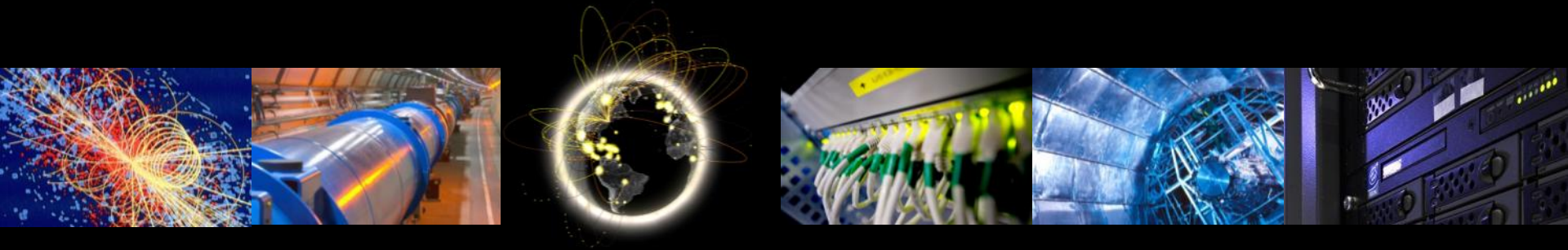


Update on OSG/WLCG Network Services

Shawn McKee, Marian Babik

2015 WLCG Collaboration Workshop
12th April 2015



Network Monitoring in WLCG/OSG

- **Goals:**
 - Find and isolate “network” problems; alerting in time
 - Characterize network use (base-lining)
 - Provide a source of network metrics for higher level services
- **Choice of a standard open source tool: perfSONAR**
 - Benefiting from the R&E community consensus
- **Tasks achieved:**
 - Monitoring in place to create a baseline of the current situation between sites
 - Continuous measurements to track the network, alerting on problems as they develop
 - Developed test coverage and made it possible to run “on-demand” tests to quickly isolate problems and identify problematic links



perfSONAR Deployment

259 perfSONAR instances
registered in GOCDB/OIM
233 Active perfSONAR instances
171 Running latest version (3.4.2)



- Initial deployment coordinated by WLCG perfSONAR TF
- Commissioning of the network followed by WLCG Network and Transfer Metrics WG

perfSONAR Current Release

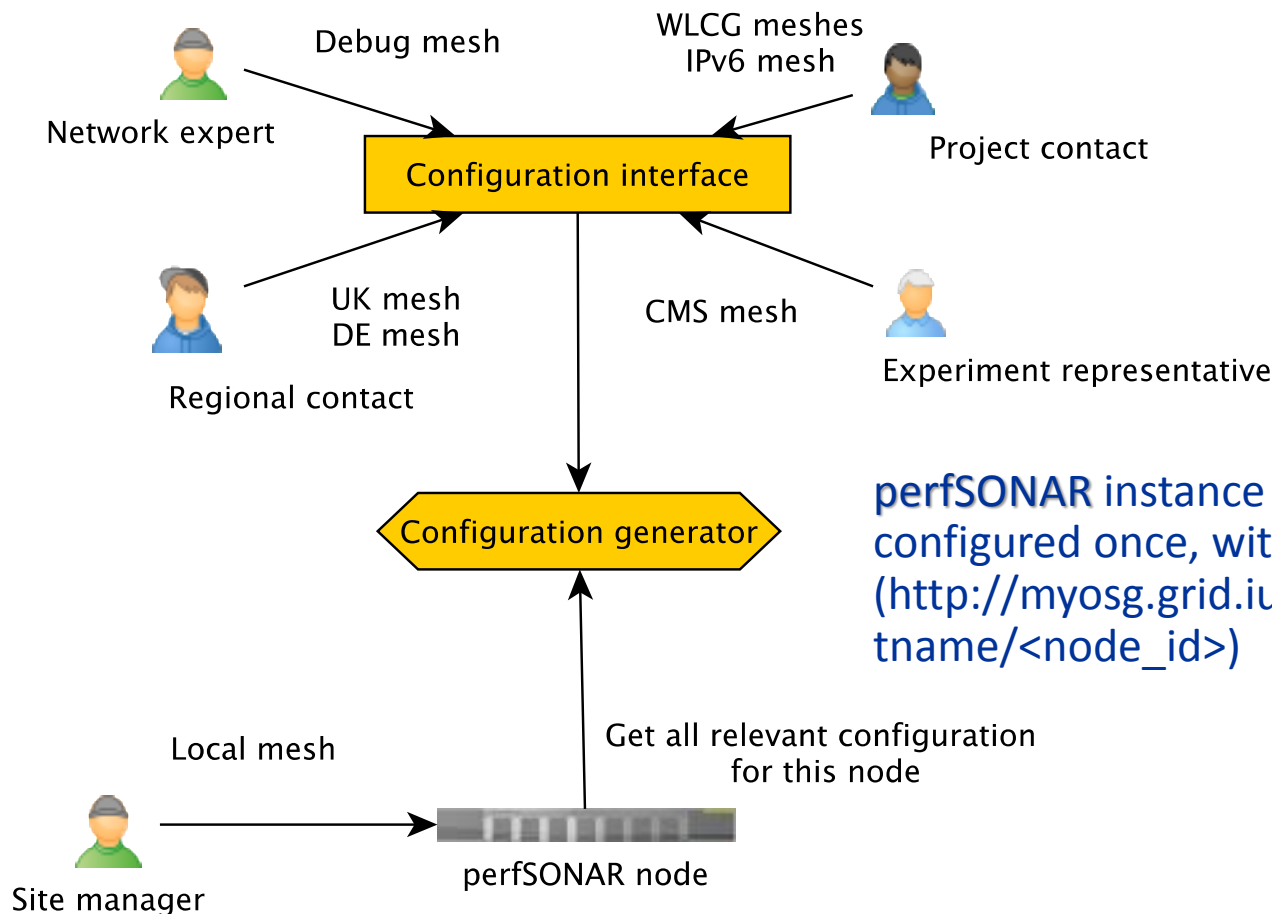
- The WLCG deployment scale and monitoring has helped identify issues in the perfSONAR toolkit
 - **Version 3.3.x** had numerous issues including security problems in the underlying OS and tests that wouldn't consistently run
 - **Version 3.4** fixed all known issues we identified in 3.3 but introduced a few new problems: excessive logging, more memory use, regular tests stopping and a few minor bugs
 - **Current version 3.4.2** has addressed all the new issues we know about. See release notes at <https://www.perfsonar.net/release-notes/version-3-4-2/>
- We track deployment status daily at http://grid-monitoring.cern.ch/perfsonar_report.txt

