Laboratoire d'Annecy-le-Vieux de Physique des Particules

perfSONAR en pratique

Supervision réseau multi-domaine
dans le cadre de WLCG & LCG-France

Sources & Références :

- perfSONAR-PS http://www.perfsonar.net/
- Déploiement perfSONAR au sein de WLCG/OSG
- WLCG Ops Network & Transfer Metrics WG Laurent Caillat, Frédéric Schaer, FC
- GGUS : WLCG Network Throughput SUPPORT UNIT
- perfSONAR-MDM http://services.geant.net/PerfSONAR/

Université Chambéry Savoie

cnrs IN2P3 Les deux infinis

Contenu*

- Objectifs & Use Case des expériences

  les finalités du déploiement de perfSONAR

- Déploiement dans les sites LCG-FR

  l'infrastructure dont on dispose

- Evolutions récentes et développements en cours

Objectifs

- Monitoring réseau de bout en bout : optimiser l'efficacité des transferts de données et des *workflows des expériences*
- Deux activités du WG WLCG perfSONAR menées en //
  1. Déploiement et Opérations perfSONAR
  2. « Network et Transfer Metrics »

  L'une des premières actions a été de réunir les Use Cases des expériences

  Suivi-Contact FR : Laurent Caillat, Frédéric Schaer, FC

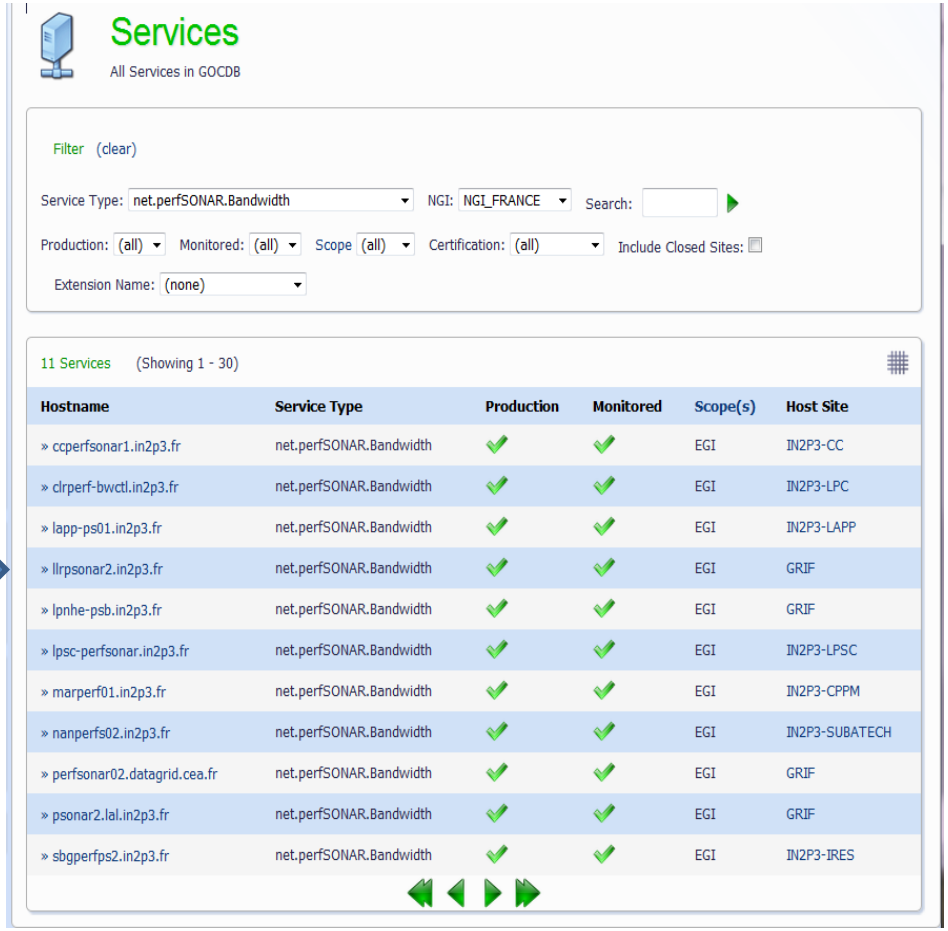
** Toutes les imprécisions, éventuelles erreurs sont de mon seul fait*

