perfSONAR Monitoring for LHCONE

Shawn McKee/University of Michigan

LHCONE Meeting – APAN

Nantou, Taiwan

August 13th, 2014

Overview of Talk

- Review: Current Status & Why are we doing this?
- Problems we have and must address
- Moving forward
- Summary, Question and Discussion

Review: Testing the LHCONE Network

- Why do we want the network tests between LHCONE sites?
  - Broadly: to identify problems on the network paths between sites
  - To find (to the extent possible) the location of problems
    - In time and physical or logical location
  - To alert when significant changes occur
  - To set expectations about what is possible and expected
  - To provide network metrics to existing and future services
- We have been using the **perfSONAR** because it is standardized, supported broadly in R&E networks (and now many regionals and sites) and provides scheduled measurement of standard network metrics
- The **perfSONAR-PS toolkit** provides both a scheduled test capability as well as “on-demand” testing and is easy to deploy via ‘netinstall’ or by booting from CDROM

Review of Where We Are

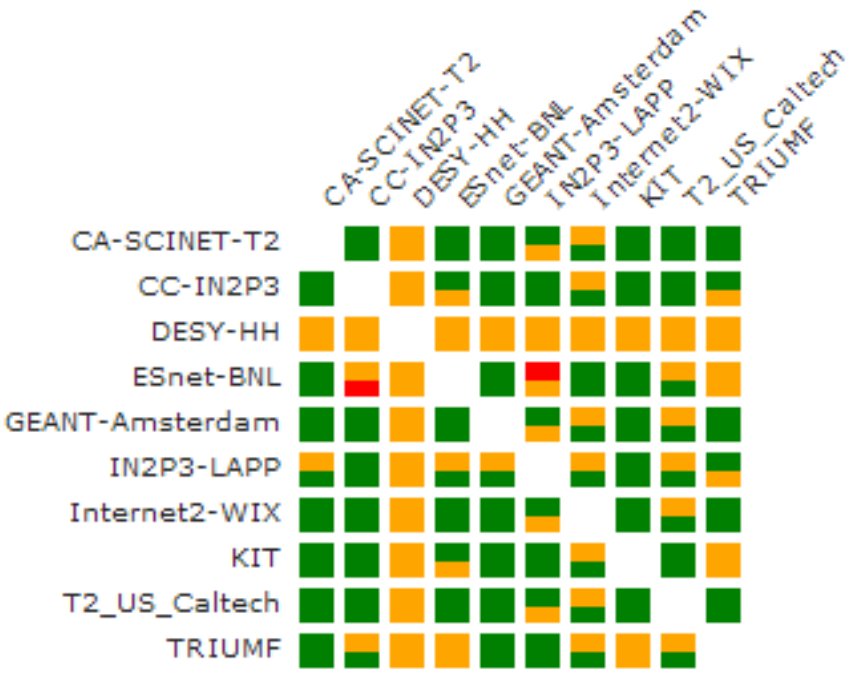
- A brief review of where we are and how we got here:
  - LHCOPN adopted perfSONAR-PS toolkit in June 2011; deployed by September 2011
  - WLCG deployed perfSONAR-PS toolkit at ALL sites (April 1 2014)
  - LHCONE needed specific monitoring:
    - How are things working?
    - Use of 100G ANA circuit…things OK or not?
    - proposed using the same system as WLCG **for this specific use case**
    - **Use existing sites + instrument specific PoPs for LHCONE**
  - Currently we have 13 "LHCONE sites" with a full-mesh of bandwidth and latency tests: http://maddash.aglt2.org/maddash-webui/index.cgi?dashboard=LHCONE%20testing%20sites
- Even at this (sampled) scale we have challenges
  - Bad: Too much "orange" (missing measurements)…
  - Good: That this may be identifying issues for LHCONE (firewalls,routing)

LHCONE perfSONAR-PS Setup

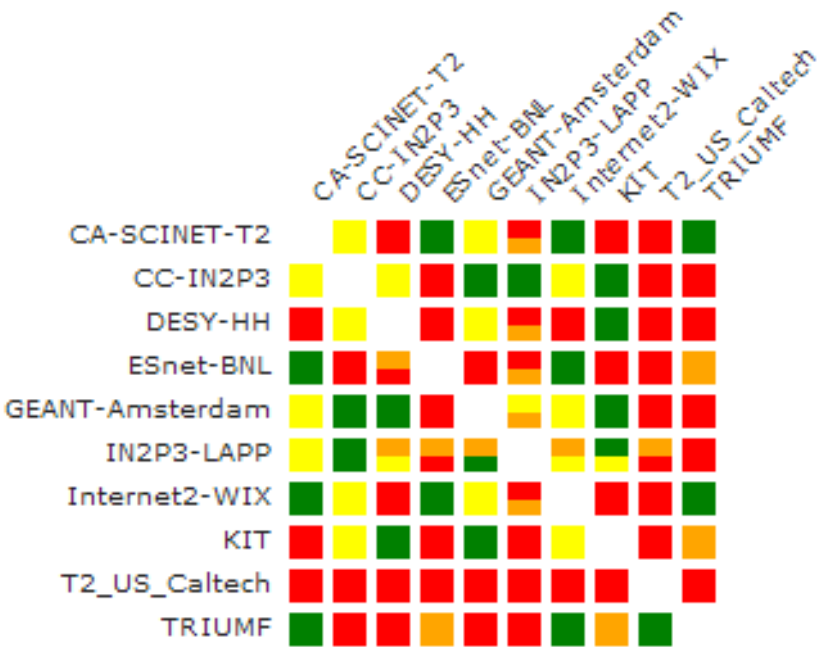
- We want to measure (to the extent possible) the entire network path between representative LHCONE sites.
  - We want to locate perfSONAR-PS instances as close as possible to the storage resources associated with a end-site. Goal is to ensure we are measuring the same network path to/from the storage.
  - Selected network PoPs should also be instrumented
- There are two separate instances that should be deployed: latency and bandwidth
  - The **latency instance** measures one-way delay by using an NTP synchronized clock sending 10 Hz packets to each target
  - The **bandwidth instance** measures achievable bandwidth via a short test (20-60 seconds) per src-dst pair every 2 hour period
- We also run a traceroute to all partner sites 1/hour
  - Critical for understanding which path was active during tests

LHCONE Network Matrices: 28Apr2014

OWAMP (Latency)



BWCTL (Bandwidth)



No packet loss, packet loss>0.01

BW>0.9 Gb, 0.5<BW<0.9 Gb, BW<0.5 Gb

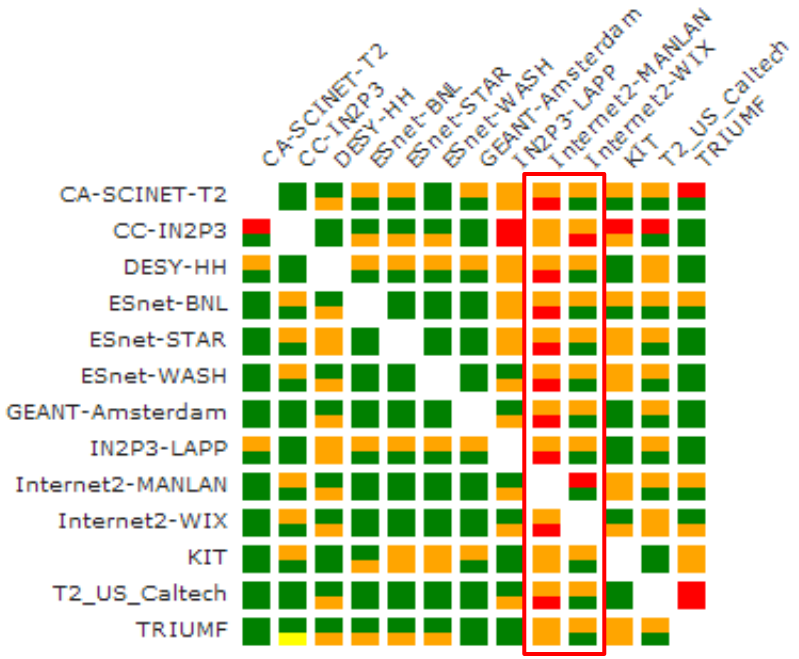
Main issue was too much "orange" indicating missing measurements/data

Sources are "row", Destination is "column"

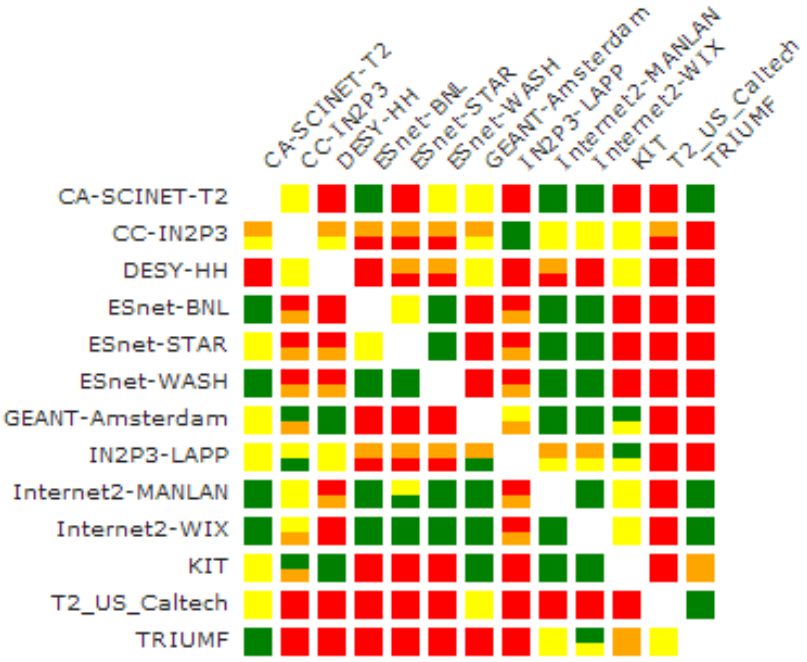
Each box split into two regions indicating where the test is run: top corresponds to "row", bottom to "column"

