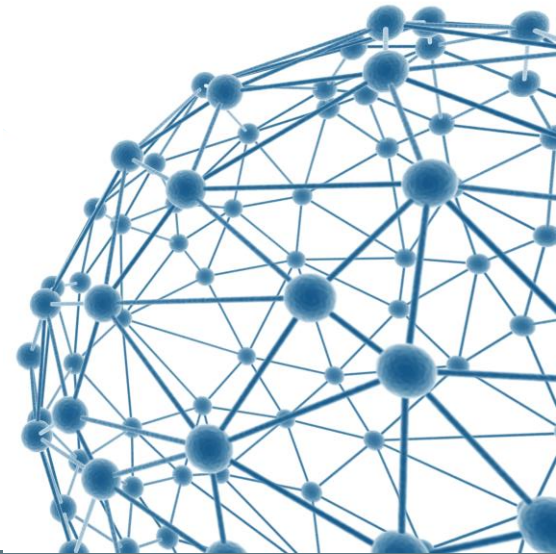




# Deployment of a WLCG network monitoring infrastructure based on the perfSONAR-PS technology

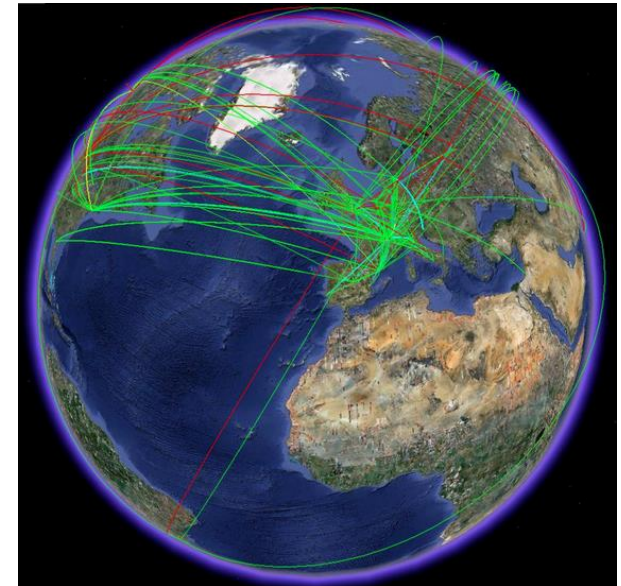
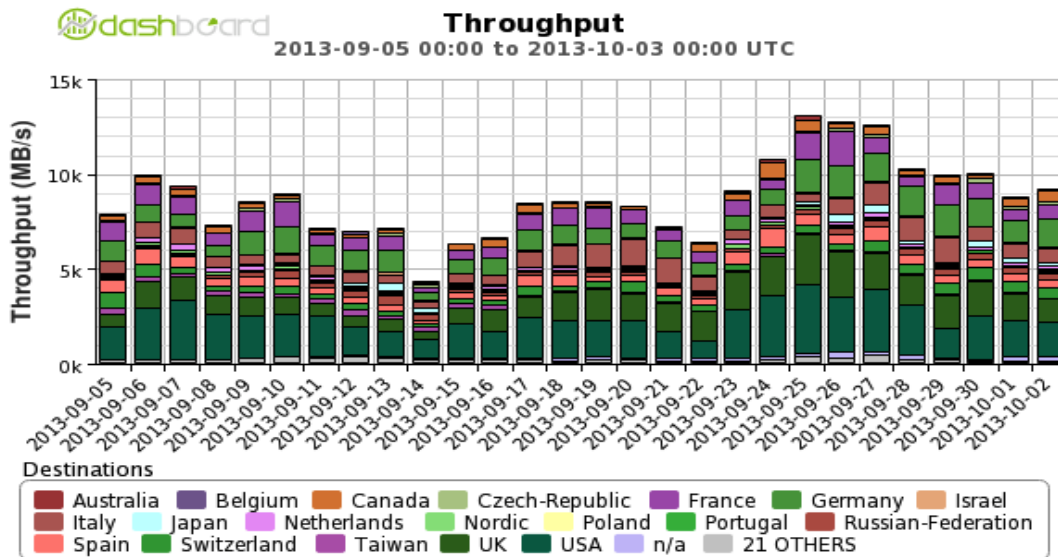
Campana (CERN-IT/SDC), McKee (Michigan)