perfSONAR Metrics and Meshes

- Tests are organized in meshes – set of instances that test to each other
- perfSONAR regular tests currently configured:
 - **Traceroute**: End to end path, important to understand context of other metrics (Full WLCG mesh/hour)
 - **Throughput**: Notice problems and debug network, also help differentiate server problems from path problems (Full WLCG mesh/week)
 - **Latency**: Notice route changes, asymmetric routes and watch for excessive **Packet Loss** (Regional meshes, Continuous, 10Hz)
- perfSONAR is a testing framework, new tests and tools can be integrated as they become available
 - From iperf to iperf3, traceroute to tracepath
- Dynamic reconfigurations now possible
 - Creation, modification of meshes
 - Test frequency and parameters
- Additional perfSONAR nodes inside local network, and/or at periphery still needed (on LHCONe: MANLAN, WIX, GEANT)
 - Characterize local performance and internal packet loss
 - Separate WAN performance from internal performance

Configuration Interface


- perfSONAR instance can participate in more than one mesh
- Configuration interface and auto-URL enables dynamic re-configuration of the entire perfSONAR network



perfSONAR instance only needs to be configured once, with an auto-URL (http://myosg.grid.iu.edu/pfmesh/hostname/<node_id>)

Configuration Interface

← → ↻ <https://oim.grid.iu.edu/oim/meshconfig>

 OIM ▾ Home Certificate Topology Downtimes Virtual Organizations Support Centers Campus Grids Projects

Mesh Config Administrator

Host Groups Parameter Sets Configs Tests

Configuration to be part of ✖

* Required

Service Type * Required

Name * Required

Parameters * Required

Mesh Type * Required

Host Group A

Disable

OSG and the perfSONAR developers are working on creating a standalone “mesh-config” package to allow users and sites to manage and maintain their own meshes

MaDDash for Metrics Visualization

<http://psmad.grid.iu.edu/maddash-webui/>

WLCG/OSG perfSONAR Dashboard

☰ Dashboards Settings WLCG/OSG Networking Resources

USATLAS Mesh Config - USATLAS Bandwidth Mesh Test

USATLAS Mesh Config - U

■ Throughput >= 900Mbps ■ Throughput < 900Mbps

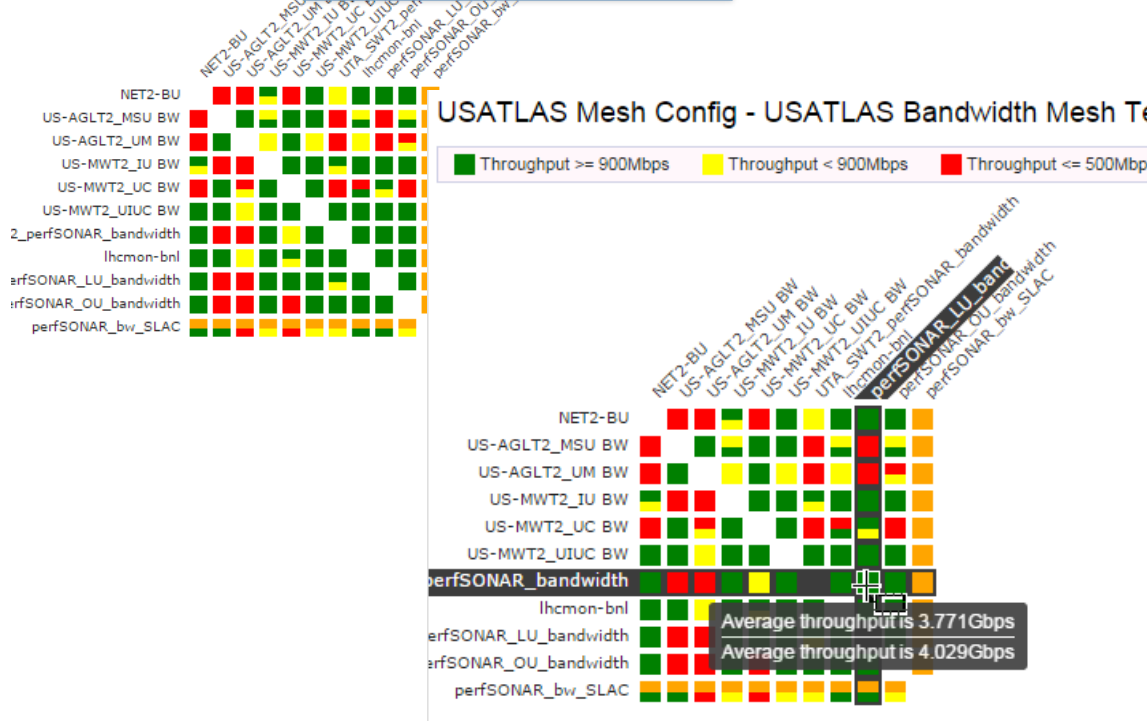
- Wiki on Deploying perfSONAR
- WLCG Prototype OMD/Check_MK
- WLCG Production OMD/Check_MK
- OSG Mesh-config
- WLCG Network and Transfer Metrics WG
- OSG Production MaDDash
- perfSONAR Homepage

