# Introduction to perfSONAR

Network Startup Resource Center

ESnet
ENERGY SCIENCES NETWORK

U.S. DEPARTMENT OF ENERGY
Office of Science

BERKELEY LAB



UNIVERSITY OF OREGON

Last updated 17th October 2016

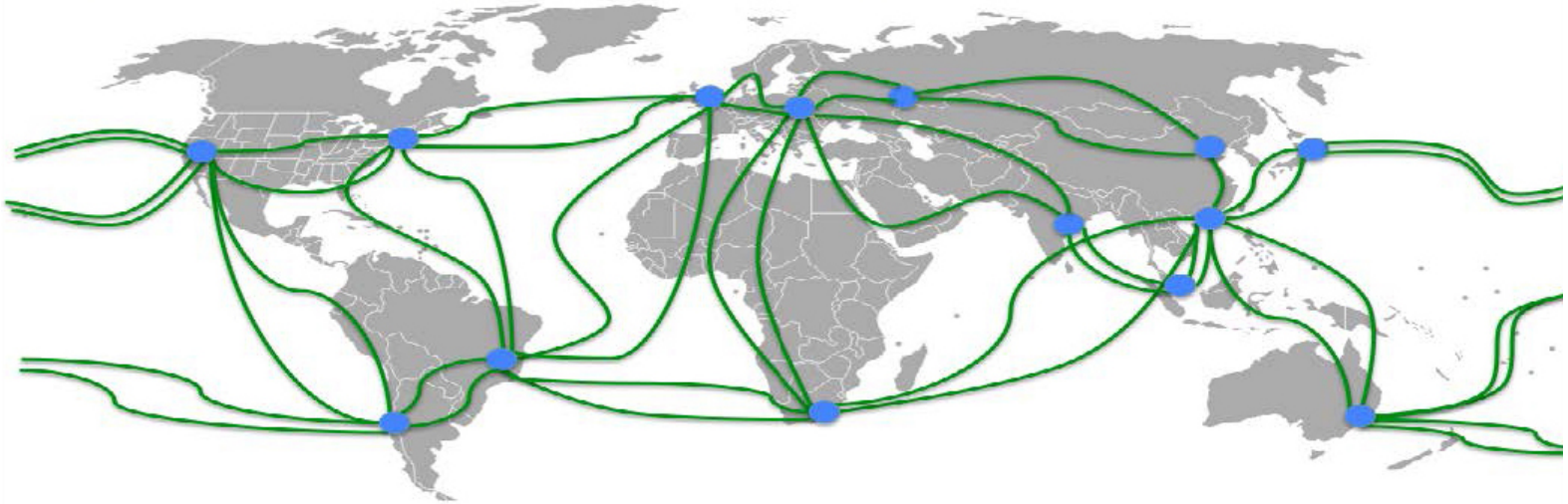
NSRC
Network Startup Resource Center

# What is it?

- ***perfSONAR*** is a network measurement toolkit that provides end-to-end statistics and helps to establish end-to-end usage expectations.
  - How fast is our circuit in reality?
  - Are there problems?
  - How do I fix problems?
- ***perfSONAR*** provides a uniform interface that allows for the scheduling of measurements, storage of data in uniform formats, and scalable methods to retrieve data and generate visualizations to help answer the three questions above.
- http://www.perfsonar.net/about/what-is-perfsonar/

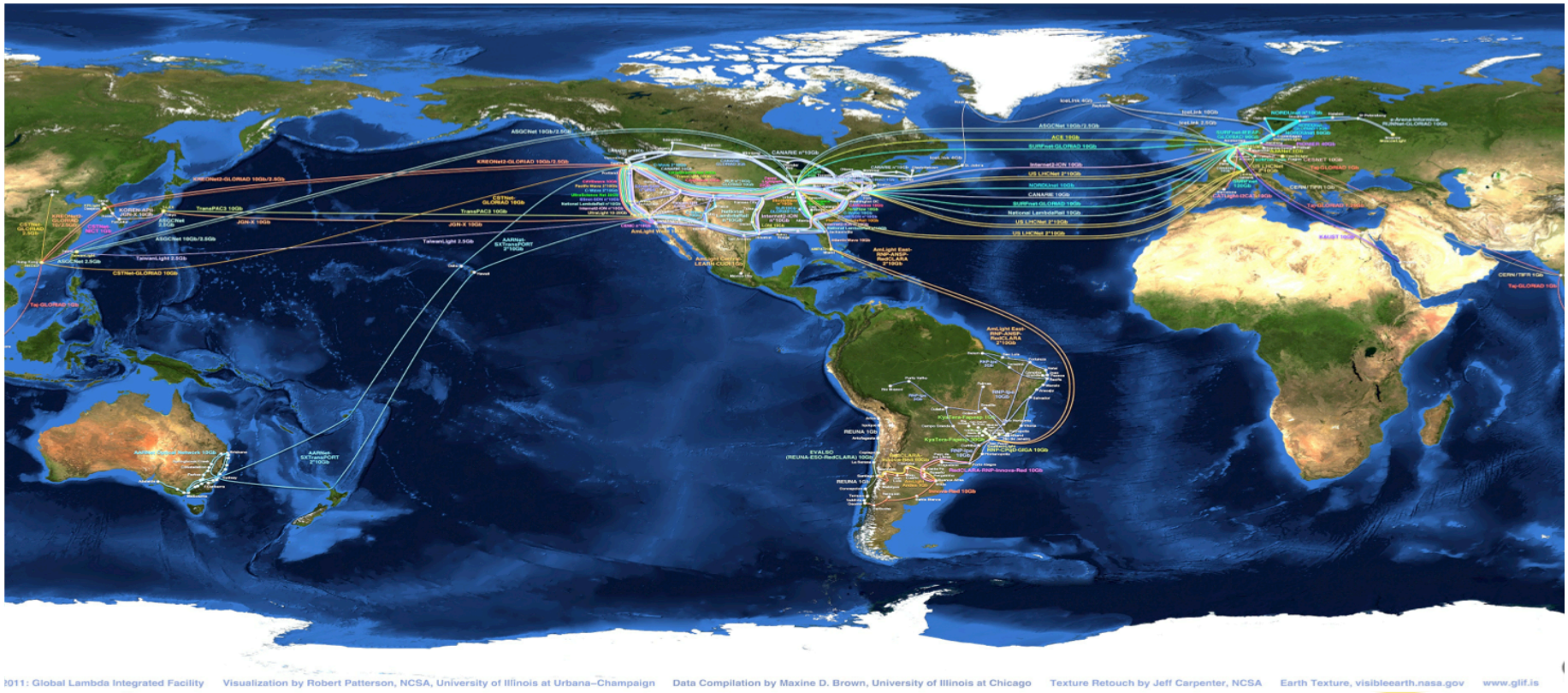
# Problem Statement

- The global Research and 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.