16 October 2013



# Network Monitoring for WLCG

- WLCG relies heavily on the underlying networks
  - Interconnect sites and resources



# Network Monitoring for WLCG

- We discovered that end-to-end network issues can be difficult to spot and debug
  - Insufficient tools to detect network failures and diagnose
    - Sometimes noticed only by the application
  - Multiple “owners” (administrative domains)
- The famous “BNL-CNAF network issue”
  - [https://ggus.eu/ws/ticket\\_info.php?ticket=61440](https://ggus.eu/ws/ticket_info.php?ticket=61440)
  - 7 months, 72 entries in the ticket, lots of real work from many people

# Network Monitoring Infrastructure

- The WLCG service needs to guarantee effective network usage and rapid solution of network issues
- Using “standard” tools comes many benefits
  - Quality software, supported by a large community
  - Standard metrics, familiar for network engineers
- WLCG choose perfSONAR as the basis of its network monitoring infrastructure
  - Significant experience already in USATLAS and LHCOPN

# perfSONAR and perfSONAR-PS

- **perfSONAR** is an infrastructure for network performance monitoring
  - Organized as consortium of organizations
    - building an interoperable network monitoring middle-ware
  - Defines the service types and a protocol for them to communicate
  - Develops the software packages to implement the services
- **perfSONAR-PS** is an open source development effort based on perfSONAR
  - targeted at creating an easy-to-deploy and easy-to-use set of perfSONAR services
  - Comes with all-in-one solution (CD or USB) or single packages for CentOS 5 and 6

# perfSONAR-PS toolkit

- **Web based GUI**
  - for the administrator to configure the service and schedule the tests
  - for the user to display the measurements
  
- **Engine for execution of various test types**
  - Throughput tests (bwctl), non-concurrent
  - Ping (PingER), time stamped
  - One-Way Latency tests (owamp), time stamped
  - Traceroute
  - Network Diagnostic Tools (NDT, NPAD) on demand
  
- **A Measurement Archive stores and exposes programmatically the results**

User Tools	
Local Performance Services	
Global Performance Services	
Java OWAMP Client	<a href="#">↗</a>
Reverse Traceroute	<a href="#">↗</a>
Reverse Ping	<a href="#">↗</a>
Reverse Tracepath	<a href="#">↗</a>
Service Graphs	
Throughput	
One-Way Latency	
Traceroute	
Ping Latency	
SNMP Utilization	
Cacti Graphs	<a href="#">↗</a>
Toolkit Administration	
Administrative Information	
External BWCTL Limits	
External OWAMP Limits	
Enabled Services	
NTP	
Scheduled Tests	
Cacti SNMP Monitoring	<a href="#">↗</a>
perfSONAR Logs	<a href="#">↗</a>
Performance Toolkit	
Configuration Help	<a href="#">↗</a>
Frequently Asked Questions	<a href="#">↗</a>
About	<a href="#">↗</a>
Credits	<a href="#">↗</a>

# Early perfSONAR deployment

- perfSONAR deployment started in the OPN and USATLAS
- Test definitions statically configured on each node by the site administrator following a set of instructions
- Good for the OPN use case
  - well established list of sites
- Problematic for a broad deployment
  - Service endpoints might be changing
  - New sites might join
  - Difficult to coordinate the effort