WLCG/OSG perfSONAR Dashboard

☰ Dashboards Settings WLCG/OSG Network

Bandwidth tests between all WLCG hosts

- Belle II Mesh
- CA Mesh Config
- DE Mesh Config
- Debug Mesh (temp)
- Dual-Stack Mesh Config
- ES Mesh Config
- FR Mesh Config
- IT ATLAS Mesh Config
- IT CMS Mesh Config
- LHCONE Mesh Config
- Latin America Config
- My Sites
- ND Mesh Config
- OPN Config
- OSG Grid Operations Center Test Mesh Config
- RU Config
- TW Config
- UK Config
- USATLAS Mesh Config
- USCMS Mesh Config
- perfSONAR Testbed
- All Grids



Detailed Service Checks



Services of Host ps-latency.clumeq.mcgill.ca 12 rows /DC=com/DC=DigiCert-Grid/O=Open Science Grid/OU=People/CN=Shawn Mc...

Availability ...

State	Service	Icon
OK	perfSONAR 3.4+ Toolkit Version	
OK	perfSONAR Administrator Details	
OK	perfSONAR esmond Freshness Latency Direct	
OK	perfSONAR esmond Freshness Latency Reverse	
OK	perfSONAR esmond Measurement Archive	
psum02.agt2.org		
State	Service	Icon
OK	perfSONAR 3.4+ Toolkit Version	
OK	perfSONAR Administrator Details	
OK	perfSONAR BWCTL Bandwidth Test Controller	
WARN	perfSONAR esmond Freshness Bandwidth Direct	
WARN	perfSONAR esmond Freshness Bandwidth Reverse	
OK	perfSONAR esmond Measurement Archive	
OK	perfSONAR Homepage	
OK	perfSONAR Latitude/Longitude Configured	
OK	perfSONAR Mesh Configuration	
OK	perfSONAR NTP Service	
OK	perfSONAR Regular Testing Service	
OK	perfSONAR Toolkit Version	

Service perfSONAR esmond Freshness Bandwidth Direct, psu... 1 row /DC=com/DC=DigiCert-Grid/O=...