« Use case » des expériences

- Résumé
  - Identifier et comprendre ce qui se passe en cas de baisses de performances des transferts de données
  - Caractériser l'utilisation qui est faite des **ressources** Réseau
  - Identifier et localiser les "vrais " problèmes réseau des autres défaillances potentielles
  - Différentier les différentes sources de problèmes : réseau – transfert - stockage
    - Etre en mesure d'alerter les experts réseau
    - Disposer de procédures pour contacter les sites en cas de problèmes
  - Fournir des métriques réseau utilisable par les applications et les services des VOs
    - Identifier la "meilleure source" , type Cost-Matrix (ATLAS)
    - Caractériser les liens réseau niveau service / application

Déploiement dans les sites LCG-FR

- Supervision réseau Services perfSONAR : Bandwith, Latence
- 11 instances FR :
  - 2 hôtes distincts par site BWCTL, OWAMP déclarés dans la GOCDB
  - version pS 3.4.2
- CC-IN2P3, IRFU : dual stack IPv4/IPv6
- Avril 2015 : déploiement au LLR



Services

All Services in GOCDB

Filter (clear)

Service Type: net.perfSONAR.Bandwidth NGI: NGI_FRANCE Search:

Production: (all) Monitored: (all) Scope (all) Certification: (all) Include Closed Sites:

Extension Name: (none)

11 Services (Showing 1 - 30)

Hostname	Service Type	Production	Monitored	Scope(s)	Host Site
» ccperfsonar1.in2p3.fr	net.perfSONAR.Bandwidth	✓	✓	EGI	IN2P3-CC
» clrperf-bwctl.in2p3.fr	net.perfSONAR.Bandwidth	✓	✓	EGI	IN2P3-LPC
» lapp-ps01.in2p3.fr	net.perfSONAR.Bandwidth	✓	✓	EGI	IN2P3-LAPP
» llrpsonar2.in2p3.fr	net.perfSONAR.Bandwidth	✓	✓	EGI	GRIF
» lpnhe-psb.in2p3.fr	net.perfSONAR.Bandwidth	✓	✓	EGI	GRIF
» lpsc-perfsonar.in2p3.fr	net.perfSONAR.Bandwidth	✓	✓	EGI	IN2P3-LPSC
» marperf01.in2p3.fr	net.perfSONAR.Bandwidth	✓	✓	EGI	IN2P3-CPPM
» nanperfs02.in2p3.fr	net.perfSONAR.Bandwidth	✓	✓	EGI	IN2P3-SUBATECH
» perfsonar02.datagrid.cea.fr	net.perfSONAR.Bandwidth	✓	✓	EGI	GRIF
» psonar2.lal.in2p3.fr	net.perfSONAR.Bandwidth	✓	✓	EGI	GRIF
» sbgperfps2.in2p3.fr	net.perfSONAR.Bandwidth	✓	✓	EGI	IN2P3-IRES