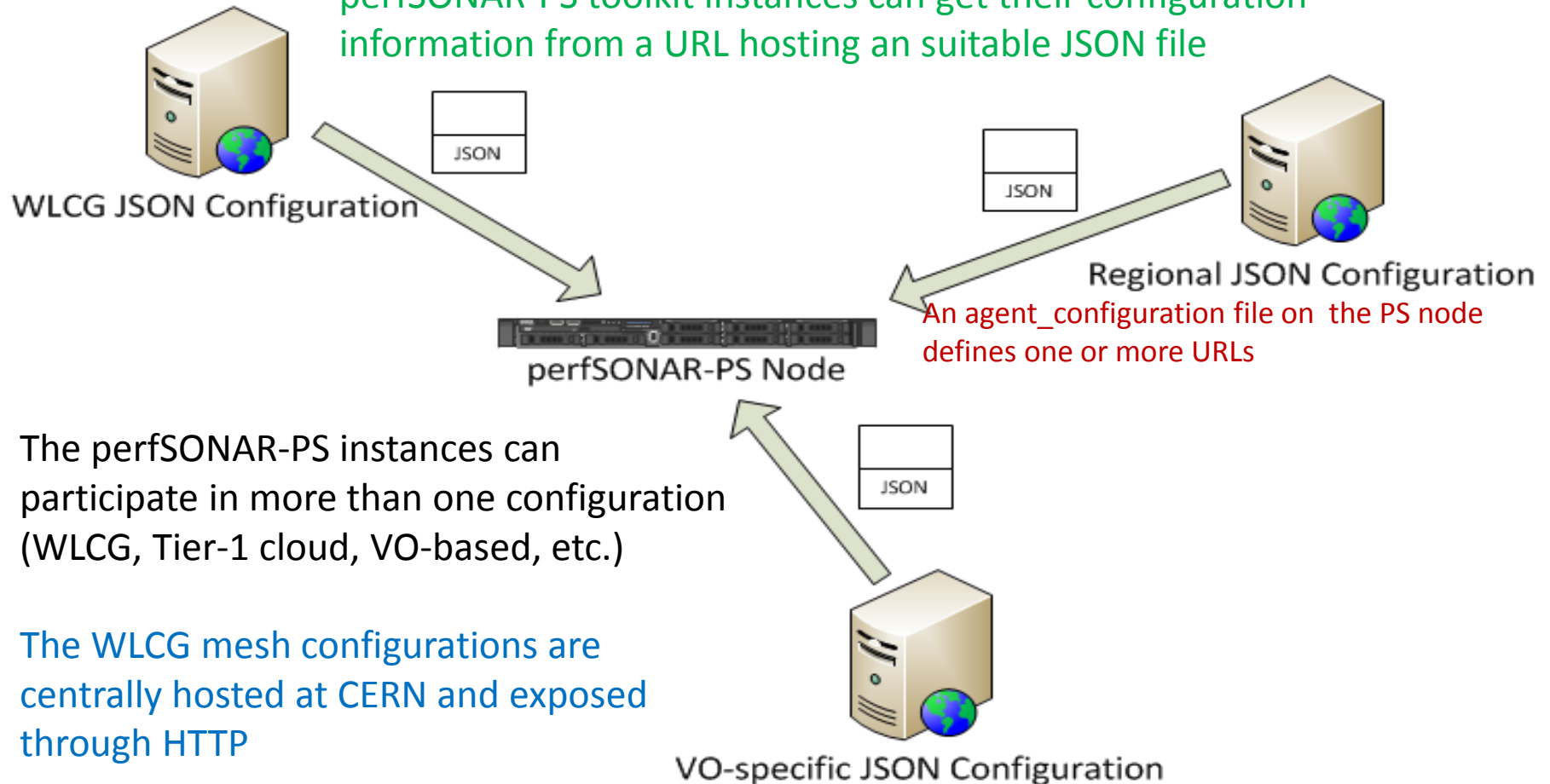
# WLCG deployment plan

- WLCG choose to deploy perfSONAR-PS at all sites worldwide
  - A dedicated WLCG Operations Task-Force was started in Fall 2012
- Sites are organized in regions
  - Based on geographical locations and experiments computing models
  - All sites are expected to deploy a bandwidth host and a latency host
- Regular testing is setup using a centralized (“mesh”) configuration
  - Bandwidth tests: 30 seconds tests
    - every 6 hours intra-region, 12 hours for T2-T1 inter-region, 1 week elsewhere
  - Latency tests; 10 Hz of packets to each WLCG site
  - Traceroute tests between all WLCG sites each hour
  - Ping(ER) tests between all site every 20 minutes



# perfSONAR-PS Mesh Example

perfSONAR-PS toolkit instances can get their configuration information from a URL hosting an suitable JSON file



The perfSONAR-PS instances can participate in more than one configuration (WLCG, Tier-1 cloud, VO-based, etc.)