# Problem Statement

- In practice, performance issues are prevalent and distributed.
- When a network is underperforming or errors occur, it is difficult to identify the source, as problems can happen anywhere, in any domain.
- Local-area network testing is not sufficient, as errors can occur between networks.

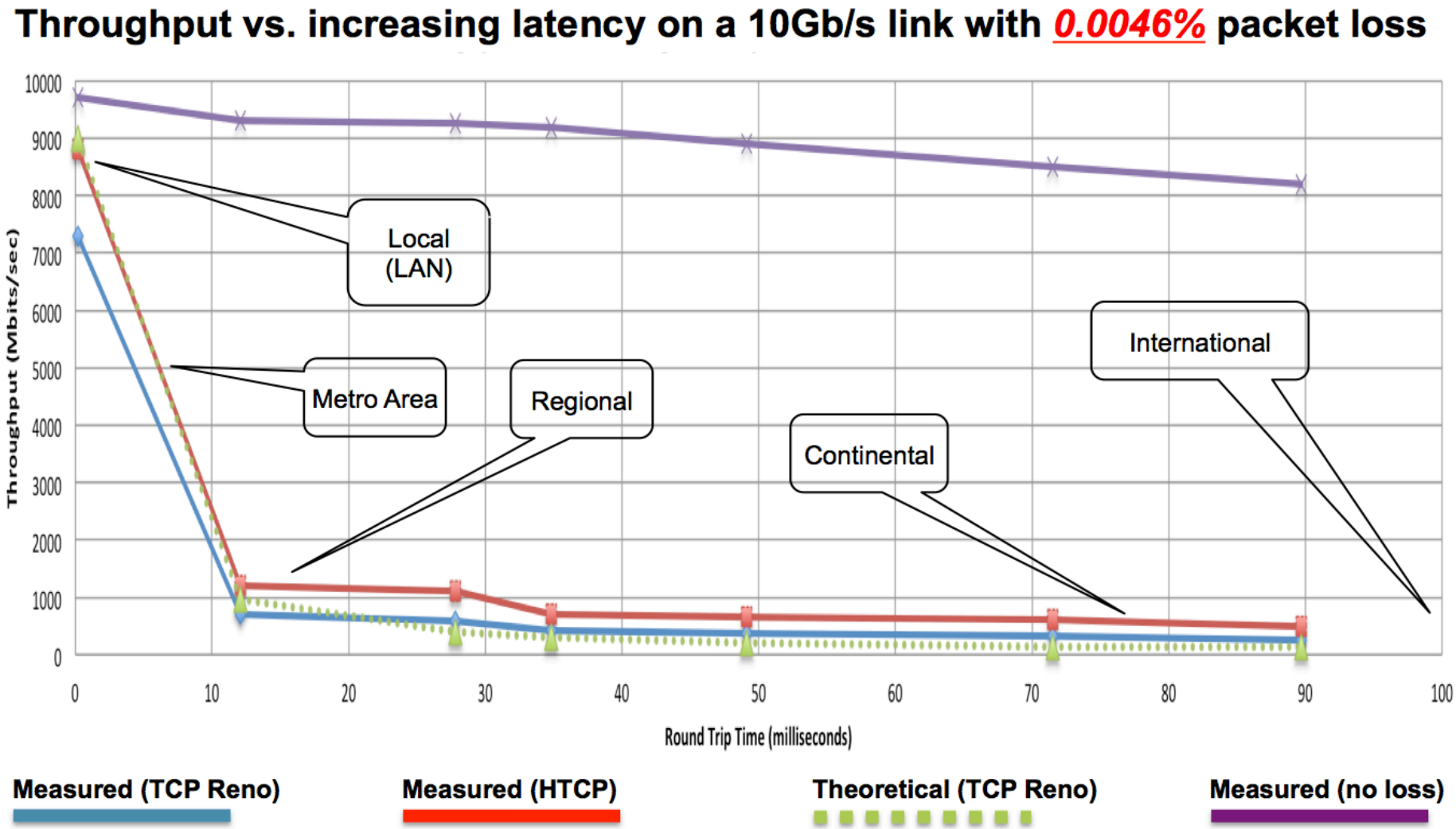
# Problem Statement
## Hard vs Soft Failures

- “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?

# What does it take to transfer data...

| 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 |

# But, just a little packet loss...



# A Solution
# perfSONAR and the perfSONAR Toolkit

- *A widely-deployed test and measurement infrastructure used by local, regional and national research networks and science facilities to actively track & troubleshoot network performance issues.*
- Open-source, community-developed software suite that:
  - Provides consistent set of proven tools to pinpoint and resolve network performance issues including soft failures across complex, multi-domain data paths.
  - Creates a standard way to visualize, publish, and archive network metrics and data for future analysis, and aids other networks in debugging issues.

# perfSONAR Toolkit

- What do we care about measuring?

- Latency: Ping, OWAMP, NDT, NPAD (don't trust Traceroute)
- Jitter: OWAMP, UDP Iperf, UDP NuQcp, NDT, NPAD
- Packet loss, duplication, unordered data: OWAMP, UDP Iperf, UDP NuQcp, NDT, NPAD
- Interface utiliization, discards, errors: - SNMP
- Achievable bandwidth: Iperf, NuQcp, NDT, NPAD
- Traveled route: Traceroute, Tracepath
- MTU feedback: Tracepath, Ping, NDT, NPAD

# Benefit: Multi-domain Insight

- Over 1700 public perfSONAR nodes are deployed globally
  - This allows you to test performance between your perfSONAR node and any other publicly available perfSONAR node. Very cool. Reverse traceroute!