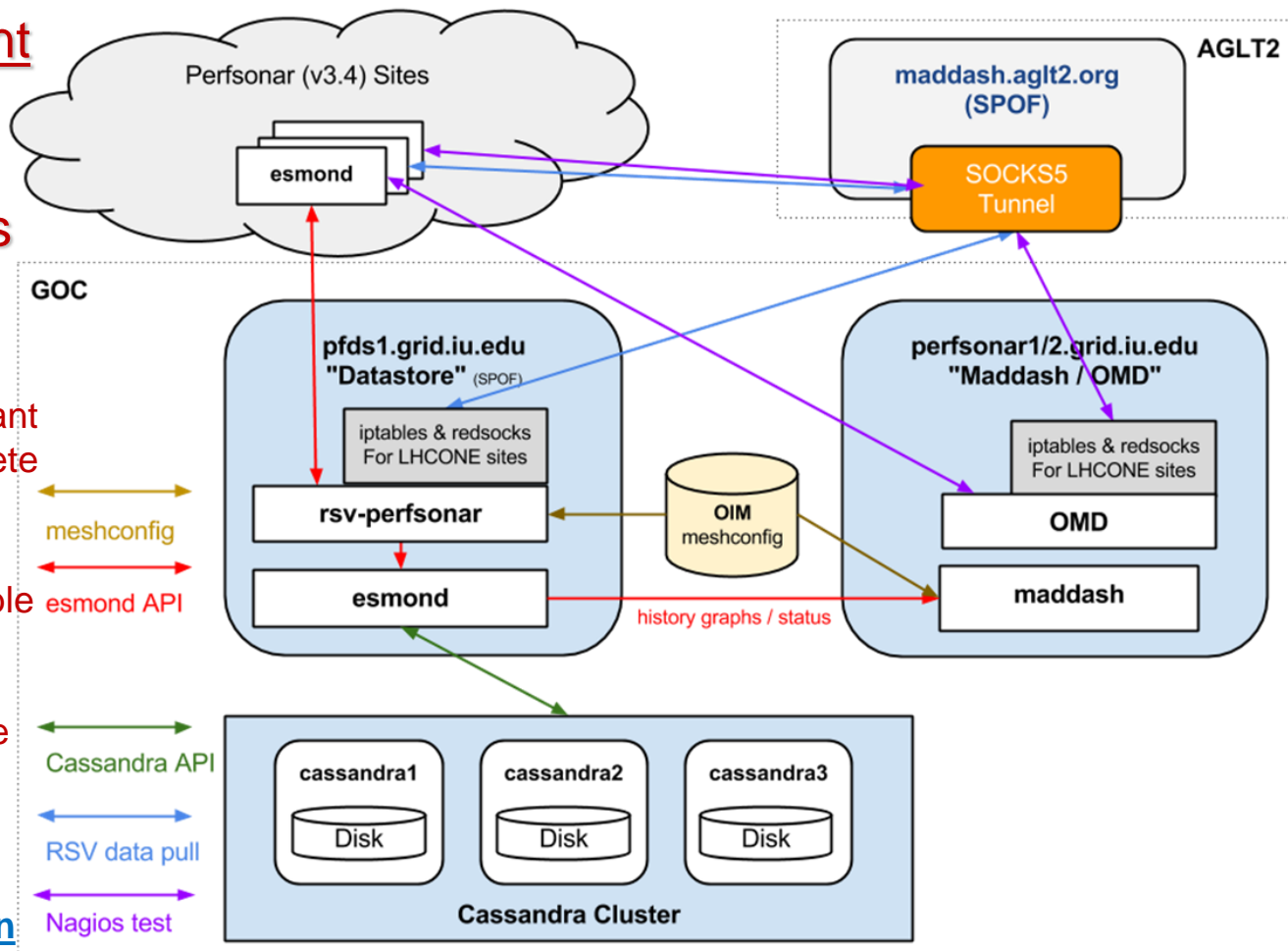
Availability ...

Site alias			
Hostname	psum02.agt2.org		
Service description	perfSONAR esmond Freshness Bandwidth Direct		
Service icons			
Service state	WARN		
Servicegroups the service is member of	sg_esmond_bw_direct, sg_esmond		
Service service level			
Service contact groups	all		
Service contacts			
Output of check plugin	WARNING Found stale hosts for certain events, time-range: 3700		
Long output of check plugin (multiline)	Time-range: 3700 Even-types checked: packet-trace Mesh (Event-type): GOC Traceroute Tests (packet-trace) Destinations count: 1 Missing destinations: perfsonar-cms2.itns.purdue.edu Mesh (Event-type): Traceroute Test Between WLCG Bandwidth Hosts (packet-trace) Destinations count: 10 Missing destinations: marperf01.in2p3.fr, perfsonar.pleiades.uni-wuppertal.de, ps02.cat.cbpf.br, perfsonar2.hep.kbfi.fr, ps02.cism.ucl.ac.be, clrperf-bwctl.in2p3.fr, perfsonar-cms2.itns.purdue.edu, perfsoar2.hep.kbfi.fr Documentation for this check can be found at https://twiki.opensciencegrid.org/bin/view/Documentation		
OK NTP synchronized	2015-01-29 20:17:01	52 min	
OK Regular Testing enabled and running	2015-02-19 10:51:38	51 min	
OK - Version 3.4.2 OK (cached:1)	2014-12-11 19:56:41	2 hrs	

OSG Network Datastore

A critical component is the datastore to organize and store the network metrics and associated metadata

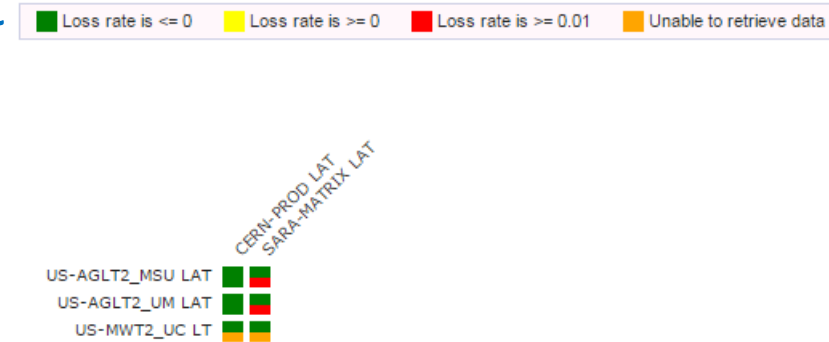
- ❑ OSG is gathering relevant metrics from the complete set of OSG and WLCG perfSONAR instances
- ❑ This data will be available via an API, must be visualized and must be organized to provide the **“OSG Networking Service”**
- ❑ **Operating now**
- ❑ **Targeting a production service by mid-summer**



Network Problem Debugging(1/4)

