Quickly identify the best suitable PS node(s) on the routes in question
  - Programmable Measurement and Troubleshooting
    - Define measurement task and conditions with software
- The Design of perfSONAR Control Plane
  - Path Discovery
  - Measurement Task Control



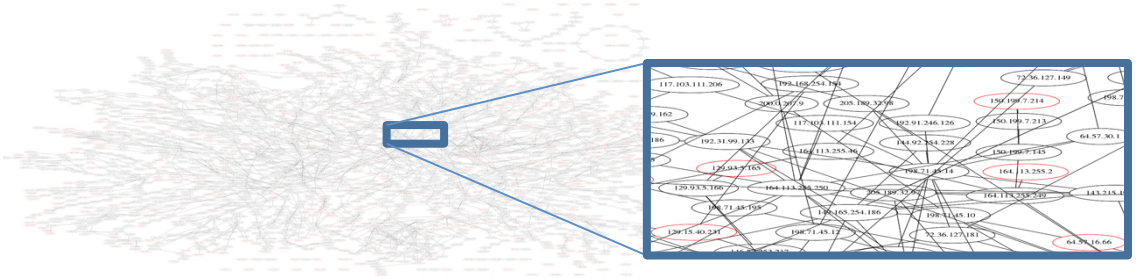
https://github.com/ACANETS/pscp

Prof. Yan Luo (yan_luo@uml.edu)

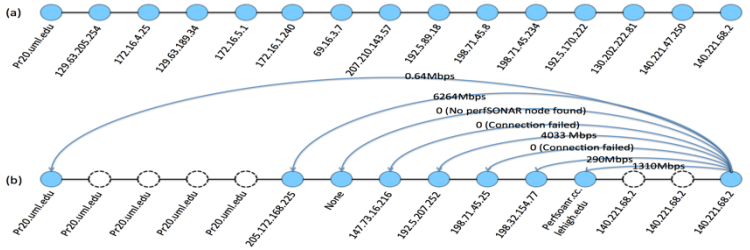
Operation and Use Case of PSCP

perfSONAR

- Obtain traceroute information from 95 perfSONAR Measurement Archives
- Build a traceroute graph based on the 1831 records



- Find a set of perfSONAR node pairs to start bandwidth tests and monitor the results
- Use Case: Diagnostic analysis and trouble-shooting a soft network link failure
  - <= 300 LOC Python code, <= 15 minutes



https://github.com/ACANETS/pscp

Prof. Yan Luo (yan_luo@uml.edu)

UMASS LOWELL
Learning with Purpose

Pythia Network Diagnosis Infrastructure (PuNDIT)

PIs: Shawn McKee (smckee@umich.edu) and Constantine Dovrolis (constantine@gatech.edu)