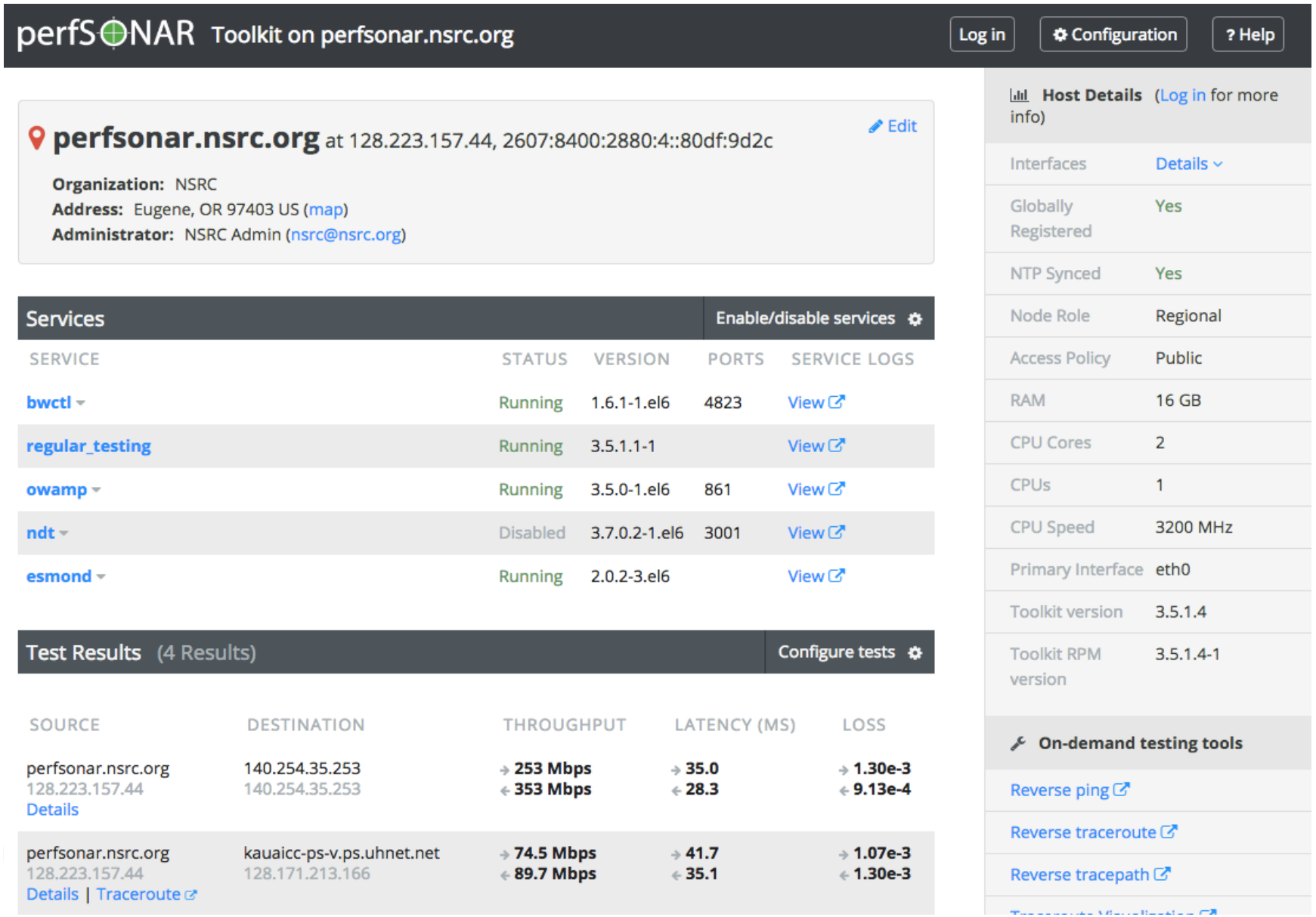
# perfSONAR Community Benefits

- i.e., Testing from client A between points B and C...

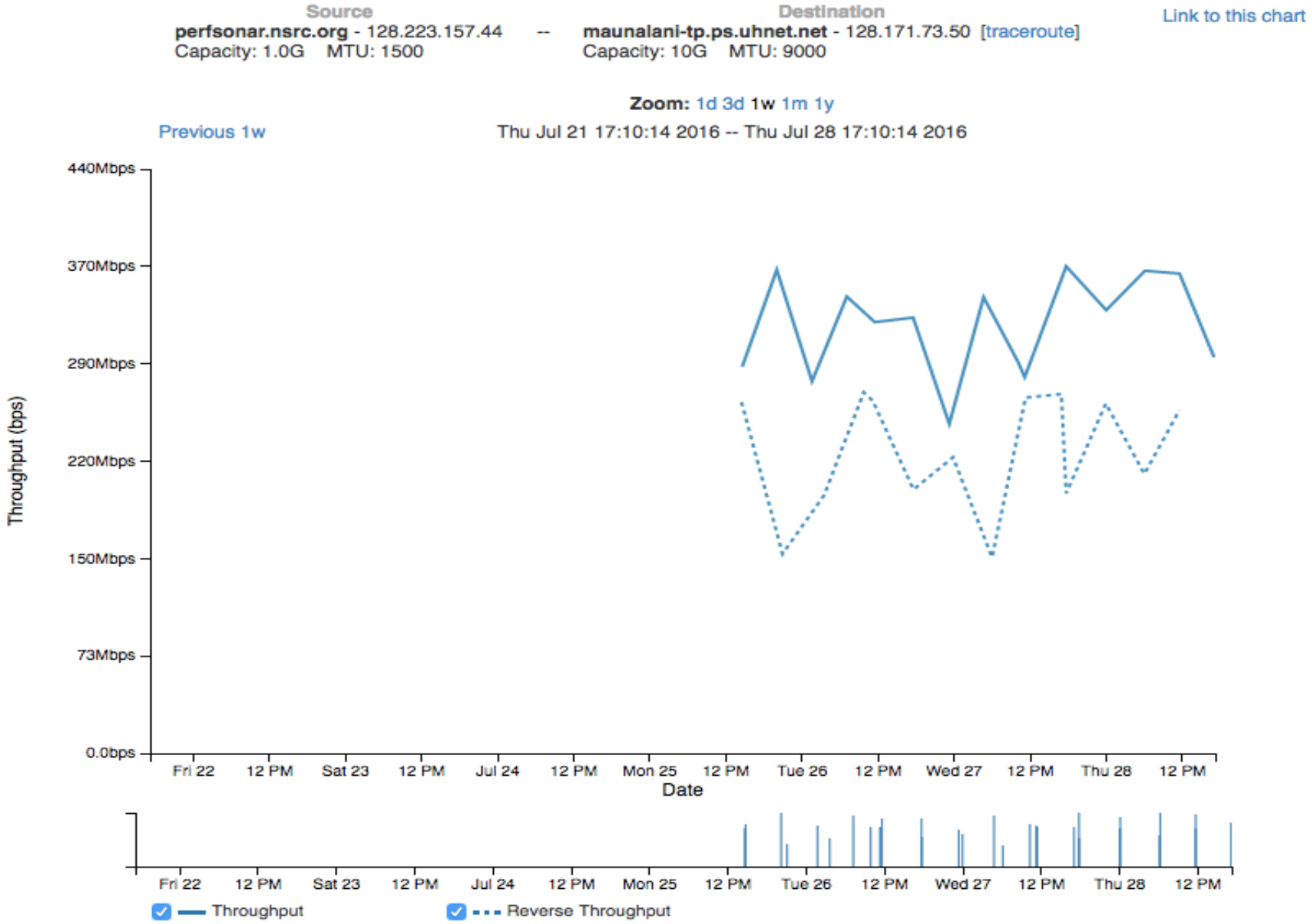
```
[root@perfsonar100 ~]# bwctl -fM -t 30 -s 202.28.229.115
bwctl: Using tool: iperf
bwctl: 37 seconds until test results available

RECEIVER START
------------------------------------------------------------
Server listening on TCP port 5204
Binding to local address 203.159.61.100
TCP window size: 0.08 MByte (default)
------------------------------------------------------------
[ 15] local 203.159.61.100 port 5204 connected with 202.28.229.115 port 38953
[ ID] Interval       Transfer     Bandwidth
[ 15]  0.0-30.1 sec   210 MBytes  6.96 MBytes/sec
[ 15] MSS size 1448 bytes (MTU 1500 bytes, ethernet)

RECEIVER END
```