The WLCG mesh configurations are centrally hosted at CERN and exposed through HTTP

<https://grid-deployment.web.cern.ch/grid-deployment/wlcg-ops/perfsonar/conf/>

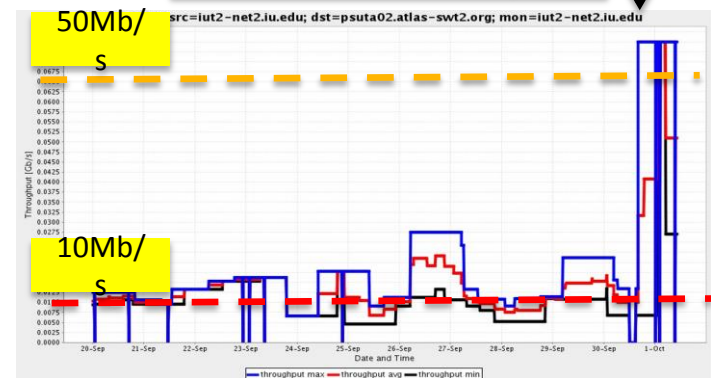
# The perfSONAR Modular Dashboard

- Centrally aggregates measurements from all PS hosts
  - Provides a web UI and REST interface
  - <http://perfsonar.racf.bnl.gov:8080/exda/>
- A new implementation maturing production quality
  - Addressing scalability issues for large meshes
  - Providing a more extensive REST API
  - Self-configuring from mesh definitions
  - Fancier ...
  - <http://perfsonar.racf.bnl.gov:8080/PsDisplay-1.0-SNAPSHOT/matrices.jsp?id=62>
- Discussions with OSG about hosting the Modular Dashboard service and automating mesh-config creation

US cloud throughput measurement

	---	0	1	2	3	4	5	6	7	8	9
0:BNLBNL-Test (lhcmn.bnl.gov)	---	0.29 0.31	0.45 0.10	0.54 0.21	0.35 0.13	0.24 0.02	0.24 0.40	0.24 0.24	0.21 0.01	0.04 0.01	0.17 0.23
1:AGLT2 (psmsu02.aglt2.org)	0.01 0.28	---	0.04 0.19	0.31 0.38	0.13 0.25	0.05 0.35	0.22 0.26	0.22 0.21	0.22 0.01	0.02 0.01	0.12 0.12
2:AGLT2 (psum02.aglt2.org)	0.07 0.32	0.24 0.93	---	0.31 0.92	0.20 0.41	0.08 0.36	0.21 0.15	0.21 0.15	0.02 0.02	0.00 0.02	0.00 0.11
3:MWT2 (lut2-net2.iu.edu)	0.03 0.25	0.03 0.35	0.04 0.09	---	0.50 0.43	0.00 0.32	0.02 0.02	0.02 0.02	0.03 0.03	0.12 0.13	0.12 0.13
4:MWT2 (mwt2-ps02.campuscluster.illinois.edu)	0.09 0.71	0.77 0.33	0.50 0.43	0.27 0.42	---	0.32 0.51	0.02 0.02	0.03 0.03	0.01 0.01	0.12 0.00	0.12 0.00
5:MWT2MWT2(UC) (uct2-net2.uchicago.edu)	0.88 0.66	0.05 1.00	0.84 0.36	0.38 0.34	0.19 0.15	---	0.02 0.76	0.84 0.03	0.22 0.11	0.14 0.11	0.18 0.18
6:NET2 (atlas-npt2.bu.edu)	0.03 0.65	0.85 0.13	0.23 0.19	0.72 0.00	0.82 0.11	0.37 0.44	---	0.57 0.08	0.11 0.11	0.14 0.00	0.14 0.00
7:SWT2 (ps2.oceph.ou.edu)	0.42 0.44	0.37 0.98	0.76 0.29	0.28 0.00	0.42 0.14	0.43 0.73	0.71 0.92	---	0.03 0.04	0.27 0.27	0.27 0.27
8:SWT2 (psuta02.atlas-swt2.org)	0.44 0.49	0.01 0.30	0.00 0.00	0.09 0.00	0.11 0.06	0.06 0.04	0.24 0.01	0.05 0.17	---	0.01 0.00	0.01 0.00
9:WT2 (psnr-bw01.slac.stanford.edu)	0.62 0.82	0.50 0.52	0.51 0.58	0.41 0.28	0.00 0.85	0.54 0.73	0.31 0.54	0.71 0.91	0.00 0.11	---	---