Déploiement dans les sites LCG-FR

http://lapp_ps01.in2p3.fr/toolkit/

Points délicats derrière nous

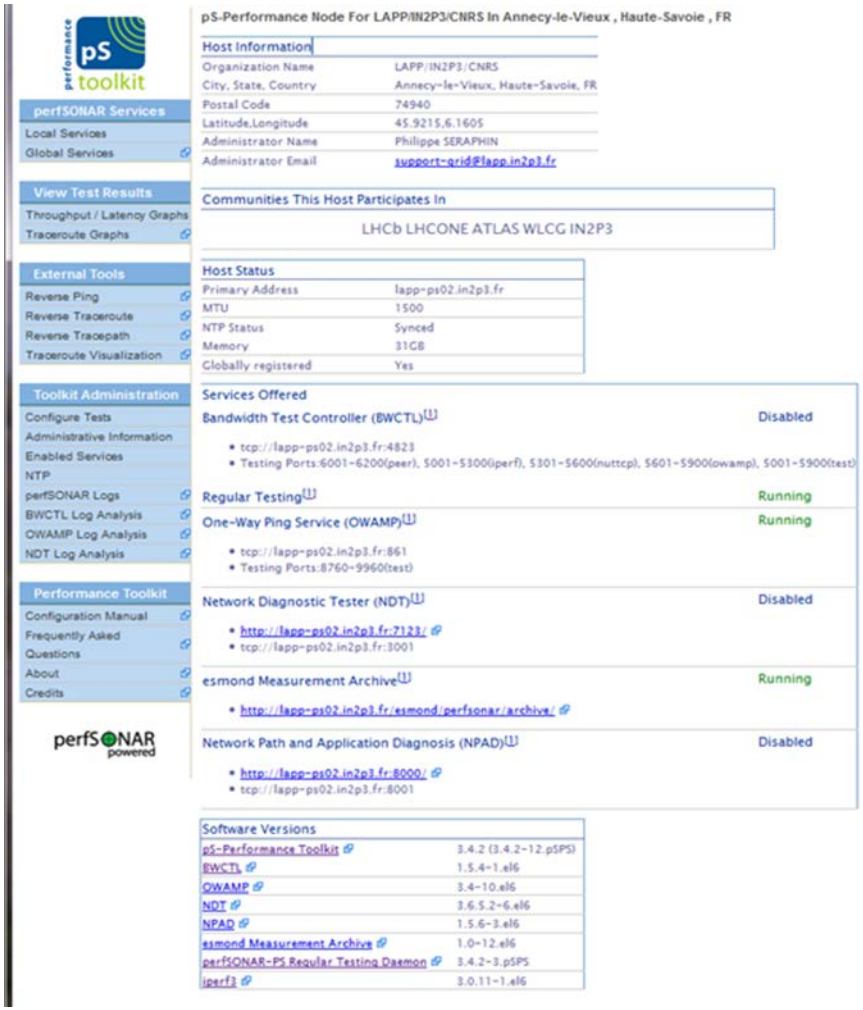
- Configuration des ports

6001-6200(peer), 5001-5300(iperf), 5301-5600(nuttcp), 5601-5900(owamp), 5001-5900(test)
8760-9960(test)

- Configuration centralisée et dynamique du maillage des instances perfSONAR (meshs) via auto-url

/opt/perfsonar_ps/mesh_config/etc
<mesh>
configuration_url
https://myosg.grid.iu.edu/pfmesh/mine/hostname/lapp-ps01.in2p3.fr
validate_certificate 0
required 1
</mesh>

http://lapp-ps01.in2p3.fr/toolkit/?format=json