- Problem with transfers from SARA to AGLT2 noted February 20th.
 - FTS transfers failing. ATLAS asked about network. Dataset had large files; transfers failed because of timeout (**1 MB/sec=3.6GB/hour; 1 hour timeout**)
 - Setup 'Debug' mesh using OSG tools to track SARA, CERN to AGLT2, MWT2
 - Jason Zurawski/ESnet ran test using perfSONAR. Bad throughput ~few hundred kbits/sec SARA->AGLT2 **network issue!**

Debug Mesh (temp) - Debug LT



Network Debugging (2/4)

- Ticket opened by AGLT2 with Internet2 NOC on February 20th
 - Poor iperf results also between SARA and MWT2
 - Routes provided both ways by perfSONAR

Topology beginning at Tue Mar 10 00:01:17 2015 (UTC -4)

Hop	Router	IP	Delay	MTU
1	ge-0-2-0-1020.grid-r1.grid.sara.nl	145.100.17.1	0.16ms	
2	geant-lhc1-gw.mx1.ams.nl.geant.net	62.40.126.161	0.279ms	
3	et-10-0-0.3019.rtr.newy32aoa.net.internet2.edu	64.57.30.225	74.946ms	
4	192.17.10.74	192.17.10.74	105.154ms	
5	64.57.30.154	64.57.30.154	120.308ms	
6	psum02.aglt2.org	192.41.230.60	107.806ms	

Topology beginning at Tue Mar 10 00:40:32 2015 (UTC -4)

Hop	Router	IP	Delay	MTU
1	192.41.230.1	192.41.230.1	0.266ms	
2	et-8-0-0.3151.rtr.chic.net.internet2.edu	64.57.30.155	5.796ms	
3	et-10-0-0.402.rtr.newy32aoa.net.internet2.edu	64.57.30.142	33.651ms	
4	64.57.30.228	64.57.30.228	119.977ms	
5	surfnet-lhc1-gw.ams.nl.geant.net	62.40.126.160	119.973ms	
6	ps.lhcopn-ps.sara.nl	145.100.17.9	119.803ms	

Topology beginning at Tue Mar 10 13:46:17 2015 (UTC -4)

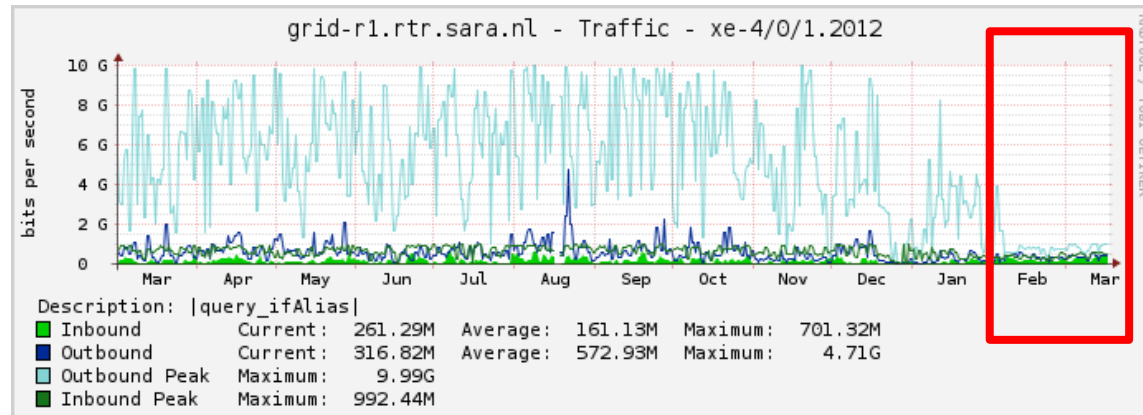
Hop	Router	IP	Delay	MTU
1	ge-0-2-0-1020.grid-r1.grid.sara.nl	145.100.17.1	0.158ms	
2	geant-lhc1-gw.mx1.ams.nl.geant.net	62.40.126.161	0.268ms	
3	et-10-0-0.3019.rtr.newy32aoa.net.internet2.edu	64.57.30.225	87.094ms	
4	192.17.10.74	192.17.10.74	102.803ms	
5	requestTimedOut	requestTimedOut		
6	psum02.aglt2.org	192.41.230.60	108.808ms	

Topology beginning at Tue Mar 10 13:43:23 2015 (UTC -4)

Hop	Router	IP	Delay	MTU
1	192.41.230.1	192.41.230.1	0.235ms	
2	esnet-lhc1-a-aglt2.es.net	198.124.80.53	6.266ms	
3	62.40.126.149	62.40.126.149	109.678ms	
4	62.40.126.148	62.40.126.148	108.904ms	
5	surfnet-lhc1-gw.ams.nl.geant.net	62.40.126.160	119.086ms	
6	ps.lhcopn-ps.sara.nl	145.100.17.9	108.313ms	

Network Debugging (3/4)

- Internet2 initially pursuing asymmetric routes and link congestion in US
- I2 opened ticket with GEANT Feb 27
- GEANT brought up LHCONe perfSONAR in Frankfurt
 - Tests to SARA(close) and AGLT2(far) showed 3x better bandwidth to AGLT2. Problem close to SARA
 - Isolated link between SARA and GEANT March 4