- *Then automate this within web GUI and build historical trends...*

# Sample perfSONAR site

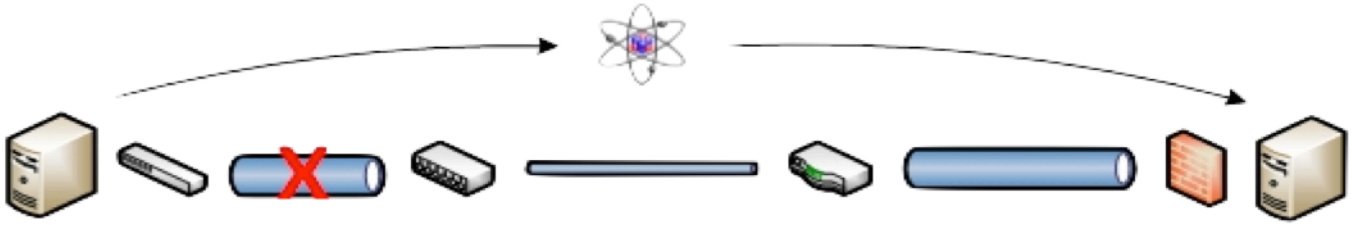


# Throughput Historical Trends



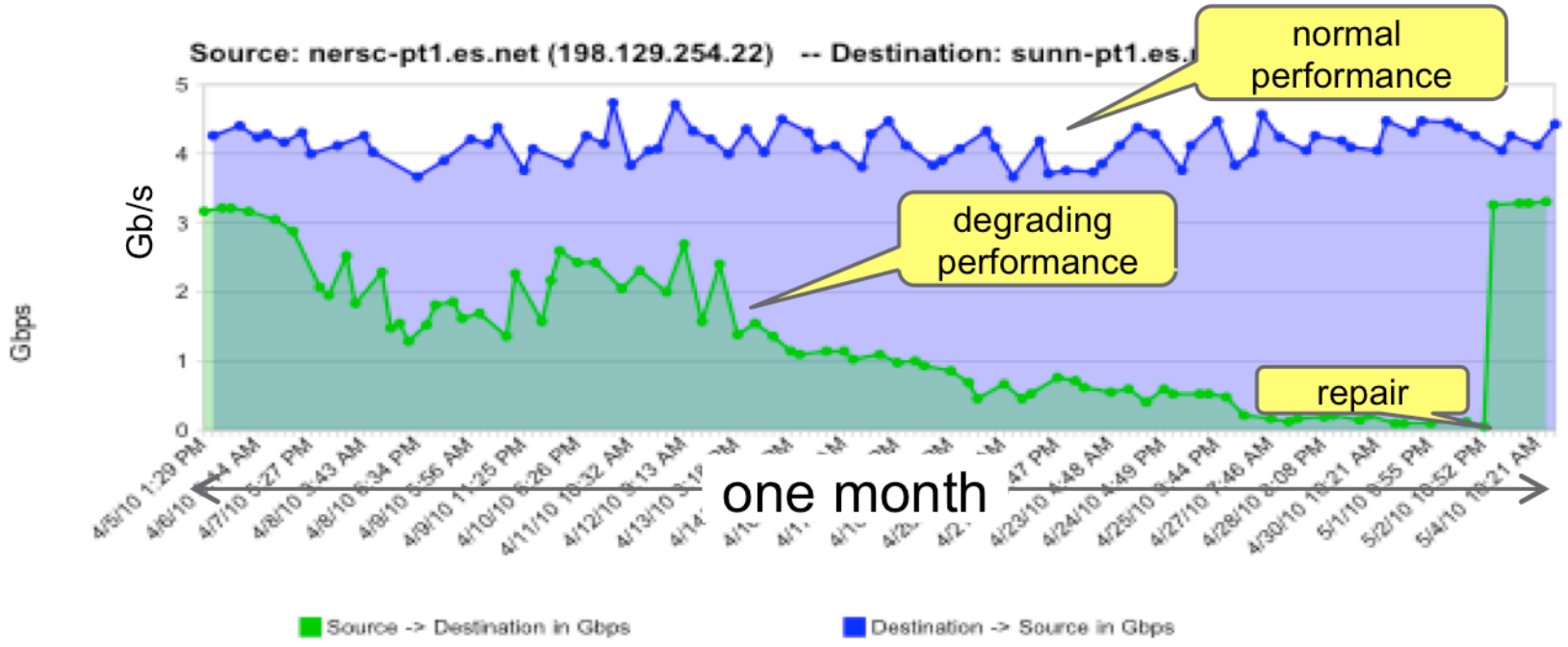
# Benefits: Finding hidden problems

- 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.