Déploiement dans les sites LCG-FR

http://marperf01.in2p3.fr/toolkit/

Points délicats derrière nous

- Configuration des ports

6001-6200(peer), 5001-5300(iperf), 5301-5600(nuttcp), 5601-5900(owamp), 5001-5900(test)
8760-9960(test)

- Configuration centralisée et dynamique du maillage des instances perfSONAR (meshs) via auto-url

/opt/perfsonar_ps/mesh_config/etc
<mesh>
configuration_url https://myosg.grid.iu.edu/pfmesh/mine/hostname/marperf01.in2p3.fr
validate_certificate 0
required 1
</mesh>

Cas inexpliqué …

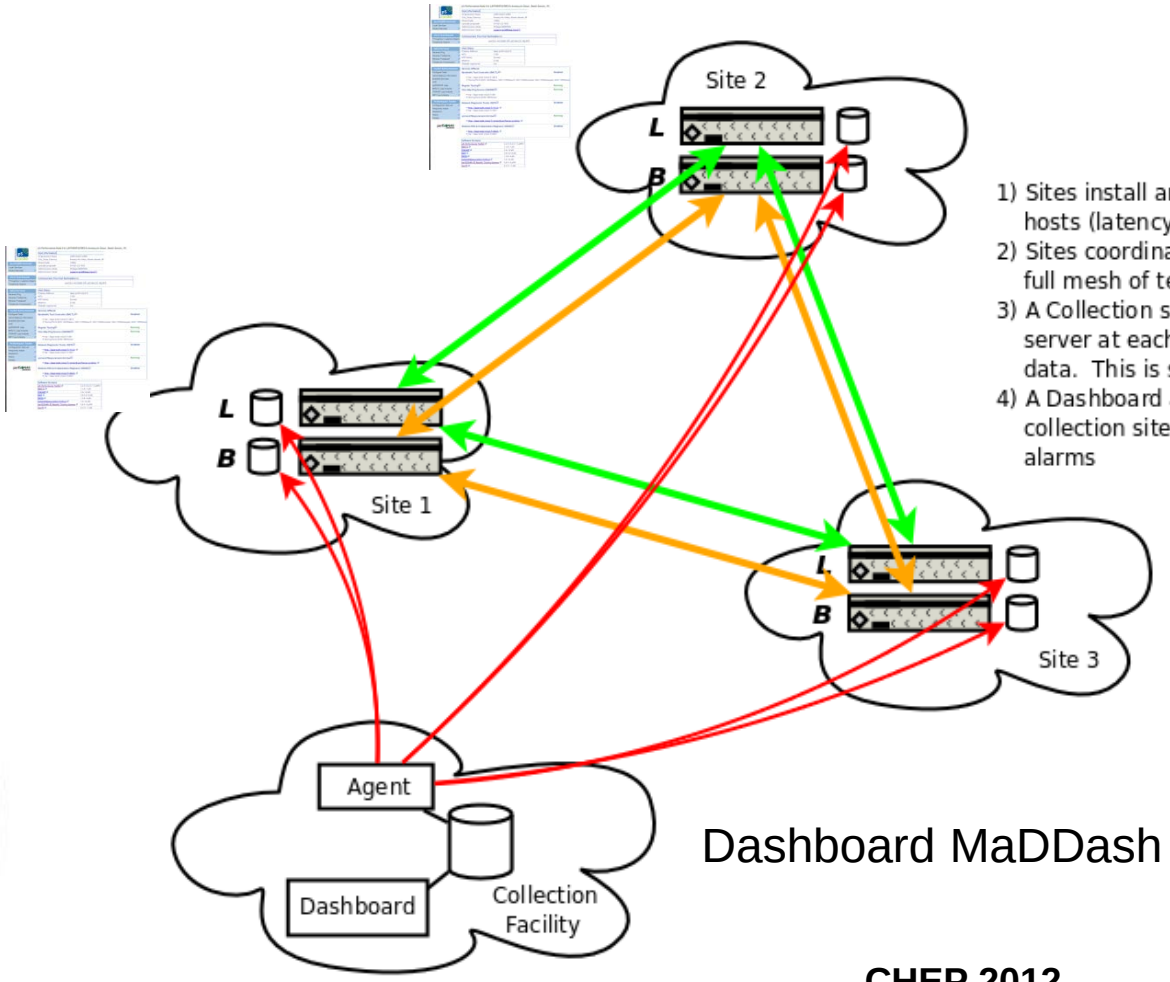
http://marperf01.in2p3.fr/toolkit/?format=json