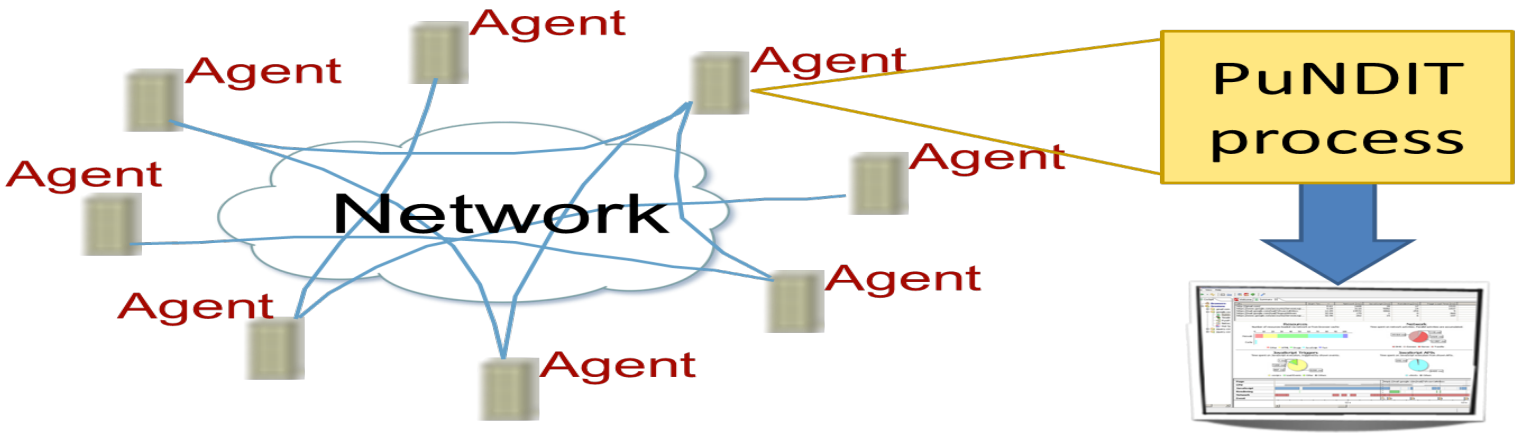
About PuNDIT

perfSONAR

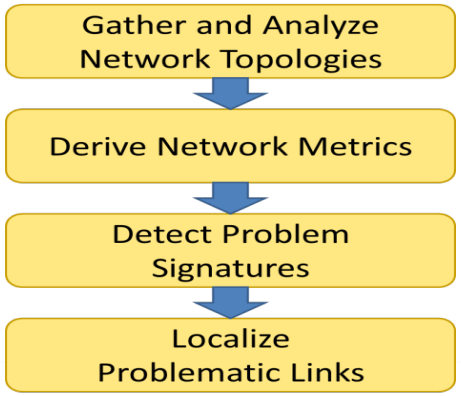
- PuNDIT is a NSF SSI project which uses perfSONAR data to identify and localize network problems (2014-2016)
  - **Goal** to automate watching/analyzing perfSONAR metrics
    - inform users/site-admins when there are real network problems they should address.
- See further details at http://pundit.gatech.edu
- User GUI mock-up http://punditui.aglt2.org/

PuNDIT Architecture

perfSONAR



- perfSONAR provides the base measurement infrastructure
  - Collects network metrics like latency, loss and reordering
  - Collects topological information
  - Adds scamper support to perfSONAR: Multipath Detection Algorithm (MDA) from the *paris-traceroute* team to handle load balanced paths
- A lightweight PuNDIT process on each host performs detection
- The central server holds event repository and runs a localization algorithm

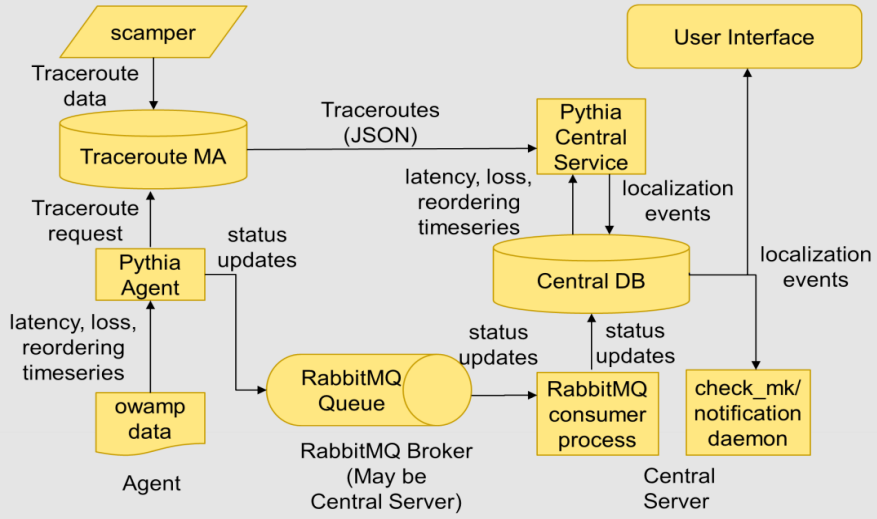


PuNDIT Details

perfSONAR

Information Flow

- 2 main message types sent across network:
  - Status updates indicate the network metrics for 5 second windows, sent in a batch every minute
  - Traceroutes in JSON from the Traceroute MA on each host
- Central Service generates localization events that trigger UI, check_mk and notification daemon



Problems Detected

Excessive Delays, Losses and Reordering	When these parameters in a 5 second window exceed user-specified thresholds
Congestion	Delay increases correlated with losses, indicating possible congestion at an interface
Route Change	A change in route resulting in a sudden and significant change in delay measurements
Route Instability	Repeated route changes, which results in poor TCP performance
End-host Context Switch	Context switching of the active measurement applications on an end-host, which may cause measurement fluctuations to be erroneously detected as problems

Email Lists and Reference Materials

Active and Growing perfSONAR Community

perfSONAR

- Active email lists and forums 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

perfSONAR

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://mail.internet2.edu/wws/subrequest/perfsonar-announce
  - Users List:
    - https://mail.internet2.edu/wws/subrequest/perfsonar-users

ESnet ENERGY SCIENCES NETWORK GÉANT INDIANA UNIVERSITY INTERNET2

perfSONAR

Useful URLs

- http://docs.perfsonar.net/
- http://www.perfsonar.net/
- http://fasterdata.es.net/
  - http://fasterdata.es.net/performance-testing/network-troubleshooting-tools/
- https://github.com/perfsonar
  - https://github.com/perfsonar/project/wiki