LHCONE Network Matrices: 11Aug2014

OWAMP (Latency)



BWCTL (Bandwidth)



No packet loss, packet loss>0.01

BW>0.9 Gb, 0.5<BW<0.9 Gb, BW<0.5 Gb

Main issue is STILL too much "orange" indicating missing measurements/data

Sources are "row", Destination is "column"

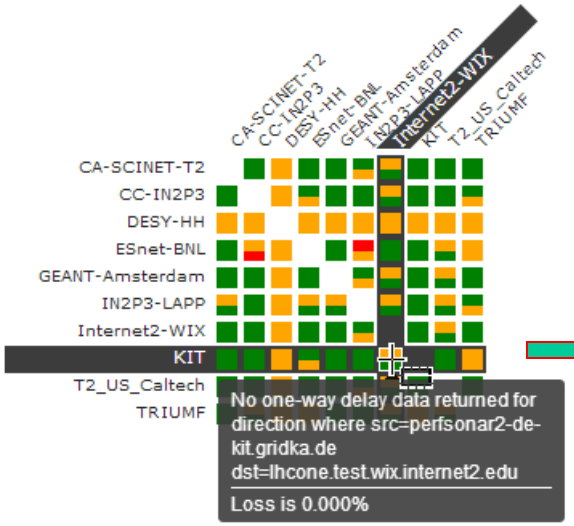
Each box split into two regions indicating where the test is run: top corresponds to "row", bottom to "column"

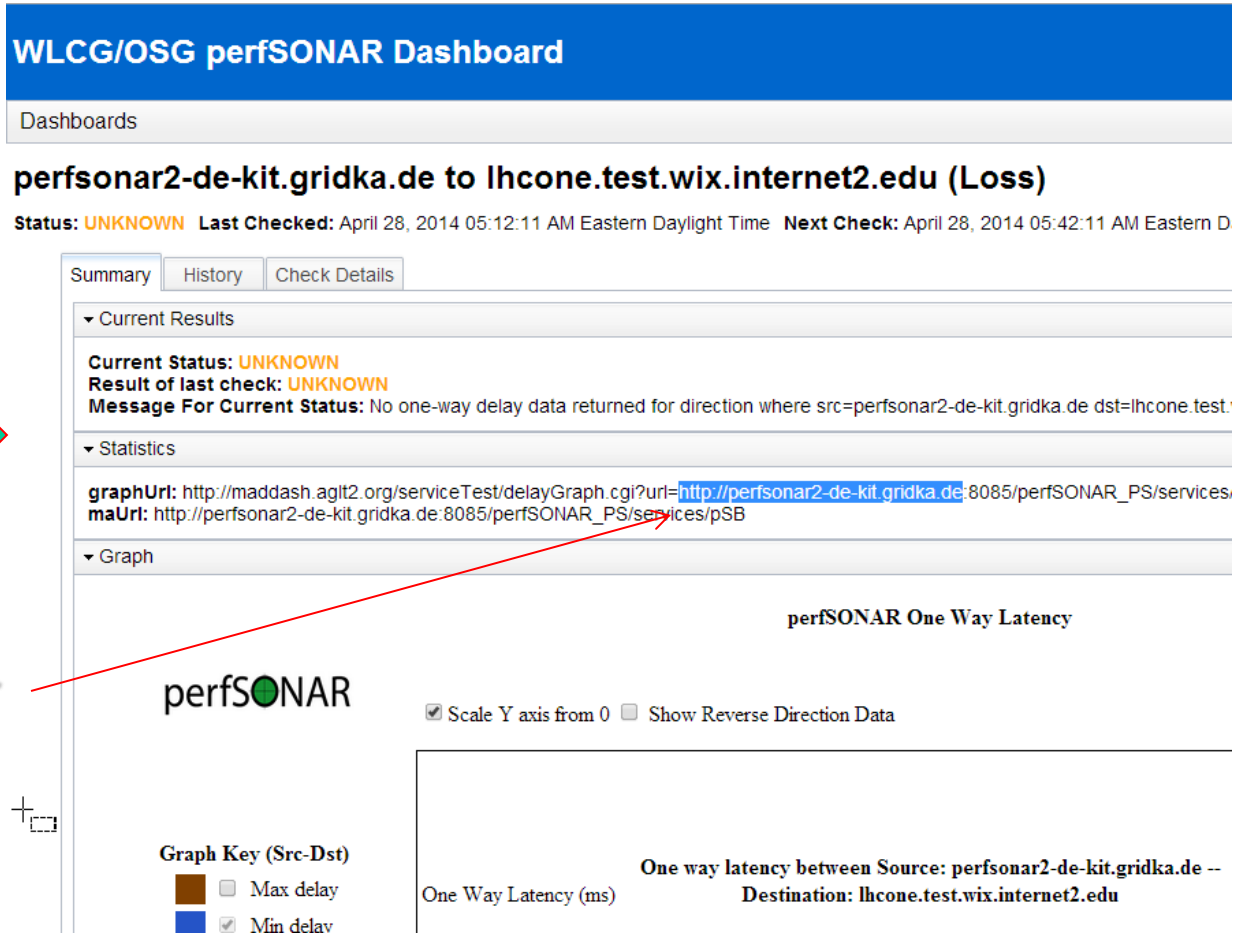
Debugging MaDDash Orange

- When MaDDash shows orange it indicates missing data. Why is the data missing?:
  1. Test for specific metric, src and dst not configured?
  2. Test unable to run? (Service down at src/dst, config or Firewall?)
  3. Result not stored (MA service down?)
  4. Result not retrievable (MA service blocked from MaDDash?)
- Remember each box in MaDDash represents a specific test (latency, packet-loss, bandwidth) between the source (row) and destination (column).
  - Top of the box is test run by "row"
  - Bottom of the box is test run by "column" (Same test src-dst)
- Let's show an example debug session…

1) Checking if Test Defined

- First we verify the test is actually configured on the MA we are querying





- Get URL base for
- Toolkit

Check Toolkit Page

Cut-n-paste base URL into browser

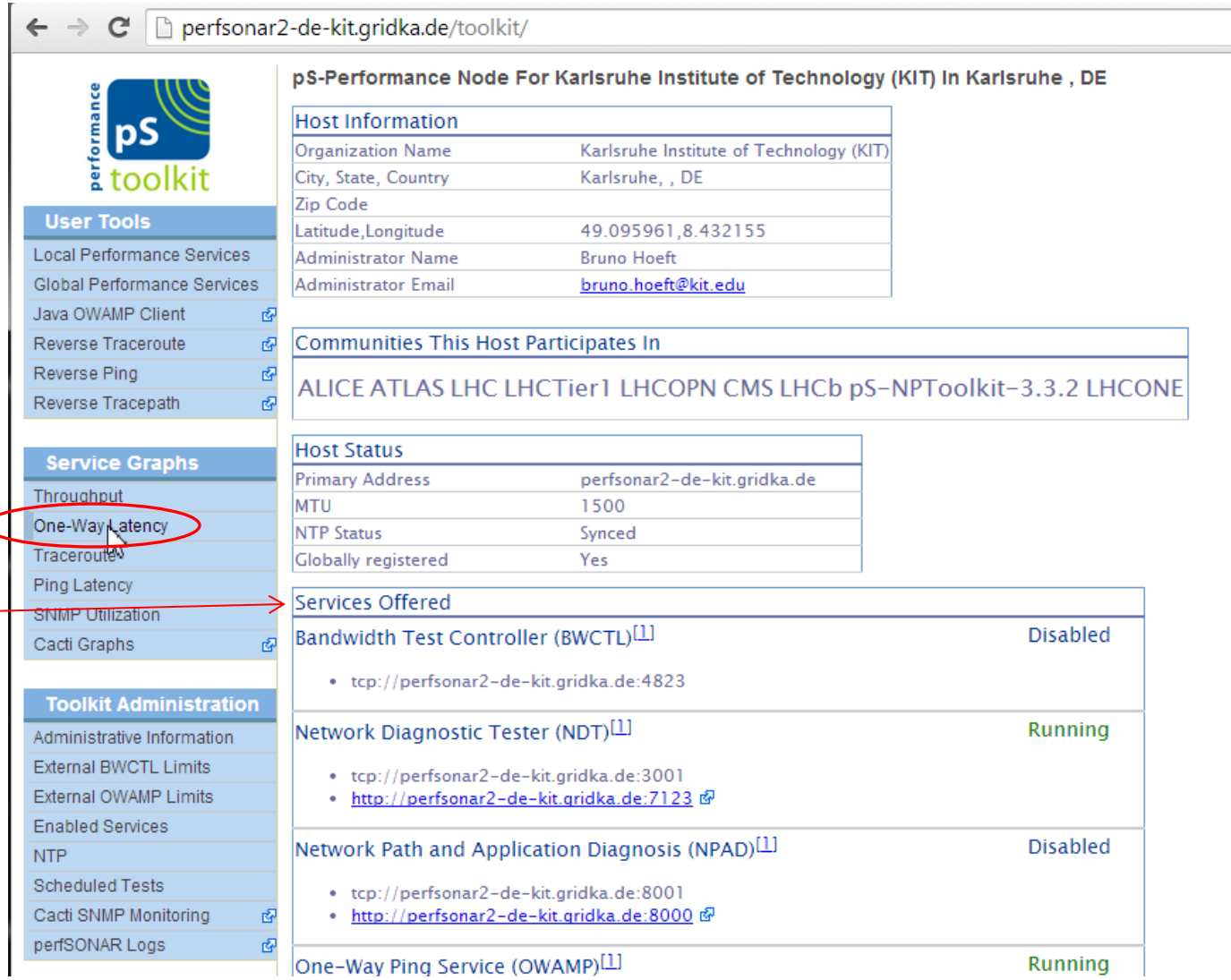
Append /toolkit

Go to homepage (*This should be in MaDDash*)

Check Latency page (click One-Way Lat.)