{
"ntp":{"synchronized":"1"},
"location":{"country":"FR",
"longitude":"5.441",
"city":"Marseille",
"latitude":"43.231",
"zipcode":"13009",
"state":"France"},
"meshes":[[]],
"services":[
{"is_running":"yes",
"addresses":null,
"version":"3.0.111.el6",
"name":"iperf3"},

…

"globally_registered":1,
"host_memory":16
}

l'infrastructure pS initiale



1) Sites install and configure 2 measurement hosts (latency testing and bandwidth testing)
2) Sites coordinate with each other to create a full mesh of testing data (orange and green lines)
3) A Collection site has probes that contact each server at each site to pull measurement (red lines) data. This is stored in local storage
4) A Dashboard application consults local data at collection site to display results/generate alarms

Dashboard MaDDash

CHEP 2012
Shawn McKee/University of Michigan

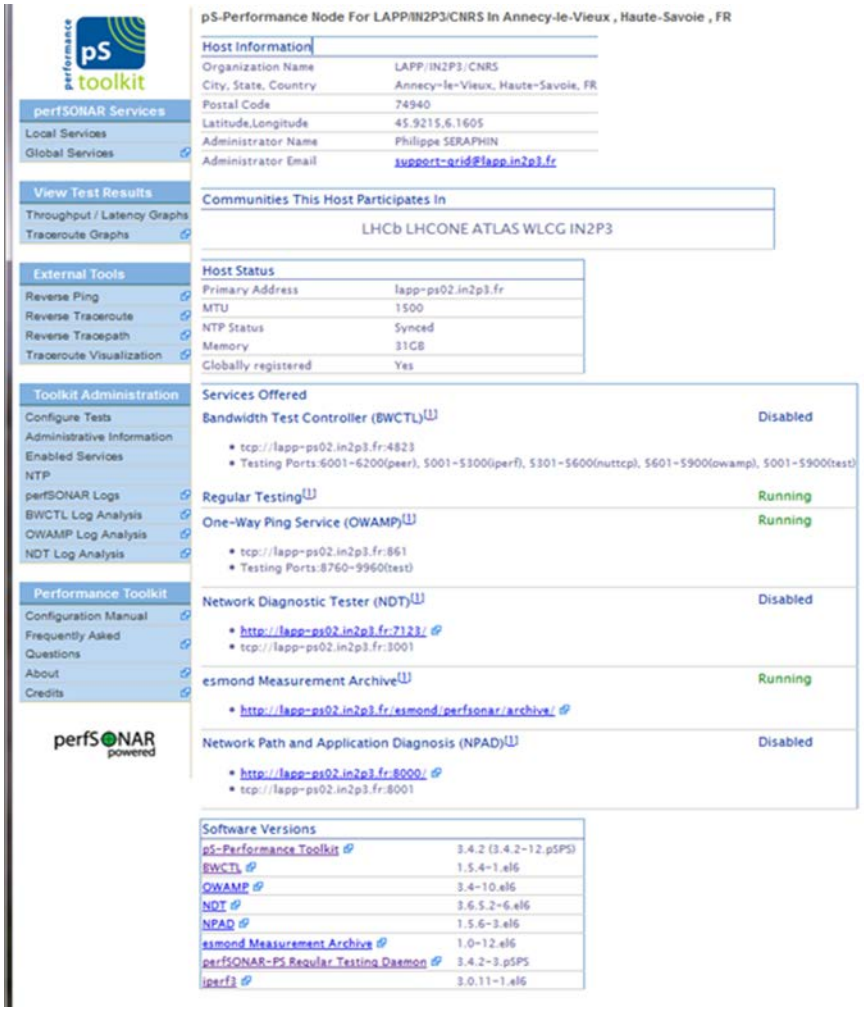
Utilisation de l'interface pS

http://hostname.in2p3.fr/toolkit/

- Possibilité de contrôle Local
- Consultation des résultats des tests réguliers archivés localement
- Mesures :
  1. Bande-Passante (BWCTL)
  2. Latence/pertes (OWAMP)
- Infos traceroute/tracepath pour connaitre la topologie

Documentation détaillée

http://www.perfsonar.net/use/



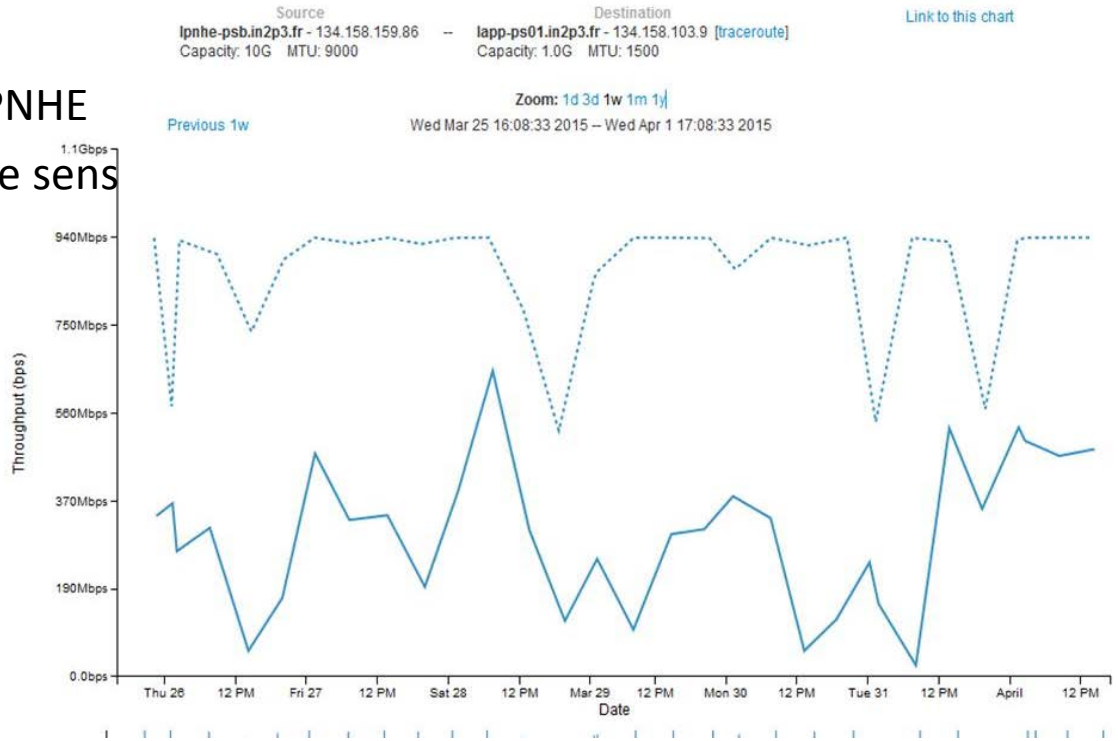
Utilisation de l'interface pS

Mesures de Bande passante initiées depuis le LAPP

- lapp-ps01.in2p3.fr interfacé à 1G MTU 1500
- lpnhe-psb.in2p3.fr interfacé à 10G MTU 9000

~ 940 Mbps LAPP => LPNHE

~ 370 Mbps dans l'autre sens



http://lapp-ps01.in2p3.fr/serviceTest/graphWidget.cgi?source=134.158.159.86&dest=134.158.103.9&url=http://localhost/esmond/perfsonar/archive/#