bwctl last 30 days



# Example of Network Monitoring

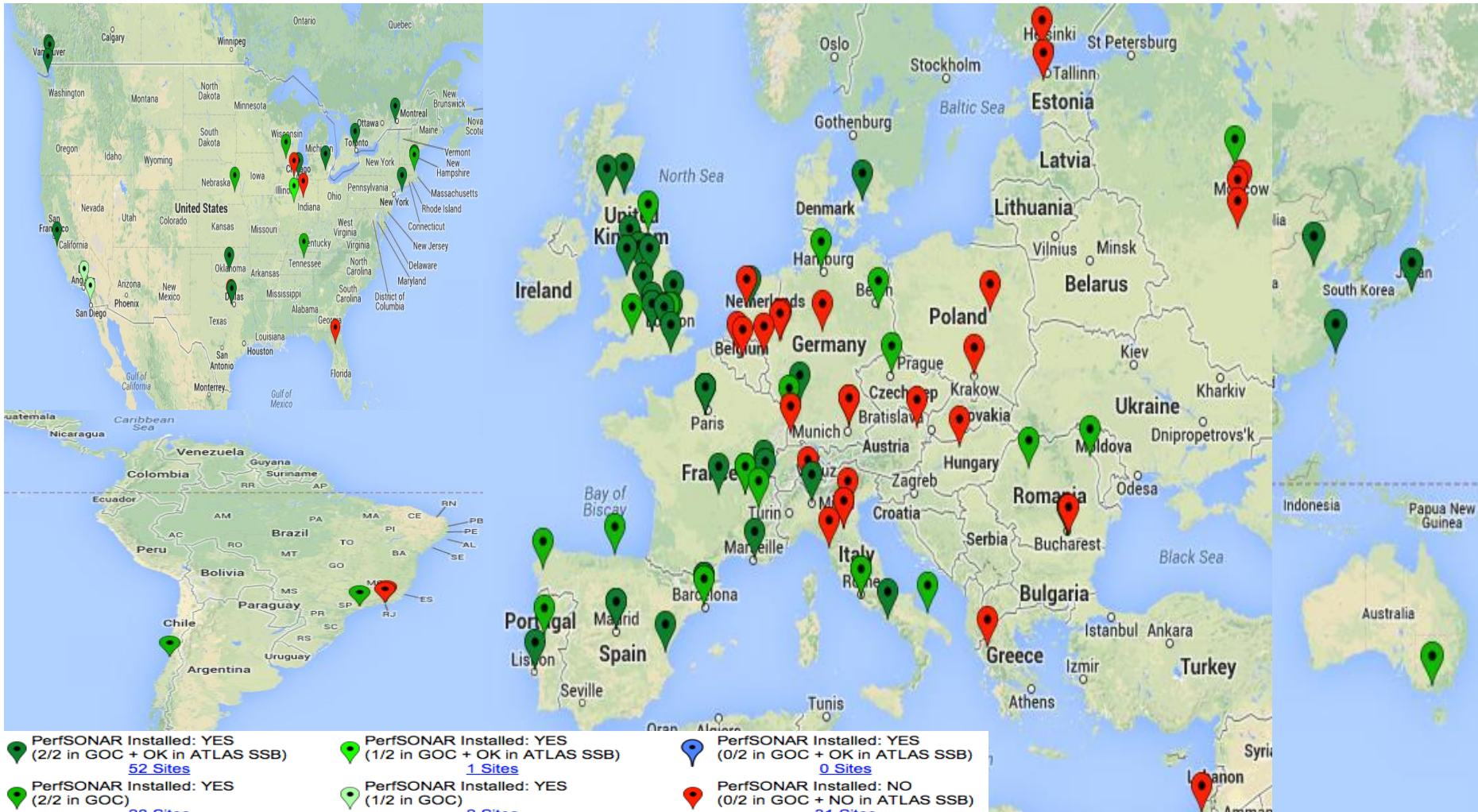
ATLAS aggregates complementary network information in the Site Status Board