But first check services are running (next page)

NOTE: it is critical we have access to the PS Toolkit Web UI; it is an import debug tool



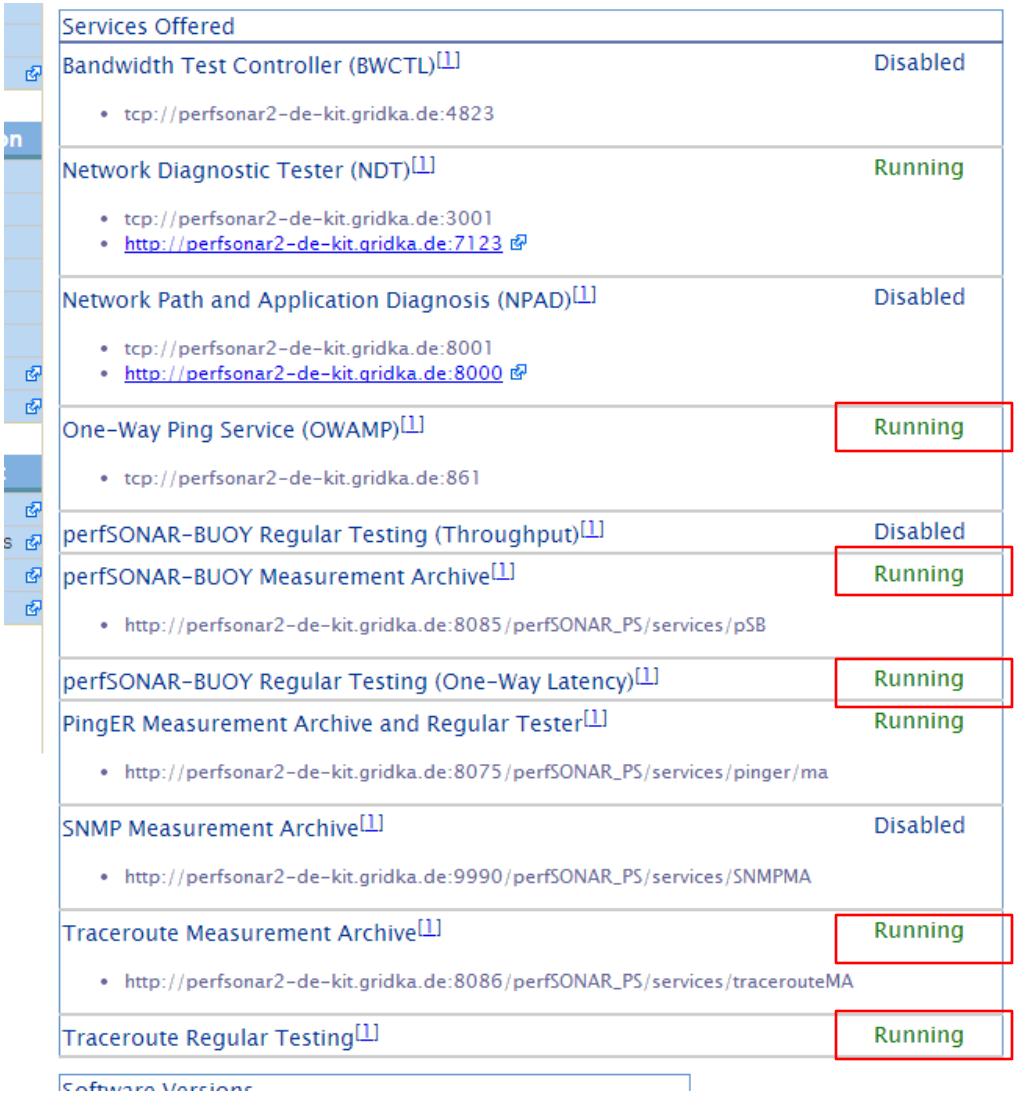
2/3) Verify Expected Services are Running

On homepage for toolkit you can check all expected services are "Running" (Green)

Services that should be running and are NOT show red "Not Running"

For OWAMP measurements we need the three services shown in red boxes to be Running

Latency nodes also run the traceroute services (also in red boxes)



1) Verify Active Test Exists

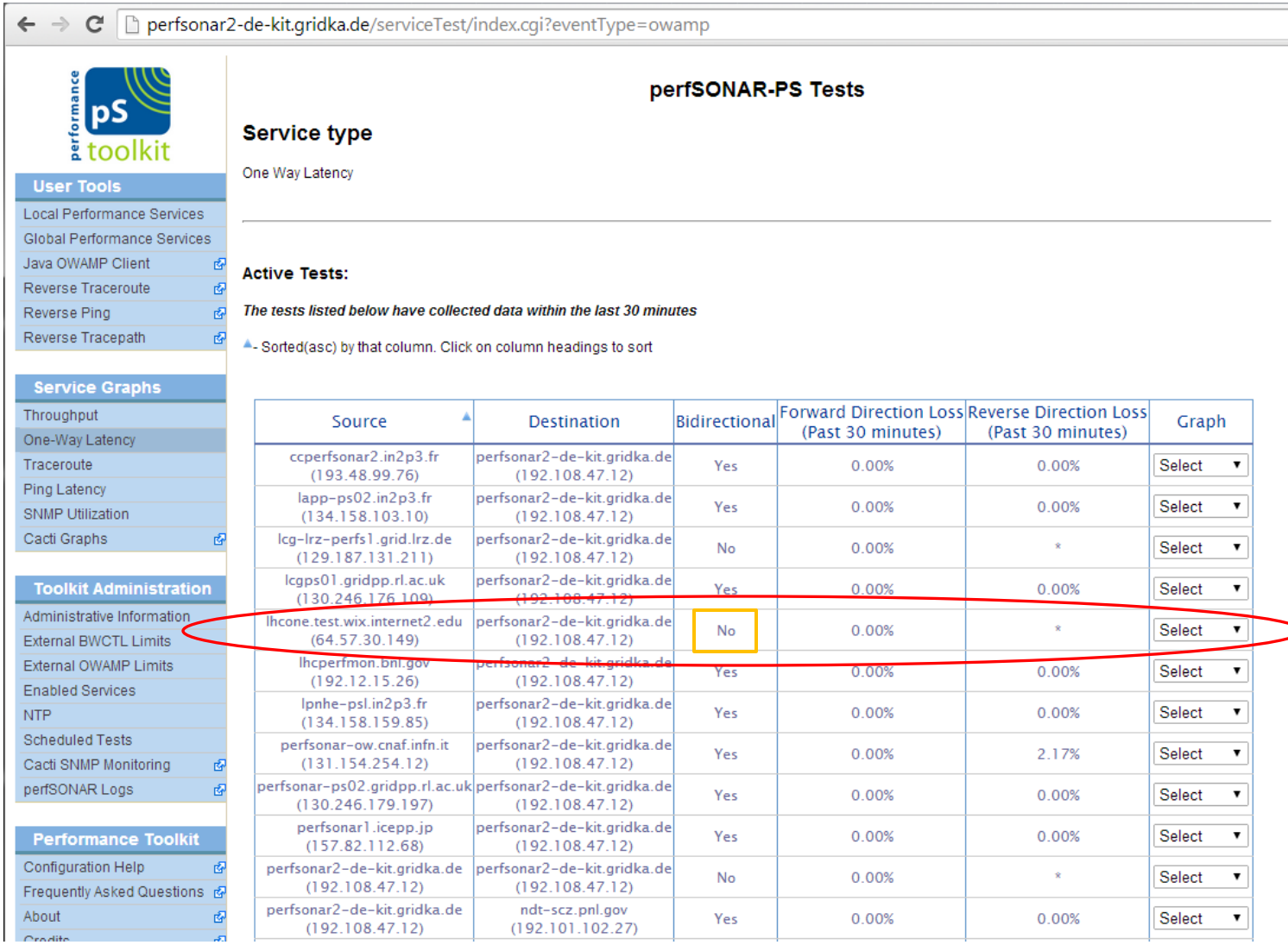
Active test **IS** defined on KIT toolkit between WIX and KIT