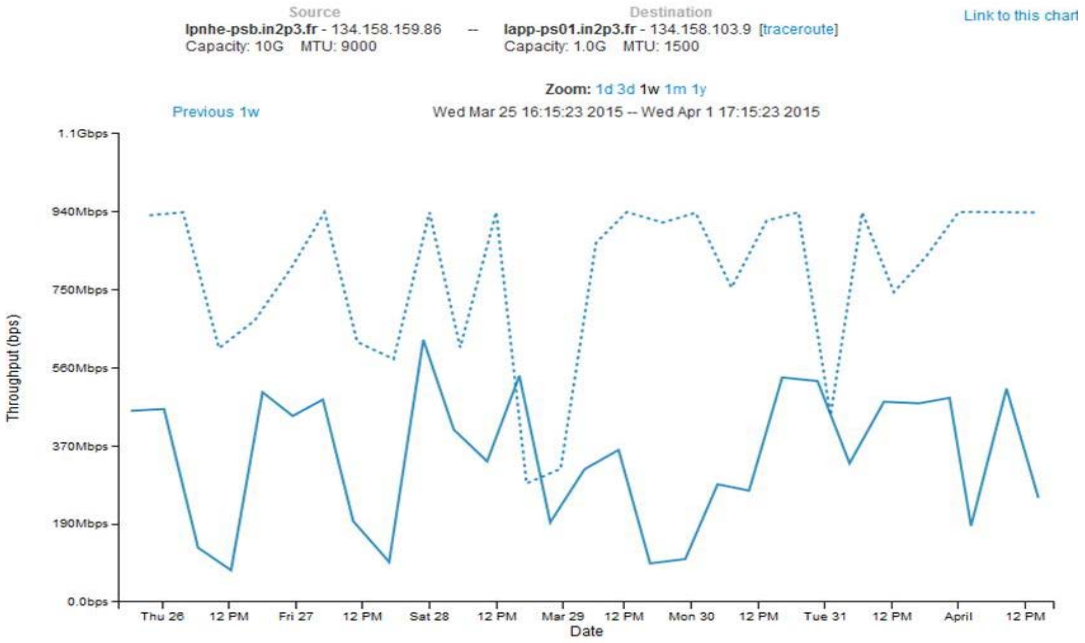
Utilisation de l'interface pS

Mesures de Bande passante initiées depuis le LAPP

- lapp-ps01.in2p3.fr interfacé à 1G MTU 1500
- lpnhe-psb.in2p3.fr interfacé à 10G MTU 9000

Même constat !
Pas de conclusions
en dépit de quelques
Investigations : MTU, TSO

Passage à 10 G envisagé

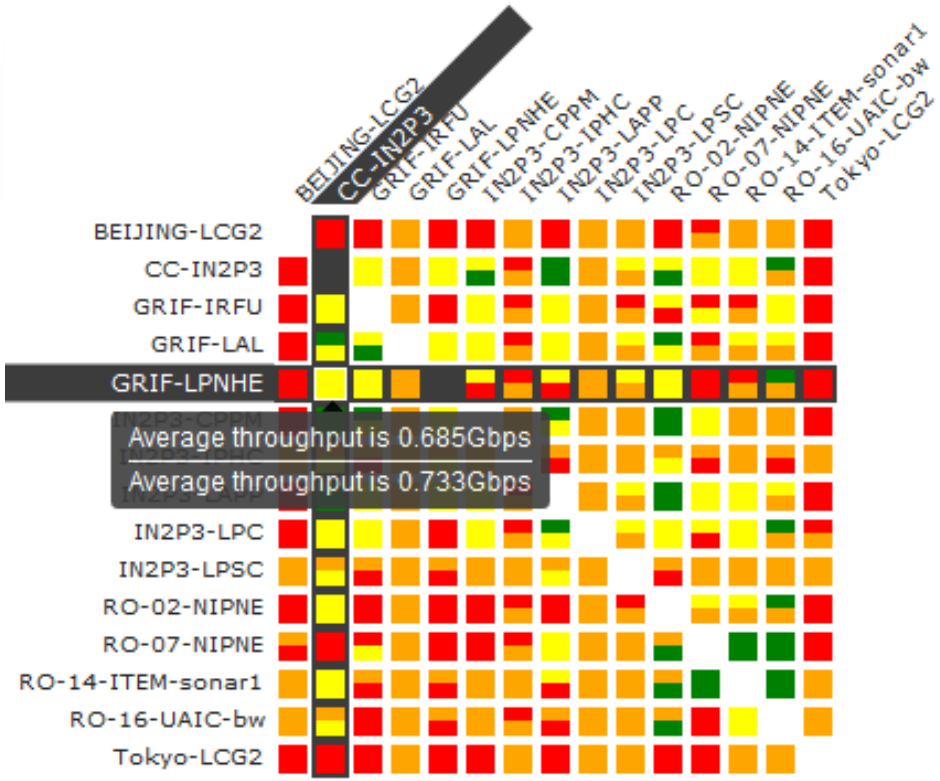


https://lpnhe-psb.in2p3.fr/serviceTest/graphWidget.cgi?source=134.158.159.86&dest=134.158.103.9&url=http://localhost/esmond/perfsonar/archive/#