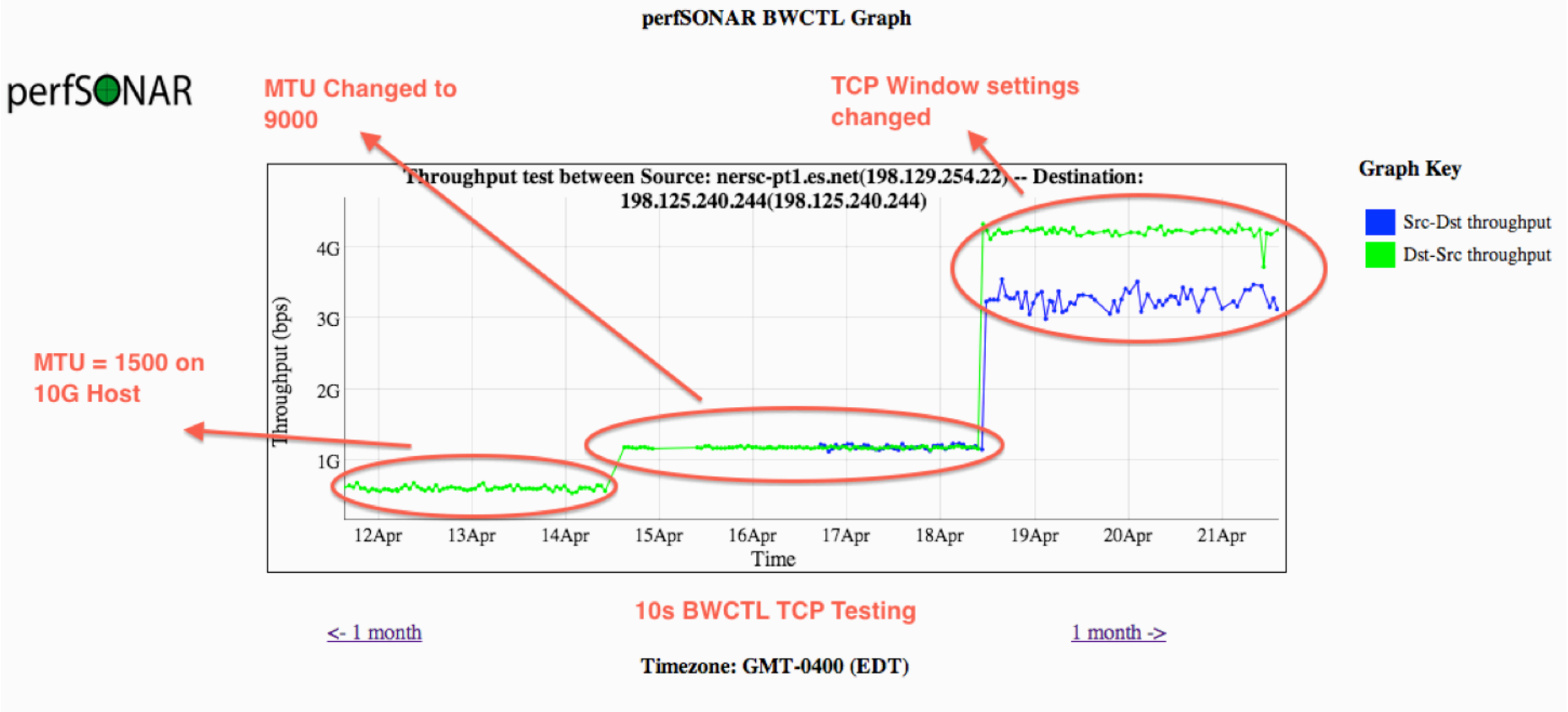
# Benefits: Finding hidden problems

- perfSONAR is designed to pinpoint and identify soft failures to accelerate resolution.
- Example: Find and replace failing optics:

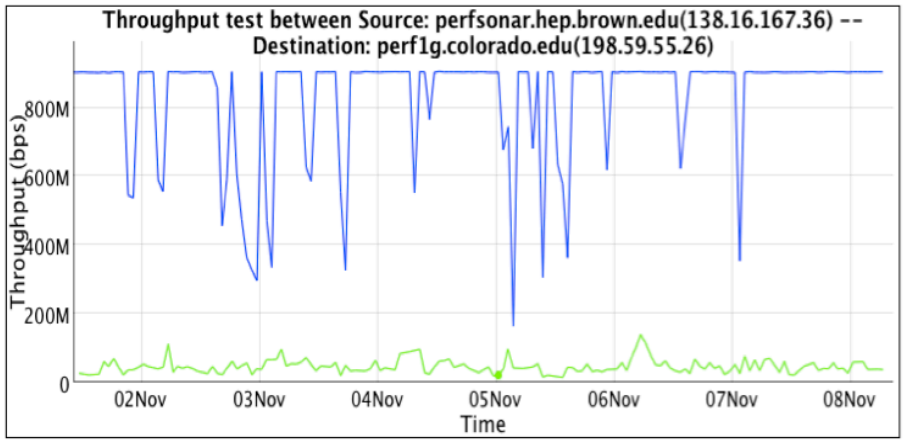


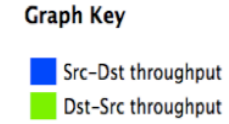
# Benefits: Finding hidden problems

- perfSONAR is designed to pinpoint and identify soft failures to accelerate resolution.
- Example: Host Tuning:



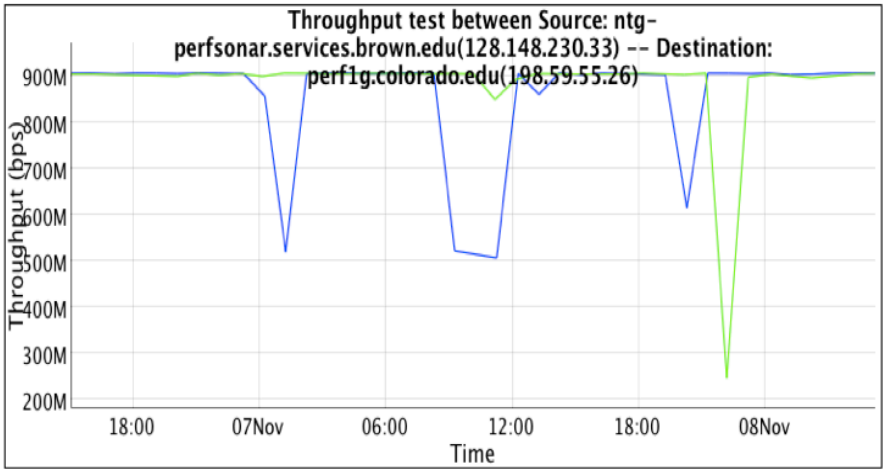
# Benefits: Finding hidden problems

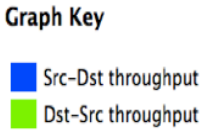



**Inside the firewall**

- One direction severely impacted by firewall
- Not useful for science data

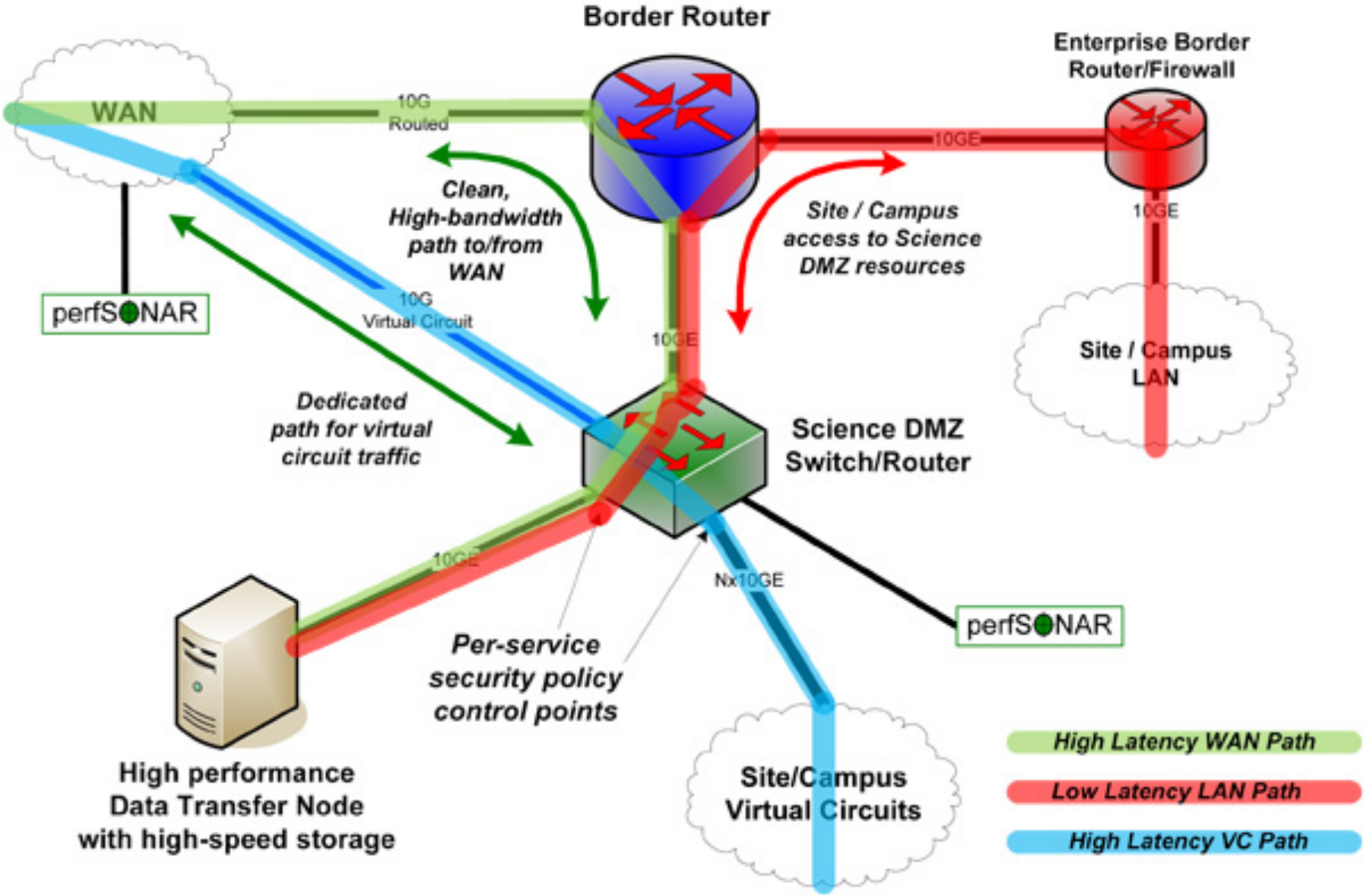



- **Outside the firewall**

- Good performance in both directions

# Firewalls and throughput



# Some local results
# powstream long-term-loss

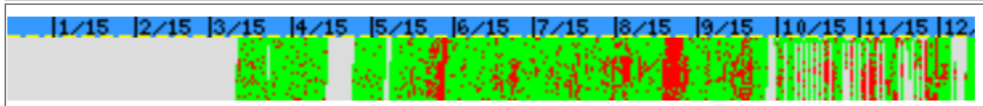
src: 172.25.91.228 dst: 192.168.71.246

| Router: networksl3-itc | |
|---|---|
| **Ingress** | **Egress** |
| 172.25.91.227 TenGigabitEthernet1/6 | 192.168.221.77 Vlan667 |
| physint: | physint: TenGigabitEthernet1/2 |
| speed dup | speed 10G dup full |

| Router: temp-manoa6504 | |
|---|---|
| **Ingress** | **Egress** |
| 192.168.221.89 Vlan667 | 192.168.76.250 Vlan3253 |
| physint: TenGigabitEthernet2/8 | physint: GigabitEthernet4/2 |
| speed 10G dup full | speed 1G dup full |

| Router: agscil3hub | |
|---|---|
| **Ingress** | **Egress** |
| 192.168.76.251 vlan.3253 | 192.168.44.234 vlan.3117 |
| physint: ge-0/0/1.0 | physint: xe-0/0/27.0 |
| speed 1G dup full | speed 10G dup full |

| Router: Mitchelll3bldg | |
|---|---|
| **Ingress** | **Egress** |
| 192.168.44.235 Vlan3117 | |
| physint: TenGigabitEthernet0/2 | physint: |
| speed 10G dup full | speed dup |




[GRAPH]

172.25.91.228 ps-node genpop01
TO >>>
dst: 192.168.71.246 armv7l-6c-ec-eb-ac-86-bc swarm-inactive
rtr: 192.168.71.192/26, MitchellL3Bldg Mitchell BET 4D Vlan442
swx: ,

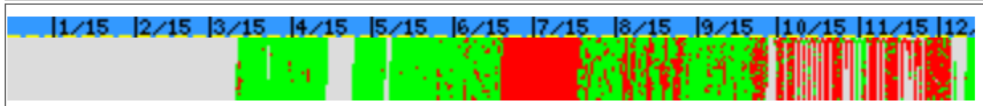
src: 192.168.71.246 dst: 172.25.91.228

| Router: Mitchelll3bldg | |
|---|---|
| **Ingress** | **Egress** |
| 192.168.71.193 Vlan442 | 192.168.44.235 Vlan3117 |
| physint: | physint: TenGigabitEthernet0/2 |
| speed dup | speed 10G dup full |

| Router: agscil3hub | |
|---|---|
| **Ingress** | **Egress** |
| 192.168.44.234 vlan.3117 | 192.168.76.253 vlan.3254 |
| physint: xe-0/0/27.0 | physint: xe-0/0/0.0 |
| speed 10G dup full | speed 10G dup full |

| Router: manoa6506-1 | |
|---|---|
| **Ingress** | **Egress** |
| 192.168.76.252 Vlan3254 | 192.168.221.80 Vlan667 |
| physint: TenGigabitEthernet5/8 | physint: TenGigabitEthernet5/5 |
| speed 10G dup full | speed 10G dup full |

| Router: networksl3-itc | |
|---|---|
| **Ingress** | **Egress** |
| 192.168.221.77 Vlan667 | |
| physint: TenGigabitEthernet1/2 | physint: |
| speed 10G dup full | speed dup |