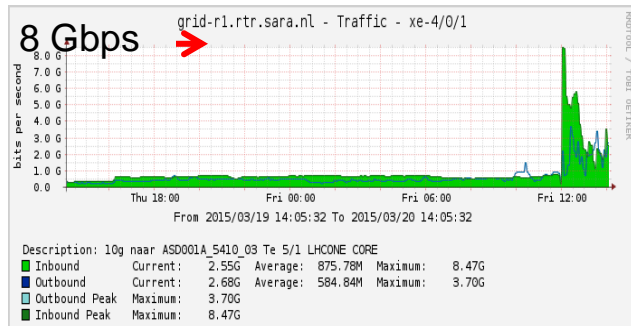


SARA noticed problems once we highlighted the link

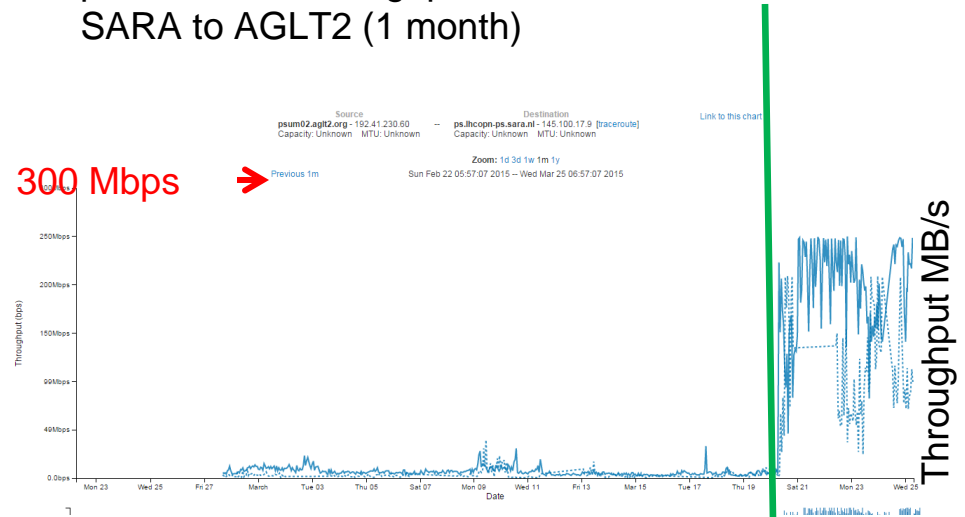


Problem Found/Fixed – Mar 20

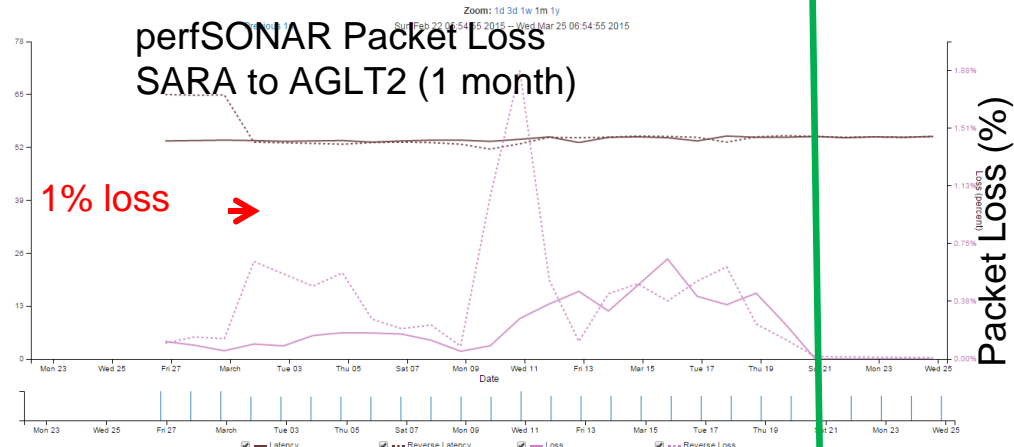
- 2 weeks of link-debugging
- Problem identified and fixed Mar 20
 - Bandwidth “policing” for LHCONE. More than a year ago, LHCONE setup 1 Gbps. *Never enforced.*
 - Turned on policing start of February.
 - **Changed BW 10Gbps March 20...fixed**



perfSONAR Throughput
SARA to AGLT2 (1 month)



perfSONAR Packet Loss
SARA to AGLT2 (1 month)