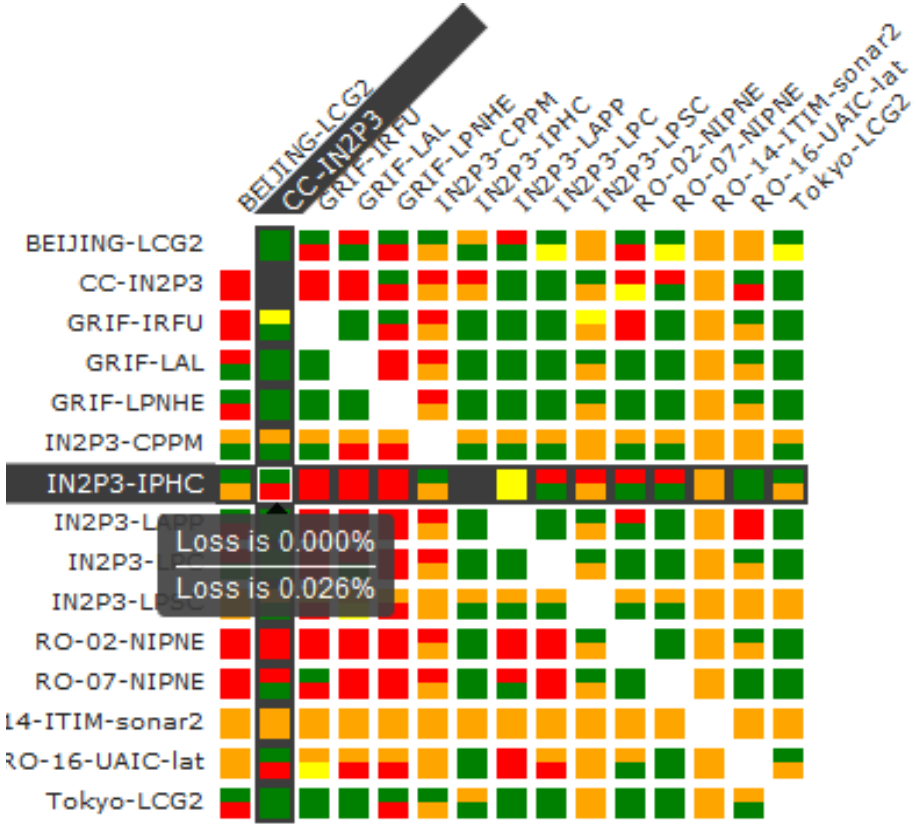
Dashboard MADDash

MaDDash https://maddash.aglt2.org/maddash-webui (prototype URL)

BWCTL (bande passante)