However the test is NOT Bidirectional

Forward direction WIX->KIT is working

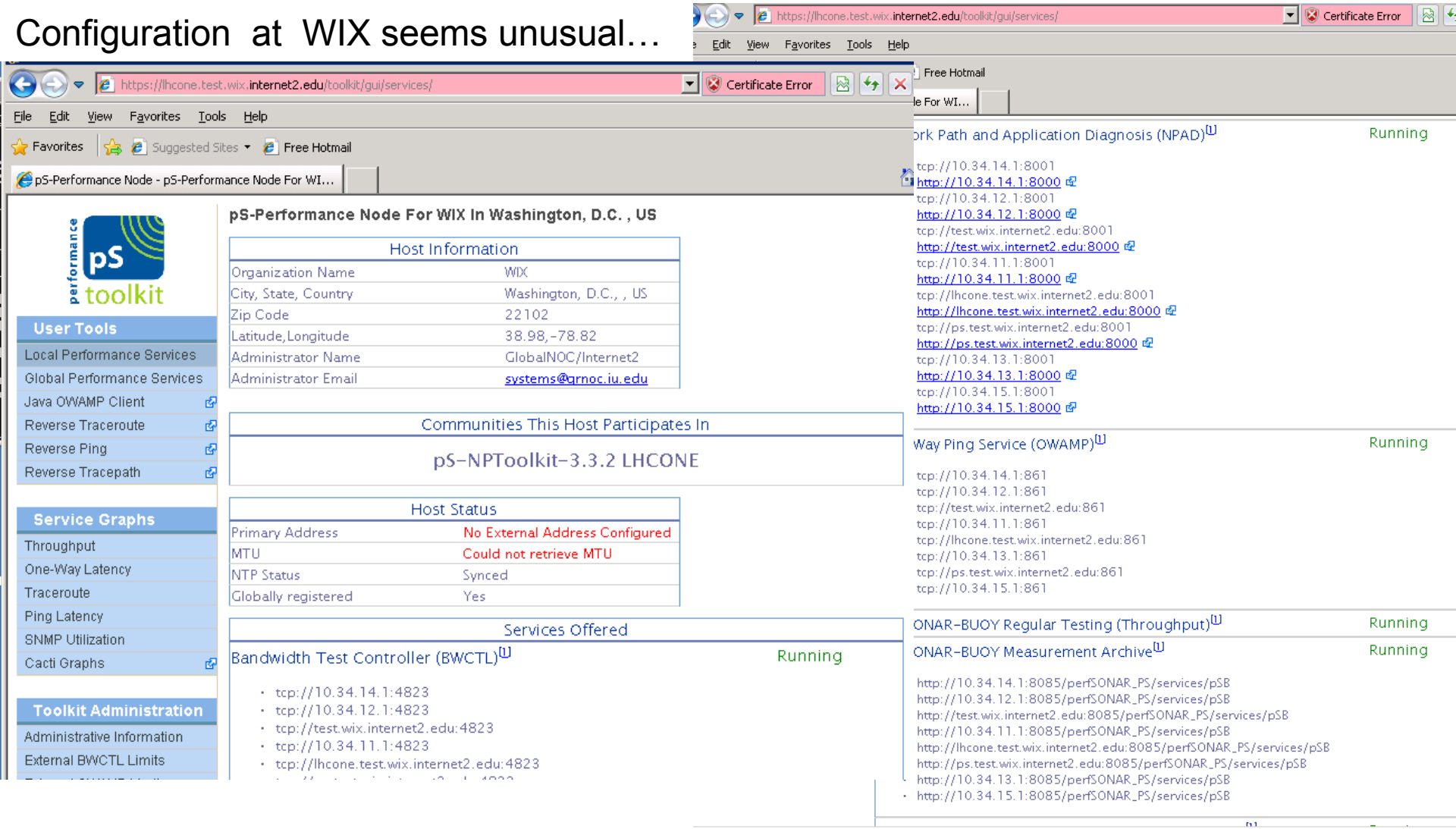
Reverse direction KIT->WIX is NOT

Why?? Still TBD



Services at WIX Seem OK…But

Configuration at WIX seems unusual…



Not sure why this test is bi-directional! Needs further work*

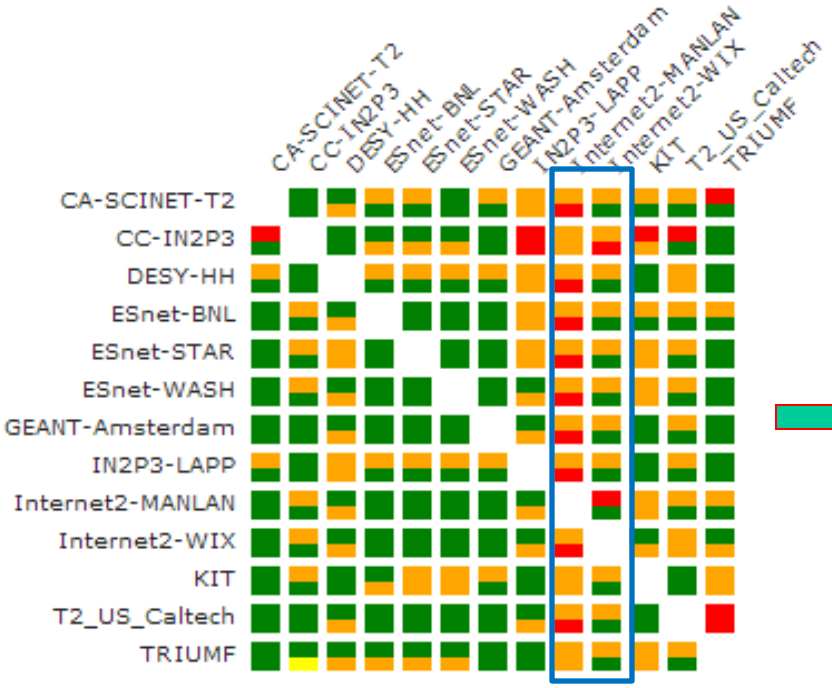
Resolution of WIX Issue

- Since the Rome LHCONE/LHCOPN meeting we have made progress on debugging WIX (and MANLAN)
  - Problem was two-fold:
    - Firewall misconfiguration (that was suspected at Rome)
    - Service configuration problem (latency measurements had too restrictive a configuration to support our level testing)
  - Fixed a few hours ago ☺
- What do we learn? Careful debugging is often needed, especially for custom installs.
  - WIX and MANLAN instances were "specially" installed, not using the standard procedure.
  - Firewalls are a continuing problem.
- **GOAL**: Let's get the remaining "orange" fixed before the Ann Arbor meeting

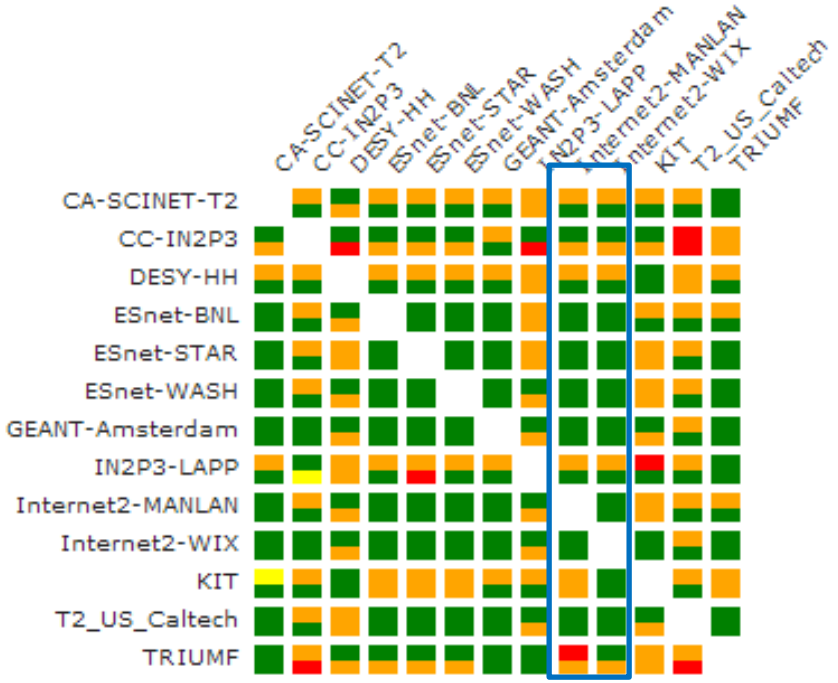
Current LHCONE Latency Mesh

Just to show…the "fixes" worked. Comparing Aug 11 with today

LHCONE OWAMP Mesh Aug 11



LHCONE OWAMP Mesh Aug 13



perfSONAR-PS Command Line Tools

- We have RPMs which provide command line tools for perfSONAR-PS documented at: https://twiki.grid.iu.edu/bin/view/Documentation/Release3/NetworkPerformanceToolkit
- Important tools for verifying functionality
  - Test bandwidth: **bwctl –s <sourcehost> -c <clienthost>**
  - Test latency: **owping <latencyserver>**
- All the perfSONAR-PS Toolkit installs should already have these tools.
- Useful to do quick tests as you **debug**, **change firewall settings**, **change service configurations**, **restart services**.

Questions Pointing at Problems

- What is the relevant, useful data we need to acquire?
  - We think we know (bw, latency, packet loss, routing)…do we?
- Is the scale of network testing implied by extrapolating our current efforts forward feasible?
  - It doesn't seem to be. Too many tests for timescales involved!
- The manageability of the current setup: How much work does it to organize adding/removing/updating sites?
  - Too much effort; too long to make changes
- How much effort is required at end-sites to keep services up and running?
  - Much more than many sites are willing to provide!
- Can the modular dashboard keep up with the large number of sites and measurements?
  - Barely, being redesigned to scale-out "wide" as required

Common Problems with perfSONAR-PS

- When things aren't working there are a few common problems to check for:
  - Services configured to run (Web UI Enabled Services)
  - Running out of disk space (logging or data (OWAMP))
  - Upgrade fills /boot (system crashes on reboot)
  - Limits are not configured correctly for LHCONE
    - OWAMP ports (must configure /etc/owampd/owampd.conf+iptables)
    - BWCTL default max time limits (should be 30s+; maybe 61s is good)
  - Firewalls blocking needed access for tests or result retrieval (use CLI tools to help debug)
  - Service unexpectedly stopping / not running (examine /var/log/perfsonar/*.log files)
  - Node not using mesh-configuration (examine Scheduled Tests)
- Use OMD to check basic services (next slide)

Monitoring Status

- MaDDash instance at http://maddash.aglt2.org/maddash-webui
  - **Has shown we still have some issues: Too much "orange" meaning data is either not be taken (configuration or firewall) or access to results are blocked**