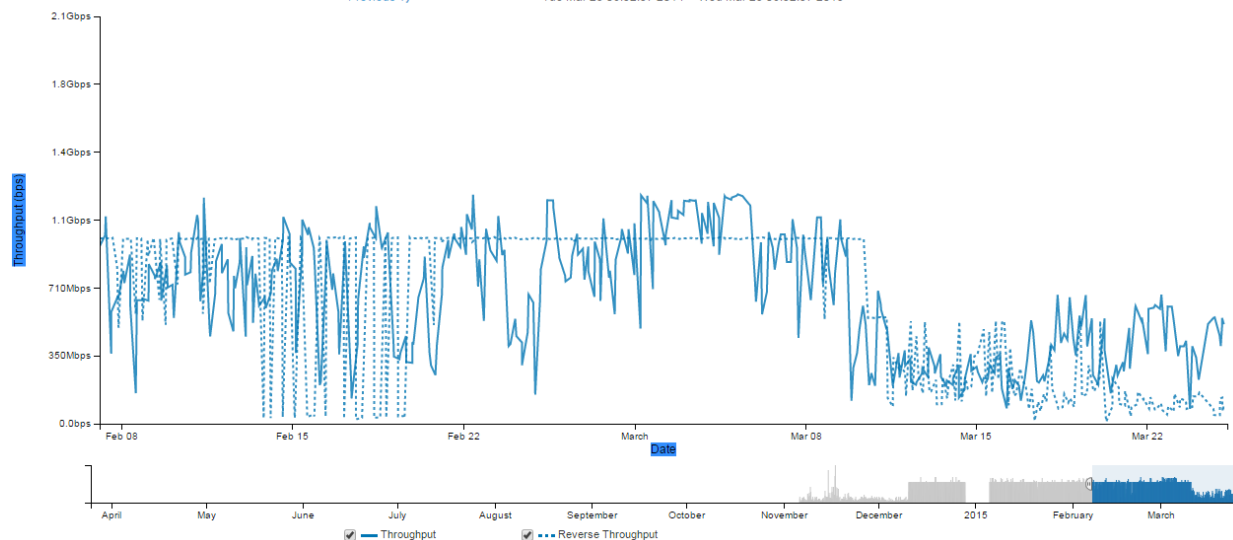
OSG Network Alerting

- What kinds of capabilities can we enable given a rich datastore of historical and current network metrics?
 - Users want "someone" to tell them when there is a network problem involving their site or their workflow.
 - Can we create a framework to identify when network problems occur and locate them? (**Must** minimize the false-positives).

Targeting this in the PuNDIT OSG satellite project

See PuNDIT Poster
<https://indico.cern.ch/event/304944/contribution/408> [#8 Poster in booth 24 Session B]

Source: psum02.aglt2.org - 192.41.230.60
Capacity: Unknown MTU: Unknown
Destination: ps2.ochep.ou.edu - 129.15.40.232 [traceroute]
Capacity: Unknown MTU: Unknown
Link to this chart
Zoom: 1d 3d 1w 1m 1y
Previous 1y
Tue Mar 25 06:32:37 2014 -- Wed Mar 25 06:32:37 2015



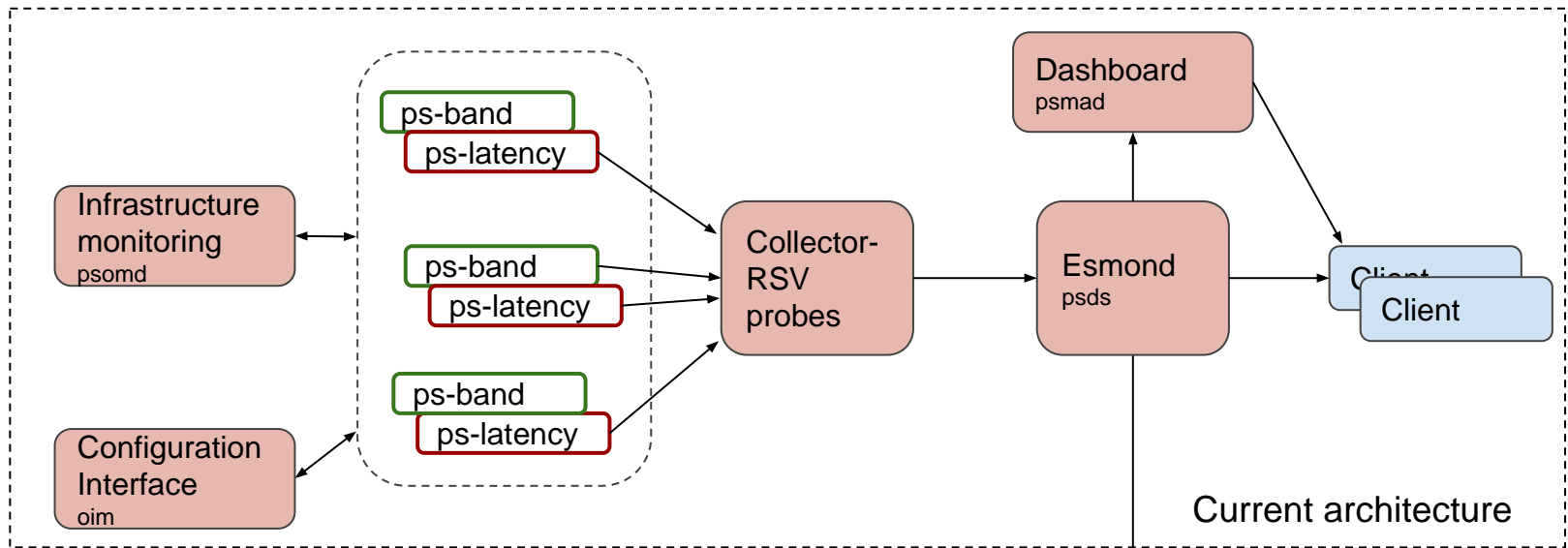
Understanding Network Topology

- Can we create tools to **manipulate, visualize, compare and analyze** network topologies from the OSG network datastore contents?
- Can we build upon these tools to create a set of next-generation network diagnostic tools to **make debugging network problems easier, quicker and more accurate?**
- Even without requiring the ability to perform complicated data analysis and correlation, basic tools developed in the area of network topology-based metric visualization would be very helpful in letting users and network engineers better understand what is happening in our networks.