[GRAPH]

src: 192.168.71.246 armv7l-6c-ec-eb-ac-86-bc swarm-inactive
rtr: 192.168.71.192/26, MitchellL3Bldg Mitchell BET 4D Vlan442
swx: ,
TO >>>
172.25.91.228 ps-node genpop01

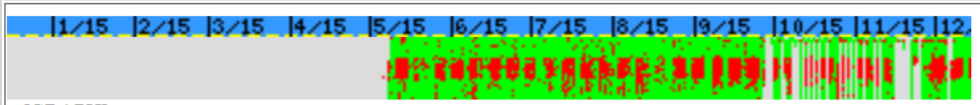
resolves router/interface information for bwtraceroute results by querying netdot

# powstream long-term loss

| Image | test |
|---|---|

src: 172.25.91.228 dst: 192.168.88.57

| Router: networksl3-itc | | Router: temp-manoa6504 | | Router: | | Router: wilsonl3bldg | |
|---|---|---|---|---|---|---|---|
| Ingress | Egress | Ingress | Egress | Ingress | Egress | Ingress | Egress |
| 172.25.91.227 TenGigabitEthernet1/6 | 192.168.221.77 Vlan667 | 192.168.221.89 Vlan667 | | 192.168.76.247 | | 192.168.44.77 Vlan3038 | |
| physint: | physint: TenGigabitEthernet1/2 | physint: TenGigabitEthernet2/8 | physint: | physint: | physint: | physint: | physint: |
| speed dup | speed 10G dup full | speed 10G dup full | speed dup | speed dup | speed dup | speed dup | speed dup |

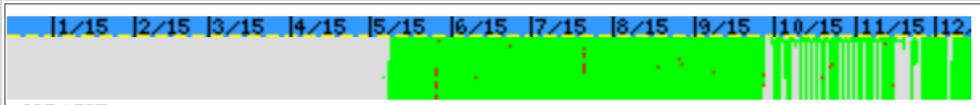



[GRAPH]

172.25.91.228 ps-node genpop01
TO >>>
dst: 192.168.88.57 rpi-140729-4 swarm-ACTIVE
rtr: 192.168.88.0/23, WilsonL3Bldg Wilson BET 5167 Vlan194
swx: WilsonBETsw1, FastEthernet0/9

src: 192.168.88.57 dst: 172.25.91.228

No trace results available

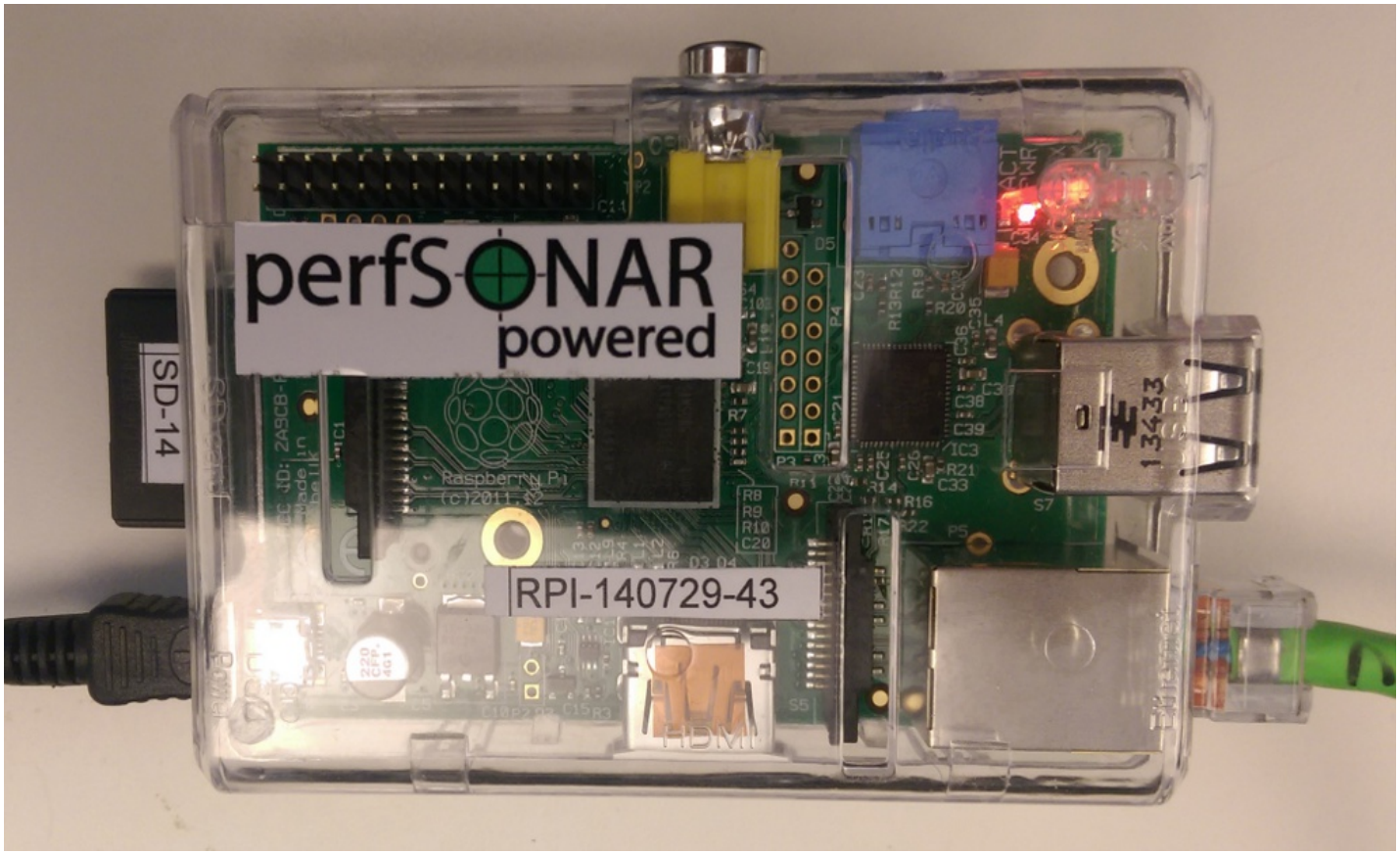



[GRAPH]