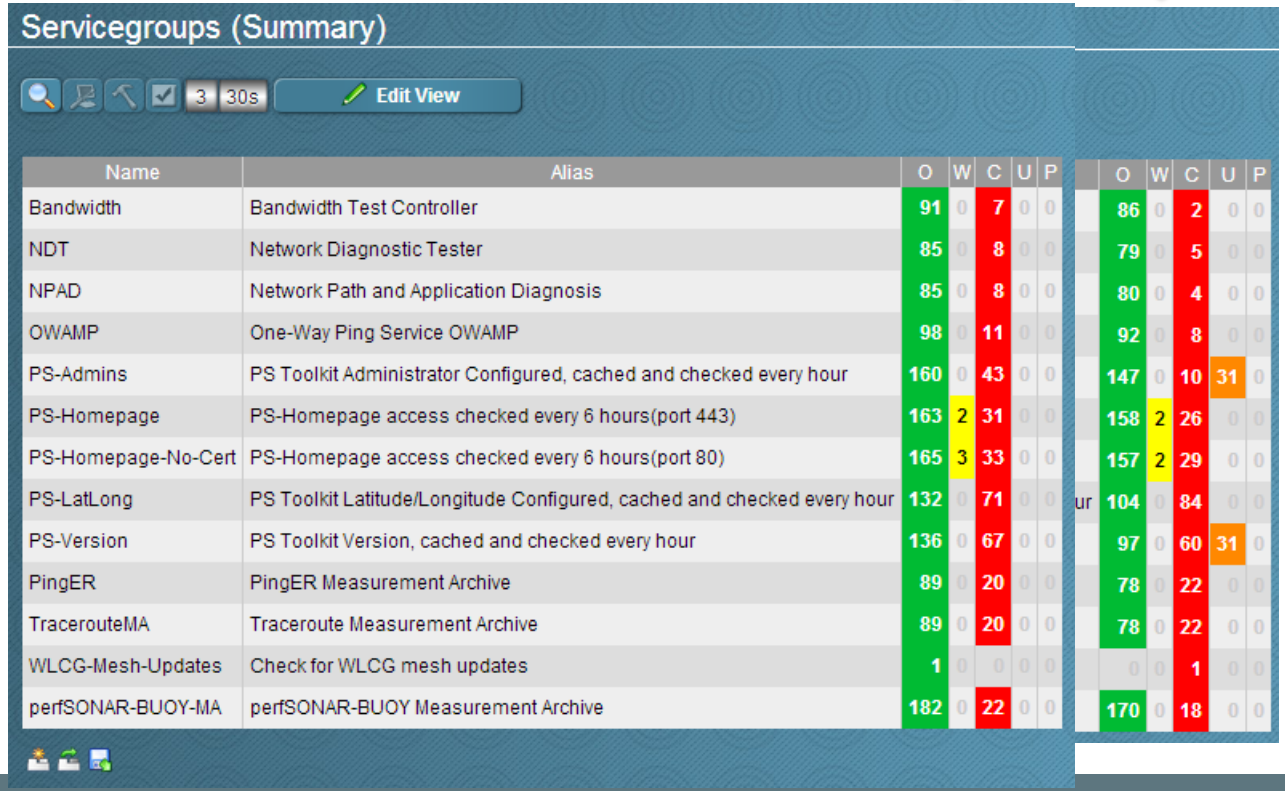
Slide From WLCG deployment TF Final Report
https://indico.cern.ch/event/309125/

April 16 March 6

Have OMD monitoring the perfSONAR-PS instances https://maddash.aglt2.org/WLCGperfSONAR/omd/

These services should migrate to OSG by fall.

This monitoring should be useful for any future work to find/fix problems.



Name	Alias	O	W	C	U	P
Bandwidth	Bandwidth Test Controller	91	0	7	0	0
NDT	Network Diagnostic Tester	85	0	8	0	0
NPAD	Network Path and Application Diagnosis	85	0	8	0	0
OWAMP	One-Way Ping Service OWAMP	98	0	11	0	0
PS-Admins	PS Toolkit Administrator Configured, cached and checked every hour	160	0	43	0	0
PS-Homepage	PS-Homepage access checked every 6 hours(port 443)	163	2	31	0	0
PS-Homepage-No-Cert	PS-Homepage access checked every 6 hours(port 80)	165	3	33	0	0
PS-LatLong	PS Toolkit Latitude/Longitude Configured, cached and checked every hour	132	0	71	0	0
PS-Version	PS Toolkit Version, cached and checked every hour	136	0	67	0	0
PingER	PingER Measurement Archive	89	0	20	0	0
TracerouteMA	Traceroute Measurement Archive	89	0	20	0	0
WLCG-Mesh-Updates	Check for WLCG mesh updates	1	0	0	0	0
perfSONAR-BUOY-MA	perfSONAR-BUOY Measurement Archive	182	0	22	0	0

	O	W	C	U	P
	86	0	2	0	0
	79	0	5	0	0
	80	0	4	0	0
	92	0	8	0	0
	147	0	10	31	0
	158	2	26	0	0
	157	2	29	0	0
ur	104	0	84	0	0
	97	0	60	31	0
	78	0	22	0	0
	78	0	22	0	0
	0	0	1	0	0
	170	0	18	0	0

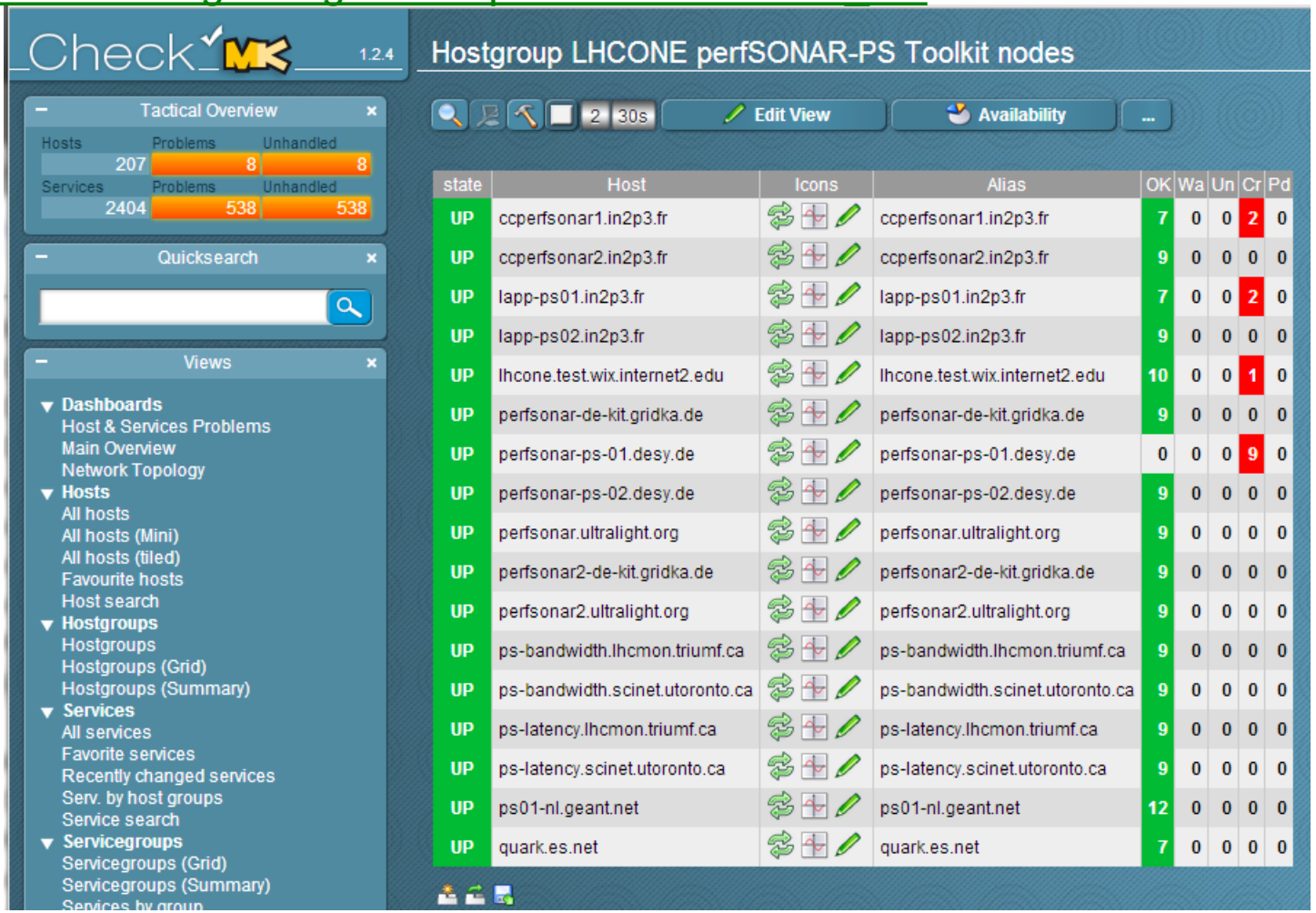
OMD for LHCONE perfSONAR-PS

http://maddash.aglt2.org/WLCGperfSONAR/check_mk

OMD (Open Monitoring Distribution) wraps a set of Nagios packages into a single pre—configured RPM