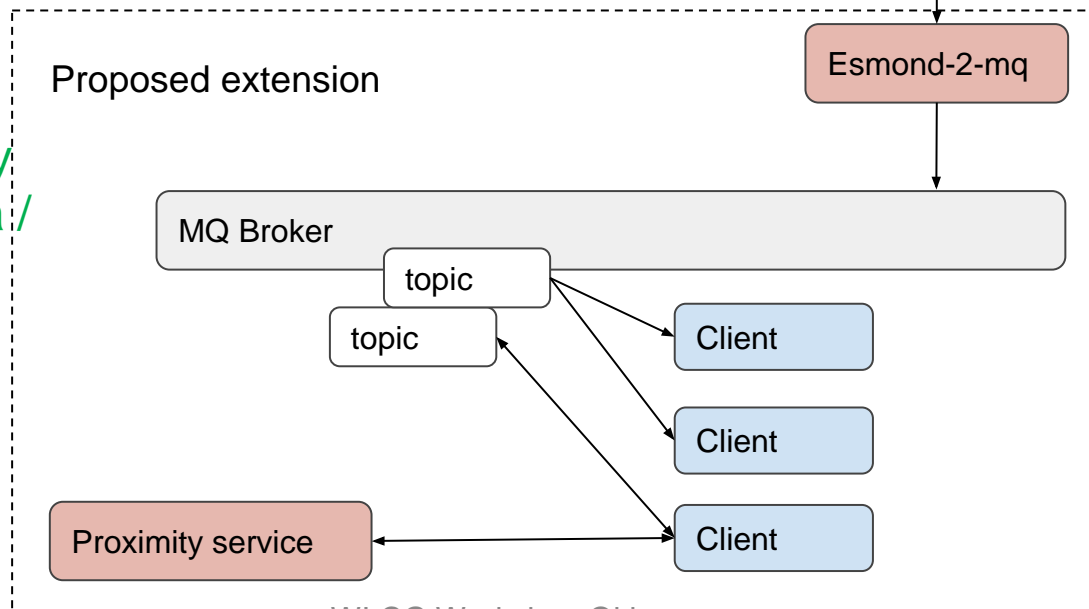
WLCG Networking and End-to-end

- Most scientists just care about the **end-to-end** results:
 - How well does their infrastructure support them in doing their science?
- Network metrics allow us to differentiate end-site issues from network issues.
- There is an opportunity to do this better by having access to **end-to-end** metrics to compare & contrast with network-specific metrics.
 - What end-to-end data can WLCG/OSG regularly collect for such a purpose? **WG is soliciting input here...**
 - Is there some kind of common instrumentation that can be added to some data-transfer tools? (NetLogger in GridFTP, having transfers "report" results to the nearest perfSONAR-PS instance?, etc)
 - FTS data-analysis project is underway (adds pS metrics)

Proposed Datastore Extension



Current architecture



Proposed extension

Being implemented by Henryk Giemza / LHCb

Goal: easy access to specific net metrics

Closing remarks

- perfSONAR widely deployed and already showing benefits in troubleshooting network issues
 - Additional deployments by R&E networks still needed
- Significant progress in configuration and infrastructure monitoring
 - Helping to reach full potential of the perfSONAR deployment
- OSG datastore – community network data store for all perfSONAR metrics – planned to enter production in Q3
- Integration projects underway to aggregate network and transfer metrics
 - FTS Performance
 - Experiment's interface to perfSONAR
- Advanced network monitoring - diagnosis and alerts based on perfSONAR, developed within NSF funded PuNDIT project

Questions?



References

- Network Documentation
<https://www.opensciencegrid.org/bin/view/Documentation/NetworkingInOSG>
- Deployment documentation for OSG and WLCG hosted in OSG
<https://twiki.opensciencegrid.org/bin/view/Documentation/DeployperfSONAR>
- New 3.4 MA guide <https://code.google.com/p/perfsonar-ps/wiki/MeasurementArchiveClientGuide>
- Modular Dashboard and OMD Prototypes
 - <http://maddash.aglt2.org/maddash-webui>
 - https://maddash.aglt2.org/WLCGperfSONAR/check_mk
- OSG Production instances for OMD, MaDDash and Datastore
 - <http://psmad.grid.iu.edu/maddash-webui/>
 - https://psomd.grid.iu.edu/WLCGperfSONAR/check_mk/
 - <http://psds.grid.iu.edu/esmond/perfsonar/archive/?format=json>
- Mesh-config in OSG <https://oim.grid.iu.edu/oim/meshconfig>
- Use-cases document for experiments and middleware
<https://docs.google.com/document/d/1ceiNITUJCwSuOuvbEHZnZp0XkWkwdkPQTQic0VbH1mc/edit>