Topology						FTS transfer (per file)						perfSONAR						FAX xrdcp rate	WAN data access (WN-SE)
SrcSite	SrcCloud	SrcTier	DstSite	DstCloud	DstTier	DDM Sonar						perfSONAR							
						AvgBRS (MB/s)	EvS	AvgBRM (MB/s)	EvM	AvgBRL (MB/s)	EvL	MinThr (MB/s)	AvgThr (MB/s)	MaxThr (MB/s)	MinPL	AvgPL	MaxPL		
Taiwan- LCG2	TW	T1	RAL-LCG2	UK	T1	0.51+/-0.63	285	7.25+/-5.02	336	7.88+/-5.47	649	0.3	0.3	0.3	0.0	65.7	329.0	n/a	
Taiwan- LCG2	TW	T1	IN2P3-CC	FR	T1	0.52+/-0.66	55886	6.34+/-2.94	6121	16.10+/-6.07	1617	0.5	0.5	0.5	600.0	600.0	600.0	1.33	
TRIUMF- LCG2	CA	T1	Taiwan- LCG2	TW	T1	0.41+/-0.41	400	1.25+/-0.24	38	2.89+/-1.30	5	0.4	0.5	0.6	0.0	0.0	1.0	n/a	
pic	ES	T1	Taiwan- LCG2	TW	T1	0.04+/-0.09	162	0.00+/-0.00	0	0.00+/-0.00	0	0.3	0.6	0.8	0.0	0.0	0.0	n/a	
FZK-LCG2	DE	T1	Taiwan- LCG2	TW	T1	0.17+/-0.24	1178	1.01+/-0.23	505	16.93+/-11.49	5	0.5	1.3	2.2	0.0	0.0	0.0	n/a	
BNL-ATLAS	US	T1	RAL-LCG2	UK	T1	0.29+/-0.51	45183	3.71+/-1.71	2697	21.06+/-15.41	879	1.5	1.7	1.9	0.0	18.2	229.0	n/a	
Taiwan- LCG2	TW	T1	FZK-LCG2	DE	T1	0.83+/-1.08	280	4.70+/-2.82	36	16.37+/-9.10	125	1.9	2.0	2.3	0.0	0.1	2.0	1.15	
INFN-T1	IT	T1	Taiwan- LCG2	TW	T1	0.29+/-0.46	540	1.87+/-0.76	6	0.00+/-0.00	0	1.7	2.0	2.3	0.0	0.0	0.0	n/a	
pic	ES	T1	RAL-LCG2	UK	T1	0.61+/-0.31	5202	6.32+/-2.22	216	20.80+/-9.55	4	1.5	2.4	2.5	0.0	57.5	357.0	n/a	
BNL-ATLAS	US	T1	IN2P3-CC	FR	T1	1.63+/-2.05	101375	15.30+/-6.78	28627	39.26+/-12.94	5481	2.5	3.3	4.4	0.0	0.0	0.0	1.34	
NDGF-T1	ND	T1	Taiwan- LCG2	TW	T1	0.09+/-0.13	4488	1.40+/-0.62	67	19.33+/-0.81	5	3.7	3.8	4.3	0.0	0.0	0.0	n/a	
IN2P3-CC	FR	T1	Taiwan- LCG2	TW	T1	0.36+/-0.57	4641	3.58+/-2.00	3840	9.12+/-6.52	1067	3.3	4.2	5.3	0.0	0.0	0.0	n/a	
FZK-LCG2	DE	T1	RAL-LCG2	UK	T1	0.47+/-0.74	70705	7.44+/-6.32	7598	14.03+/-16.74	6770	2.8	4.4	9.9	0.0	24.2	193.0	n/a	
RAL-LCG2	UK	T1	Taiwan- LCG2	TW	T1	0.06+/-0.19	13355	0.96+/-0.34	528	0.00+/-0.00	0	6.7	6.7	6.7	0.0	0.6	5.0	n/a	

WAN data access (WN-SE)



# perfSONAR Deployment Status



# Conclusions

- Network Monitoring is a key component for WLCG Operation
- We are deploying a monitoring infrastructure based on perfSONAR-PS in the scope of WLCG Operations
- 70% of the infrastructure has been deployed
- Completing the deployment and optimizing the tests are the next steps