User WLCGps

Pw at meeting ☺



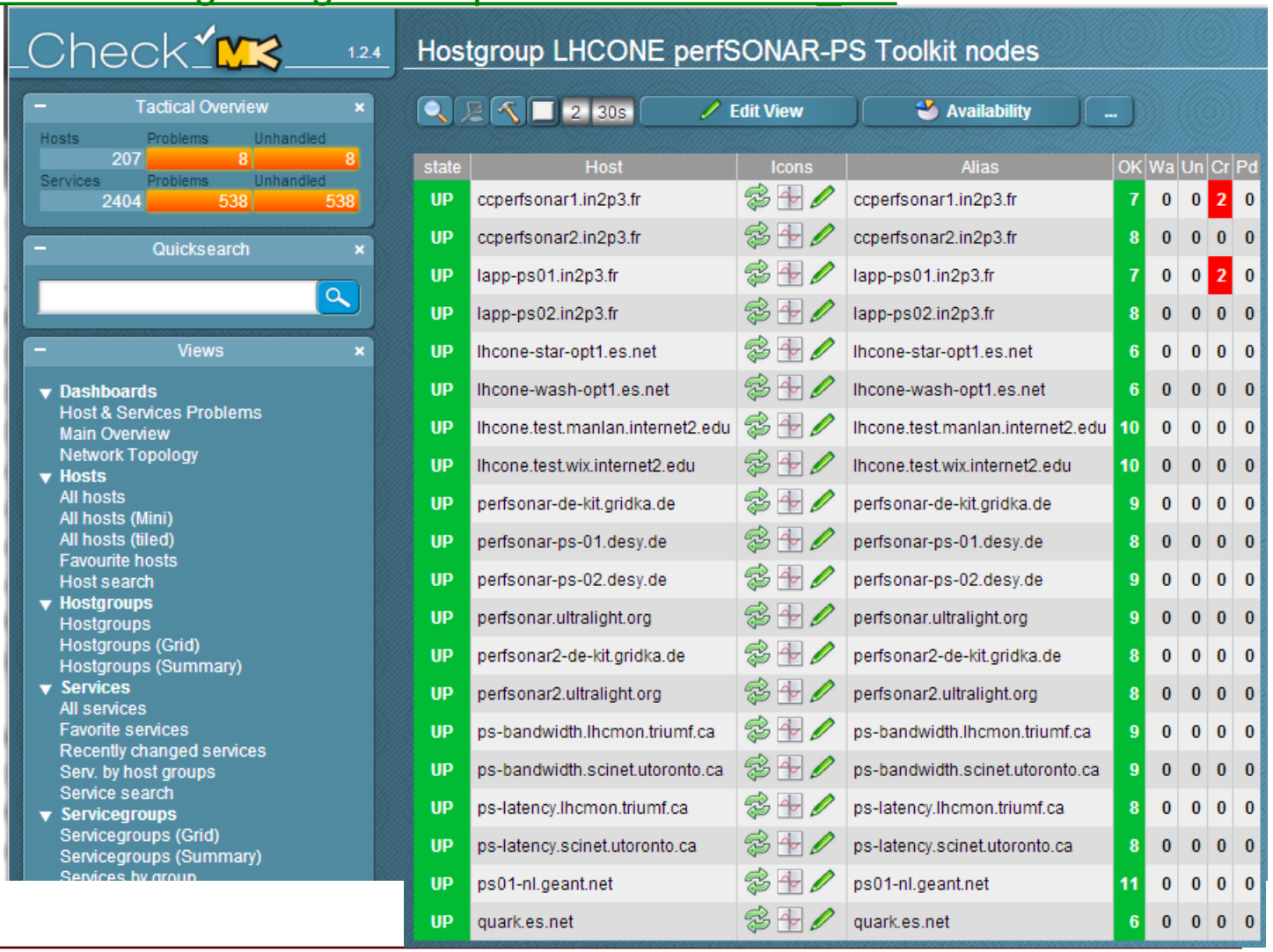
OMD for LHCONE perfSONAR-PS

http://maddash.aglt2.org/WLCGperfSONAR/check_mk

OMD (Open Monitoring Distribution) wraps a set of Nagios packages into a single pre—configured RPM

User WLCGps

Pw at meeting ☺



WLCG perfSONAR-PS Observations

- Getting working monitoring deployed is a **big** part of the battle.
  - Focusing on a set of inter-site monitoring configuration raised awareness of current shortcomings in our LHCONE infrastructure
- Two common primary problems we seem to have:
  - Firewalls block tests or MA access
  - **Services not running or not correctly configured**
- Issue with MTU setting. Suggestion for LHCONE is to use jumbo frames. We need to understand the impact on our measurements and our infrastructure.
- Test durations: 1G vs 10G. 30 seconds OK for 1G, but what about 10G? 60 seconds seems more reasonable.
- Getting alerts running: Issues with false positives.
- Higher level alarms: when, how?
- MaDDash dashboard: intro, use, future, issues.

Moving Forward for LHC(ONE) Monitoring

- Our goal should be to remove ALL orange from our LHCONE (WLCG subset) of monitoring sites
  - All LHCONE testing sites should work on this until we are consistently getting data from all scheduled tests!
- We can tweak test settings in the future to optimize
- **Adding an LHCONE test instance (or two) in Asia needs to be done.**
- Gaining experience using the metrics we are collecting
  - What is most useful?
  - What are typical use-cases for finding/fixing problems? (document)
  - Are changes needed in existing tests?
  - Are new tests providing different metrics required?

Useful URLs

- LHCOPN instructions for perfSONAR-PS (out-of-date):
  - https://twiki.cern.ch/twiki/bin/view/LHCOPN/PerfsonarPS
- LHCONE “initial” monitoring setup page
  - https://twiki.cern.ch/twiki/bin/view/LCG/PerfsonarDeployment
- Open Science Grid Networking URL
  - https://www.opensciencegrid.org/bin/view/Documentation/NetworkingInOSG
- perfSONAR tools, tips and best practices
  - http://www.usatlas.bnl.gov/twiki/bin/view/Projects/LHCperfSONAR
- MaDDash Monitoring
  - http://maddash.aglt2.org/maddash-webui/index.cgi?dashboard=LHCONE%20testing%20sites
- OMD Monitoring
  - https://maddash.aglt2.org/WLCGperfSONAR/check_mk/index.py?start_url=%2FWLCGperfSONAR%2Fcheck_mk%2Fview.py%3Fview_name%3Dhostgroups

Discussion/Questions/Comments?

There is a lot to consider.

I hope we have time for questions, discussion and comments.

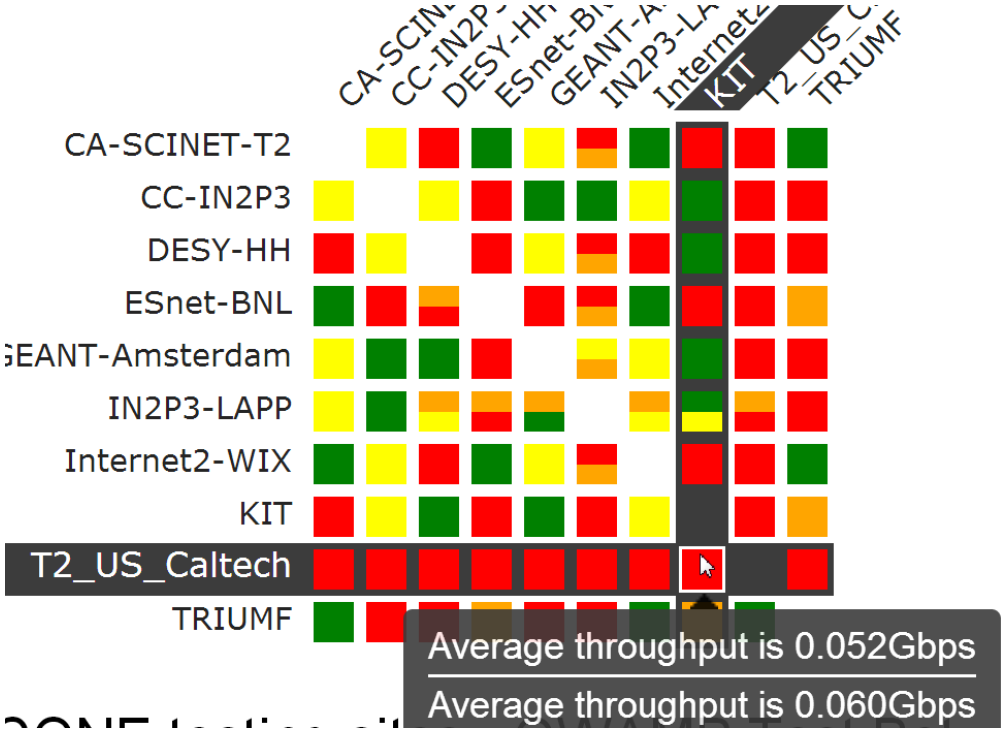
We could also do some online debugging/exploring…

Network Impact of perfSONAR

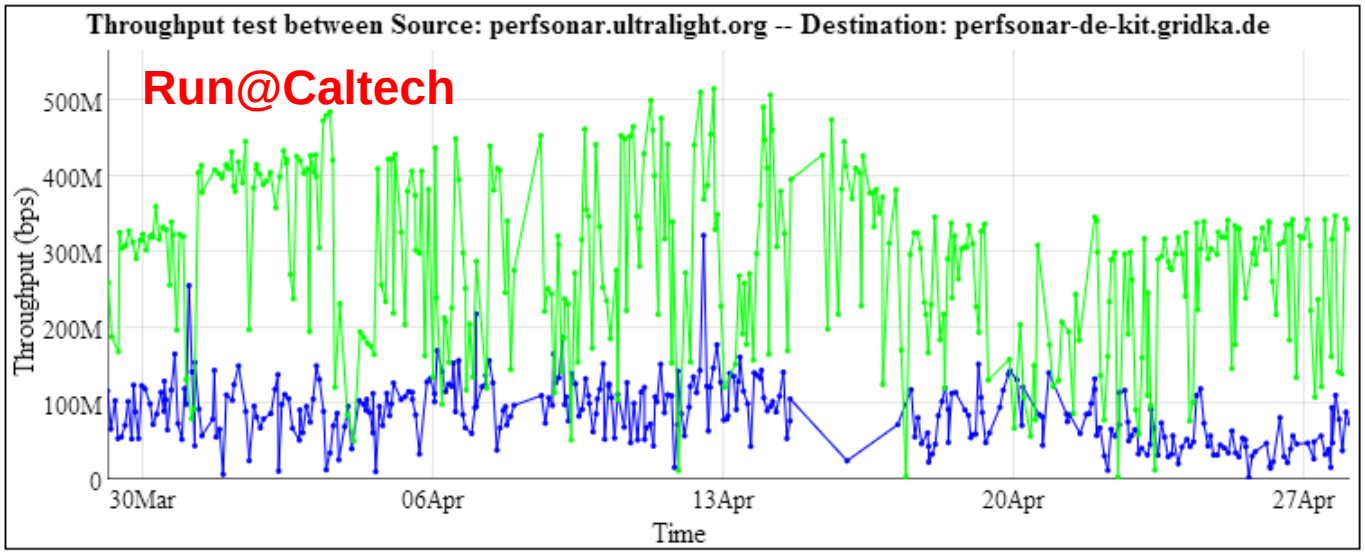
- To provide an idea of the network impact of a typical deployment here are some numbers as configured in the US
  - **Latency tests** send 10Hz of small packets (20 bytes) for each testing location. USATLAS Tier-2's test to ~10 locations. Since headers account for 54 bytes each packet is 74 bytes or the rate for testing to 10 sites is **7.4 kbytes/sec**. (Should increase?)
  - **Bandwidth tests** try to maximize the throughput. A 30 second test is run from each site in each direction once per 2 hour window. Each site runs tests in both directions. Typically the best result is around **925 Mbps on a 1Gbps link for a 30 second test**. That means we send 4x925 Mbps*30 sec every 2 hours per testing pair (src-dst) or about 7.5 Mbps average.
  - Tests are configurable but the above settings are working fine.