ESnet ENERGY SCIENCES NETWORK GÉANT INDIANA UNIVERSITY INTERNET2

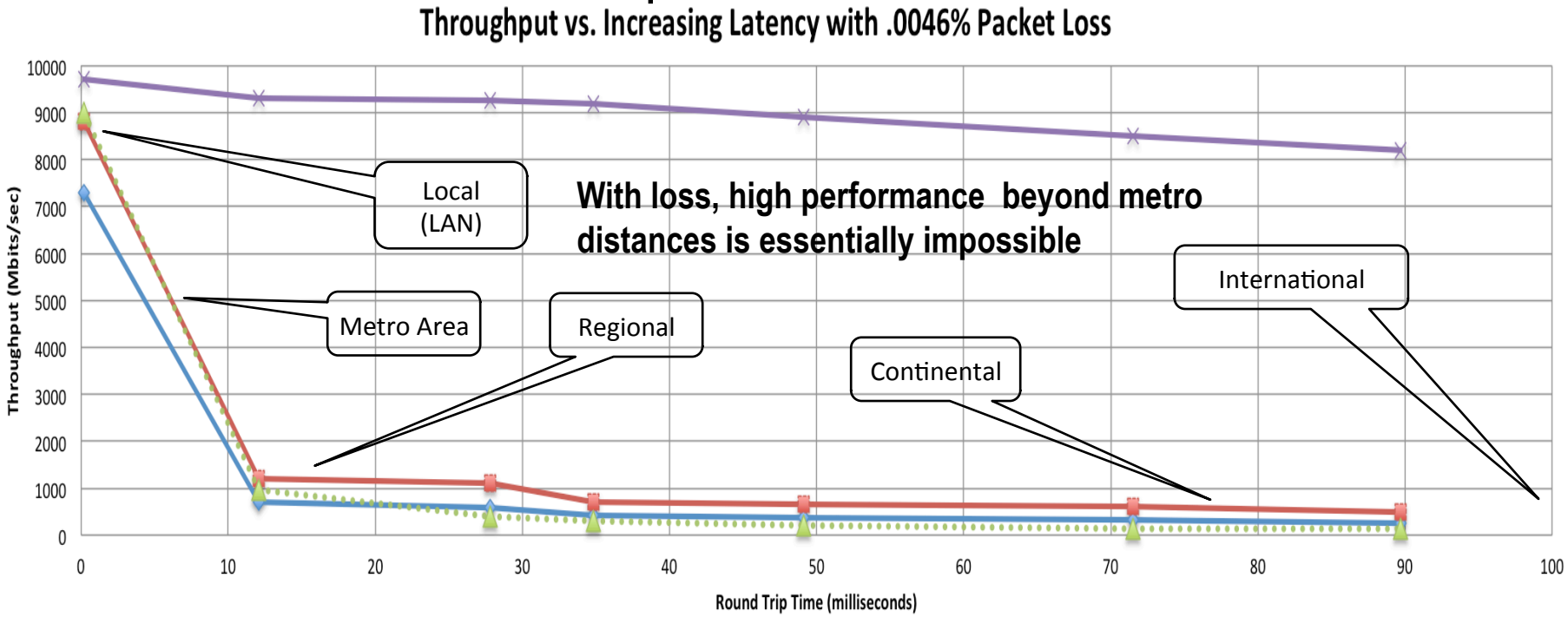
Extra Slides

bwctl features

- BWCTL lets you run any of the following between any 2 perfSONAR nodes:
  - iperf3, iperf, nuttcp, ping, owping, traceroute, and tracepath
- Sample Commands:
- `bwctl -c psmsu02.aglt2.org -s elpa-pt1.es.net -T iperf3`
- `bwping -s atla-pt1.es.net -c ga-pt1.es.net`
- `bwping -E -c www.google.com`
- `bwtraceroute -T tracepath -c lbl-pt1.es.net -l 8192 -s atla-pt1.es.net`
- `bwping -T owamp -s atla-pt1.es.net -c ga-pt1.es.net -N 1000 -i .01`

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

perfSONAR



Improved Support for Central Management

- Goals:
  - Make it easy to incorporate perfSONAR hosts into existing host management systems (puppet, chef, SaltStack, cfengine, etc.)
    - Include sample puppet config files
  - Make it easy to manage many perfSONAR hosts at a single institution
  - New rpm and debian bundles to support this

New perfSONAR Installation options

- In addition to the traditional "Toolkit" install, you now have the these additional options:
  - perfSONAR-Tools:
    - iperf3, bwctl, owamp, nuttcp, etc
    - Install this on DTNs, etc to help with troubleshooting
    - Does not support scheduled testing
    - CentOS and Debian support
  - perfSONAR-TestPoint:
    - tools plus Lookup Service registration and 'mesh agent'
    - For use in environments with a central measurement archive
    - For use on low end/older hardware (e.g.: $100 nodes)
    - Supports scheduled testing
    - CentOS and Debian support

- See: http://docs.perfsonar.net/install_options.html

New perfSONAR Installation options (cont.)

- perfSONAR-Core:
  - Includes everything except the web interface
  - Use this in environments where your site sysadmins want to fully manage the host configuration, but don't want to setup a central measurement archive
  - CentOS only
- perfSONAR-CentralManagement:
  - Includes measurement archive, test mesh manager, dashboard
  - Use this to manage a collection of perfSONAR hosts at your site/campus
  - CentOS only

New perfSONAR Installation options (cont.)

- perfSONAR-Complete
  - All perfSONAR packages
  - Use this environments where your sysadmins want to manage the install, but still use the toolkit web interface, system setting, etc
    - the toolkit install will override certain changes every update.
  - CentOS only
- Other packages to note:
  - Separate rpms/debs for iptables config, sysctl config, and ntp packages so you can add them on top of perfSONAR-Core as desired.