src: 192.168.88.57 rpi-140729-4 swarm-ACTIVE
rtr: 192.168.88.0/23, WilsonL3Bldg Wilson BET 5167 Vlan194
swx: WilsonBETsw1, FastEthernet0/9
TO >>>
172.25.91.228 ps-node genpop01

diurnal patterns in loss suggest mid-day congestion relationship

# Raspberry Pi Sensor Node
# U. Hawaii CCNIE “Swarm”



# Small ARM based devices

- **Raspberry Pi** – famous, $50, med-perf, file system on SD card, 100 Mb Ethernet, USB 2.0
- **BeagleBone Black** – $50, more perf, FS on internal flash, and/or SD card, 100 Mb, USB 2.0
- **CuBox i4** – $147, more perf, FS on SD, GigE, WiFi, USB 2.0
- **MiraBox** – $149 – most perf, FS on SD, dual GigE, WiFi, USB 3.0

# Small x86_64 devices

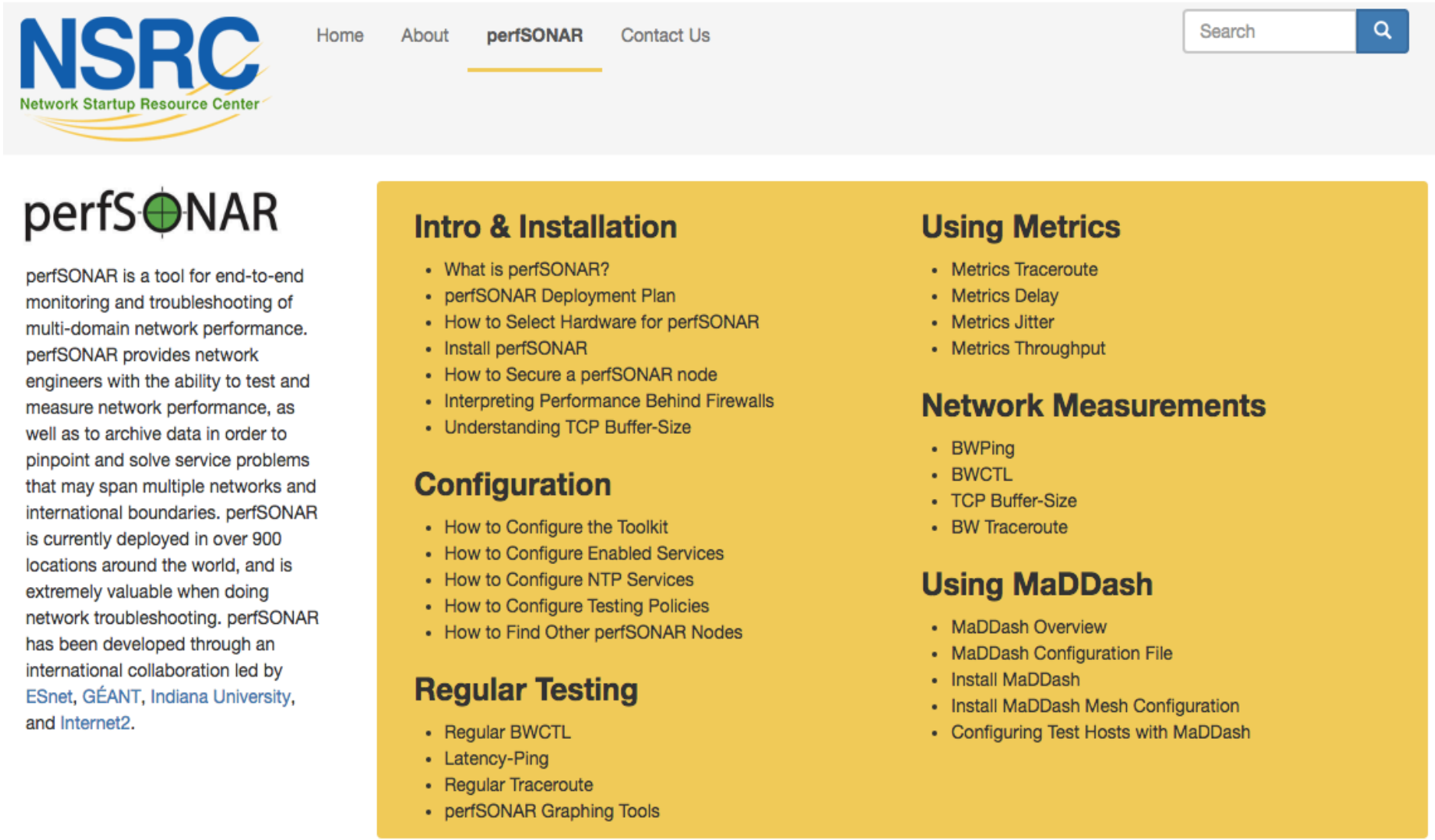
- **Liva X – gigabit Ethernet *****
  - 2 GB/32 GB $110
  - 4 GB/64 GB $290
  - (http://perfclub.org/?p=112 )
- **Gigabyte BRIX (GB-BXBT-2807)**
  - Gigabit Ethernet
  - 4 GB/120 GB - $225

# In summary

- perfSONAR helps you to test data and metrics verifies performance of current investments
  - Can you get full use out of the carrier circuit you're paying for?
  - If performance suddenly changes, what caused it?
- Data informs future investments
  - Put resources in places with demonstrated need
  - Example: if your firewall is limiting performance, buying more bandwidth won't solve the problem
- Make the case for architectural improvements
  - How might a Science DMZ or alternate architecture change performance?
  - How to evaluate potential purchases?
    - Put perfSONAR hosts on them and run tests
    - Prototype new model

# Resources

Online perfSONAR training videos



# Resources

- perfSONAR Website
  - http://www.perfsonar.net/
- perfSONAR mailing lists
  - http://www.personar.net/about/getting-help/
- perfSONAR Directory of Users
  - http://stats.es.net/Services
- perfSONAR Online Training Videos
  - https://learn.nsrc.org/perfsonar/
- FasterDate Knowledgebase
  - http://fasterdata.es.net/
- perfSONAR Checklist
  - https://code.google.com/p/perfsonar-ps/wiki/PerformanceToolkitTestingChecklist

# Primary perfSONAR Development



*For a full list of development collaborators and history of perfSONAR, please visit:*

http://www.perfsonar.net/about/who-is-involved/

# Questions?