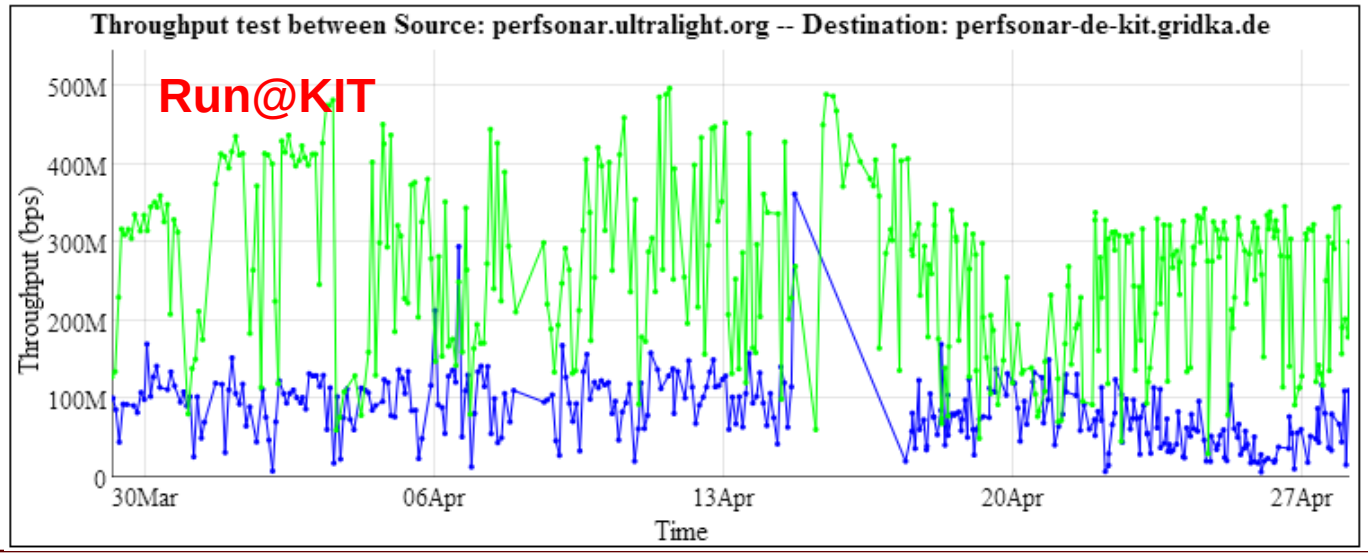
Examining Red Bandwidth

- Click top half of box to get measurement from Caltech
- Click bottom half of box to get measurement from KIT
- Hover-over to get average



BW Caltech to KIT



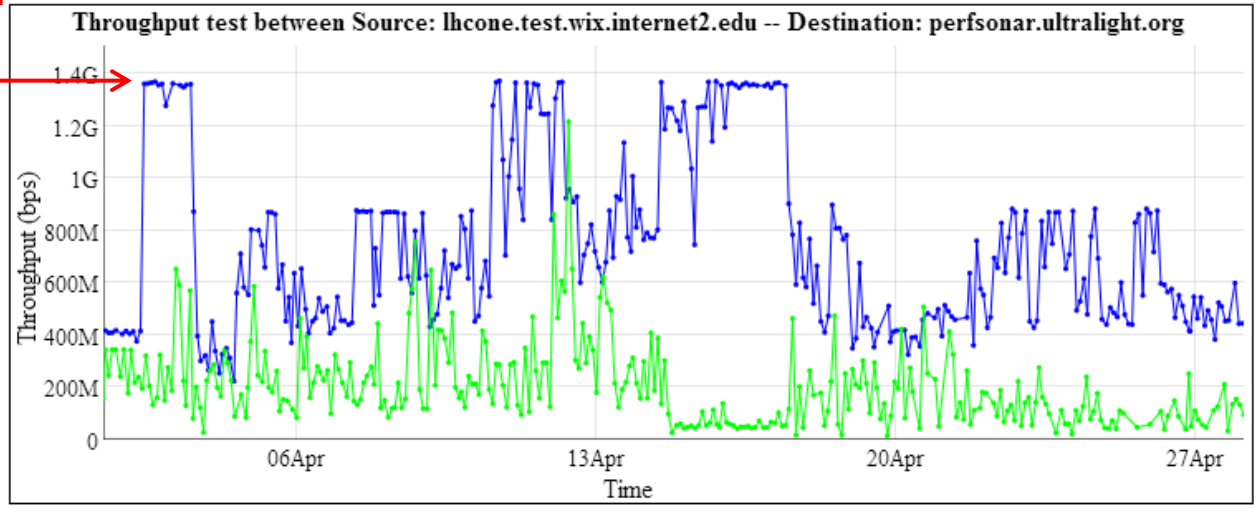


Latency Plot WIX-Caltech

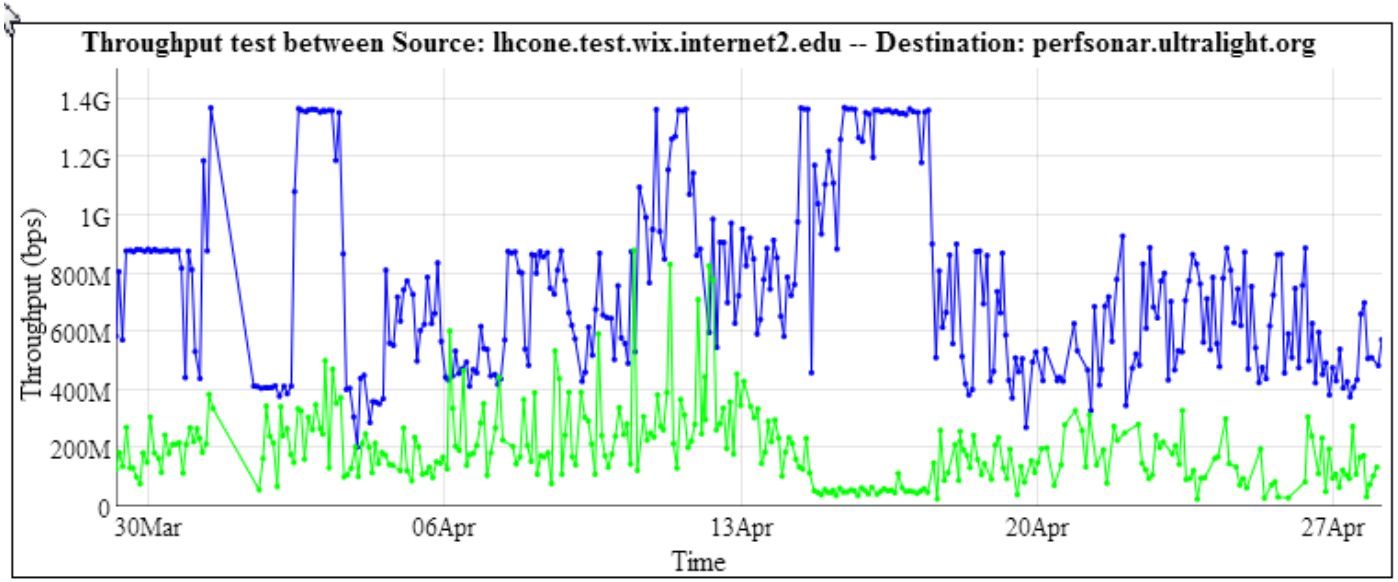
BW I2-WIX to Caltech

Sender limit? Path limit?

WIX-Cal
70 ms RTT
1.37 Gb max

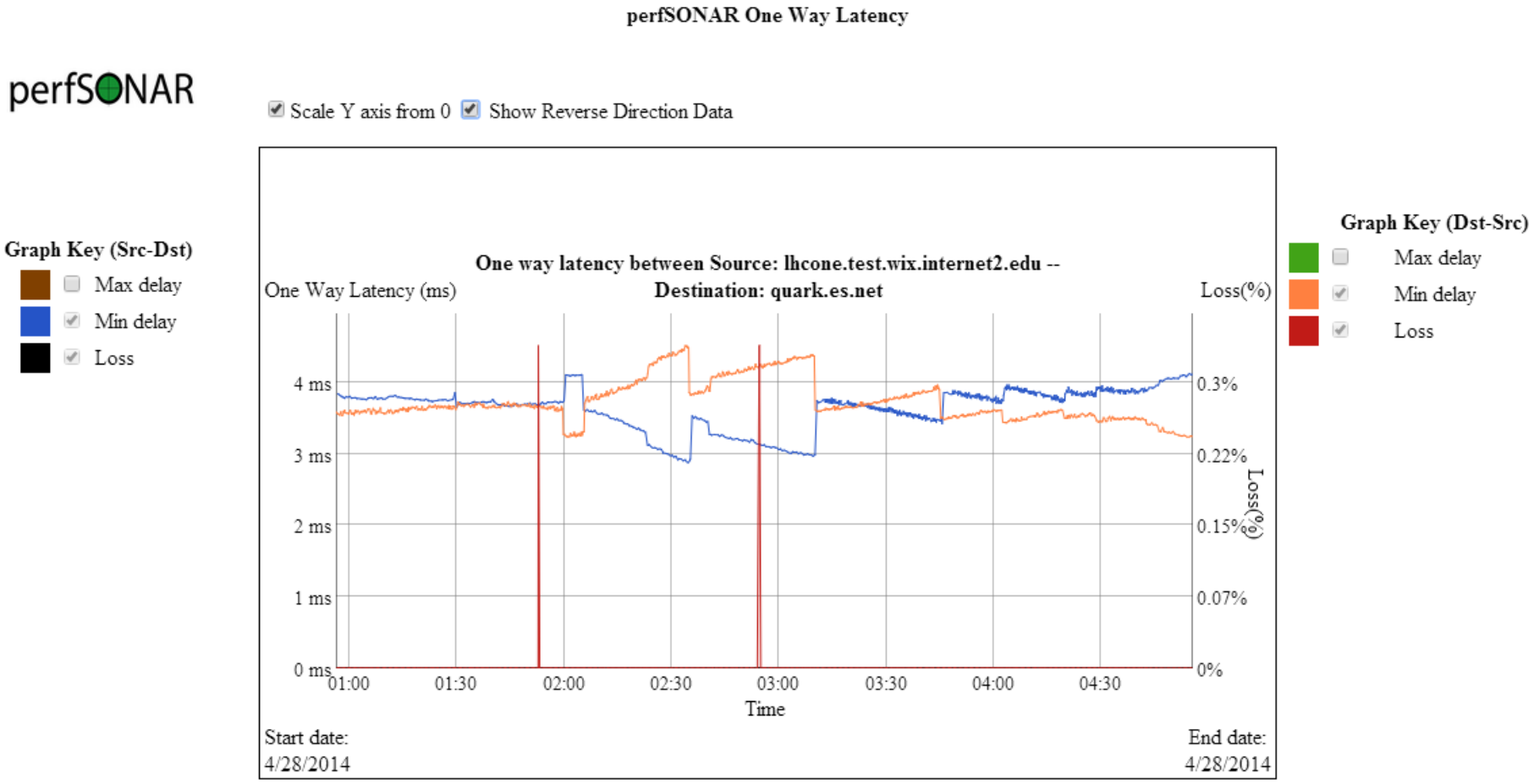


Run@WIX



Run@Caltech

Latency Plot WIX-ESnet BNL

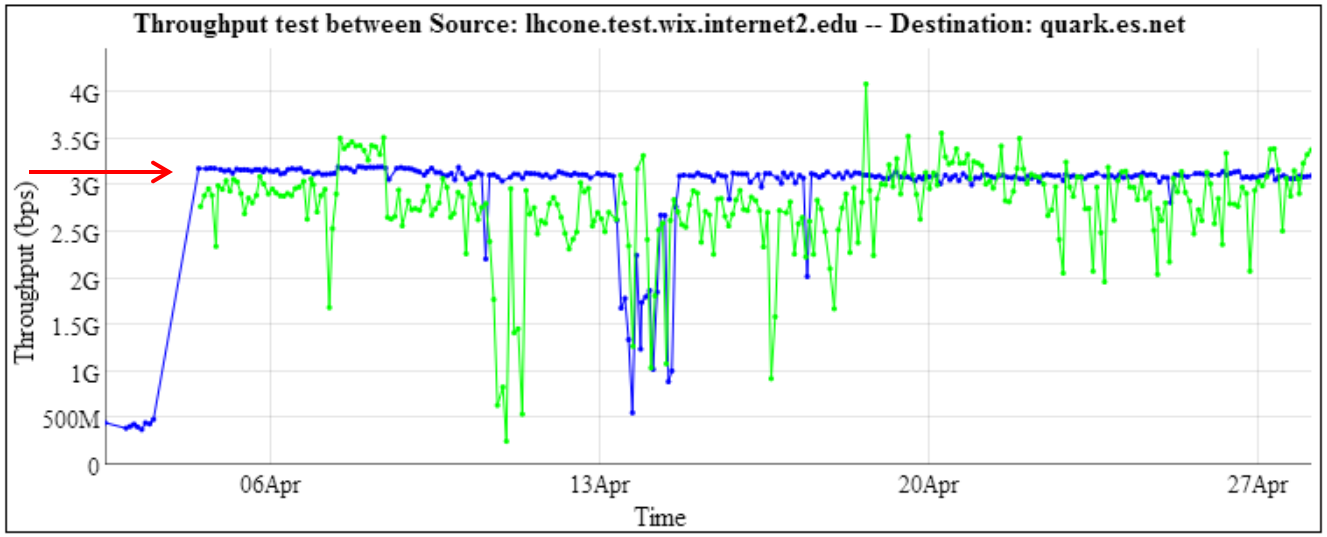


BW I2-WIX to Esnet-Quark

Sender limit? Path limit?

WIX-ESnet 7.4 ms RTT

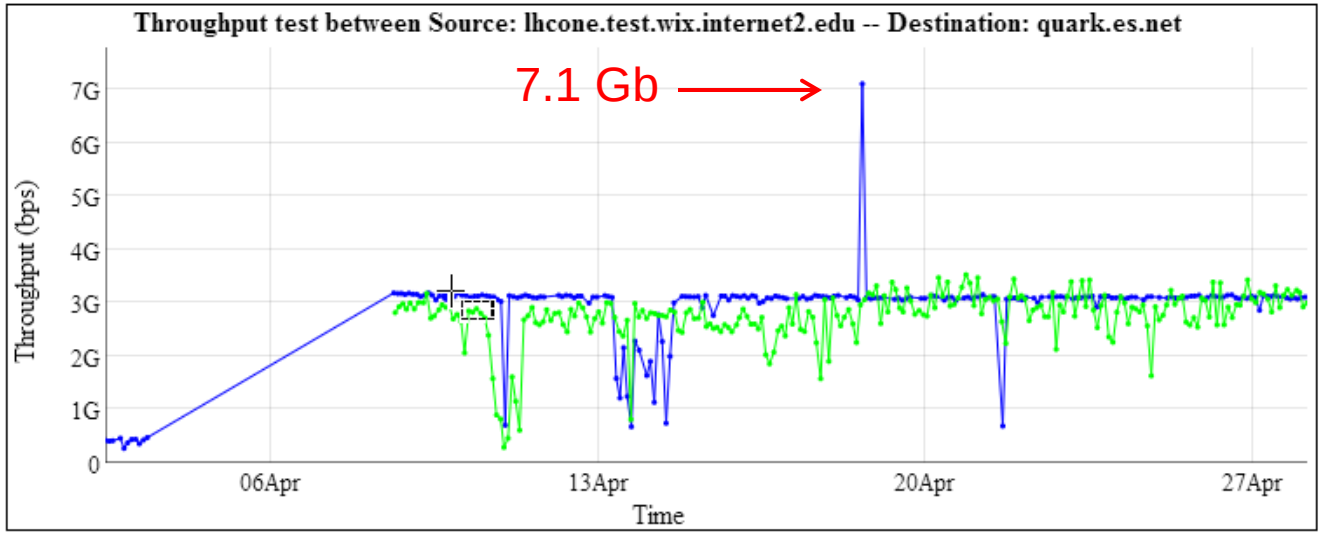
3.2 Gb max



Run@WIX

Limit doesn't scale like 1/RTT

TCP Stack?



Run@ESnet

Traceroute WIX-ESnet

Stanford University seal

SLAC logo, click here to learn more about SLAC

traceroute from 149.165.227.62 (lhcone.test.wix.internet2.edu) to 198.124.80.193 (quark.es.net) for 192.41.231.47

CGI script maintainer: *Les Cottrell*, SLAC. Script version 6.4, 8/29/2013, Jason Zurawski, Les Cottrell.
Download perl source code.
To perform a traceroute/ping/tracepath function from lhcone.test.wix.internet2.edu to the target, enter the desired target host.domain (e.g. www.yahoo.com) or Internet address (e.g. 137.138.28.228) in the box below. Note the fucntion is performed for the target's resolved Internet address.

Enter target name or address: then push 'Enter' key.

Lookup: domain name | Locating a Host | visual traceroute | Find AS's between hosts | Find AS of a host | contacting someone

Related web sites
Traceroute servers, Monitoring tutorial, Internet monitoring What is my IP address?

Please note that traceroutes can appear similar to port scans. If you see a suspected port scan alert, for example from your firewall, with a series of ports in the range 33434 - 33465, coming from lhcone.test.wix.internet2.edu it is probably a reverse traceroute from our web based reverse traceroute server. Please do NOT report this to us, it will almost certainly be a waste of both of our times. For more on this see Traceroute security issues.

```
Executing exec(traceroute -m 30 -q 3 198.124.80.193 140)
traceroute to 198.124.80.193 (198.124.80.193), 30 hops max, 140 byte packets
 1  et-9-0-0.4002.rtr.wash.net.internet2.edu (64.57.30.148)  0.513 ms  0.508 ms  0.504 ms
 2  xe-0-3-0.2001.rtr.newy32aoa.net.internet2.edu (64.57.30.225)  6.006 ms  6.006 ms  6.000 ms
 3  64.57.30.229 (64.57.30.229)  6.505 ms  6.508 ms  6.802 ms
 4  quark-gw.es.net (198.124.80.194)  7.748 ms  7.749 ms  7.743 ms
 5  quark.es.net (198.124.80.193)  7.624 ms  7.543 ms  7.542 ms
traceroute -m 30 -q 3 198.124.80.193 140 took 0secs. Total time=1secs.
```