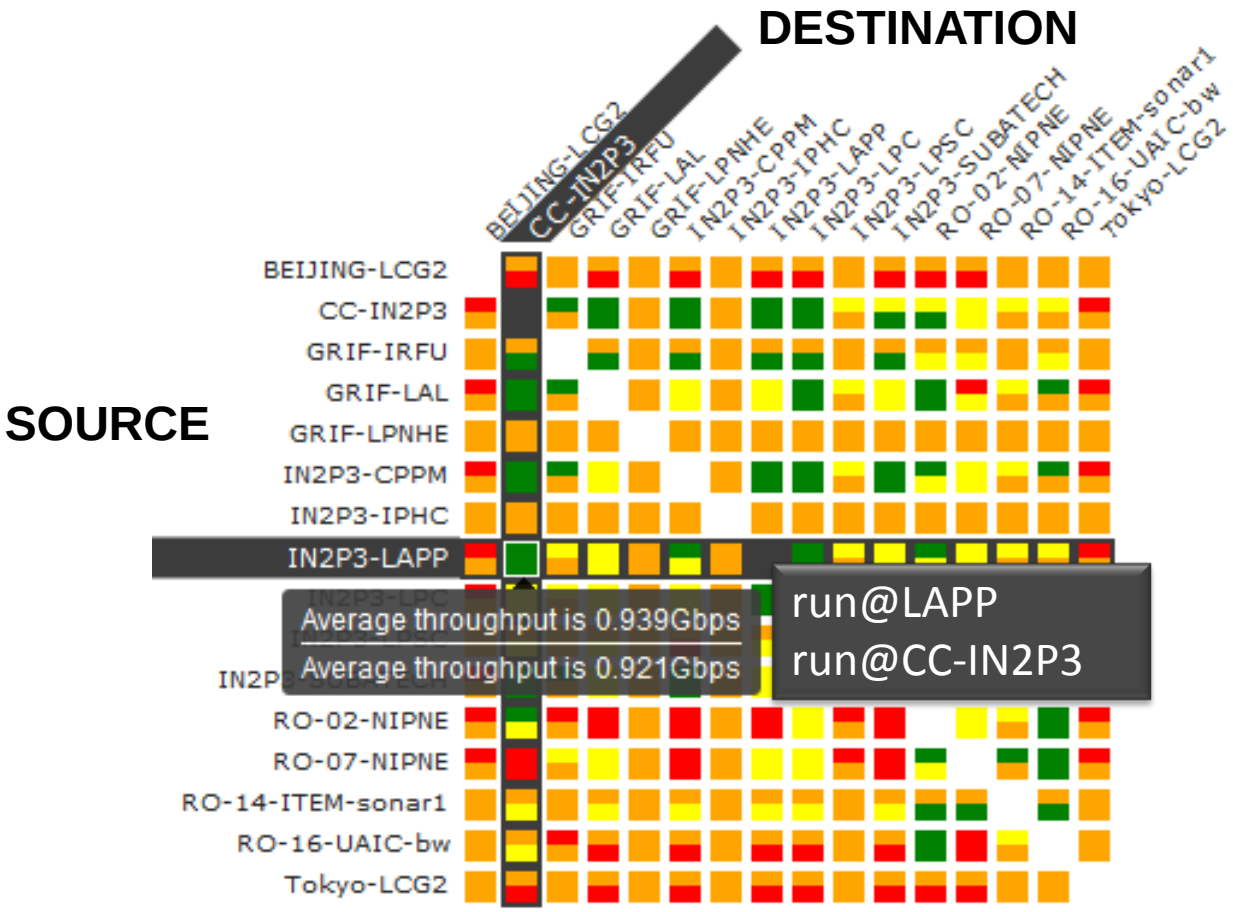
OWAMP (pings/pertes)



FR mesh Situation présentée par F.Schaer en avril 2014

Dashboard MADDash

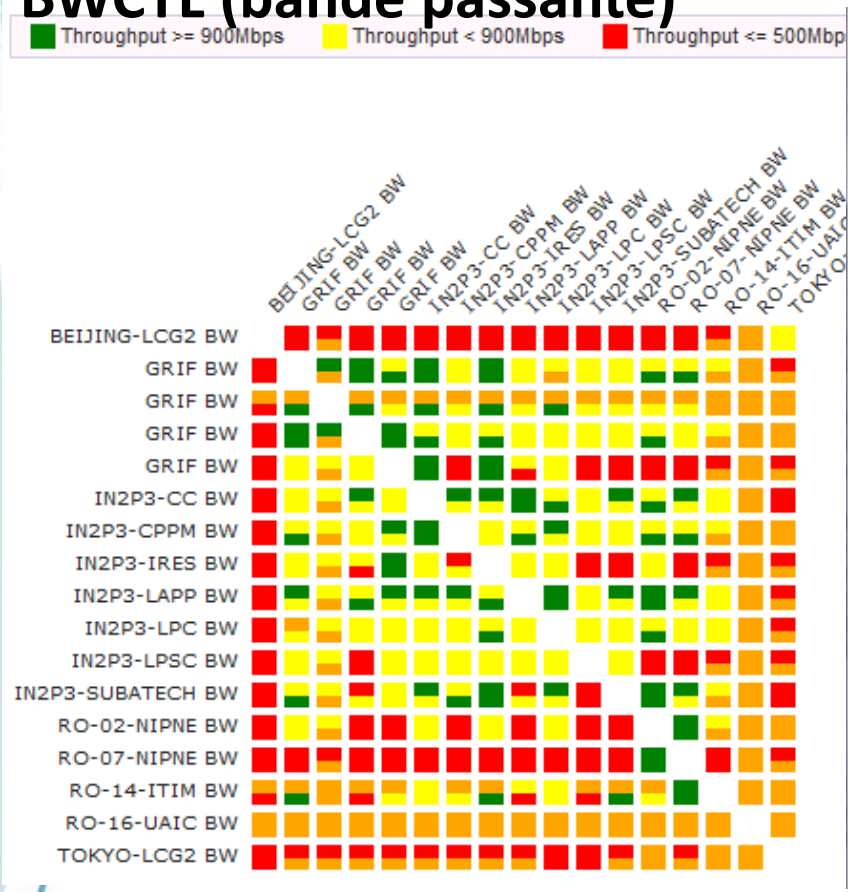
MaDDash https://maddash.aglt2.org/maddash-webui (prototype URL)



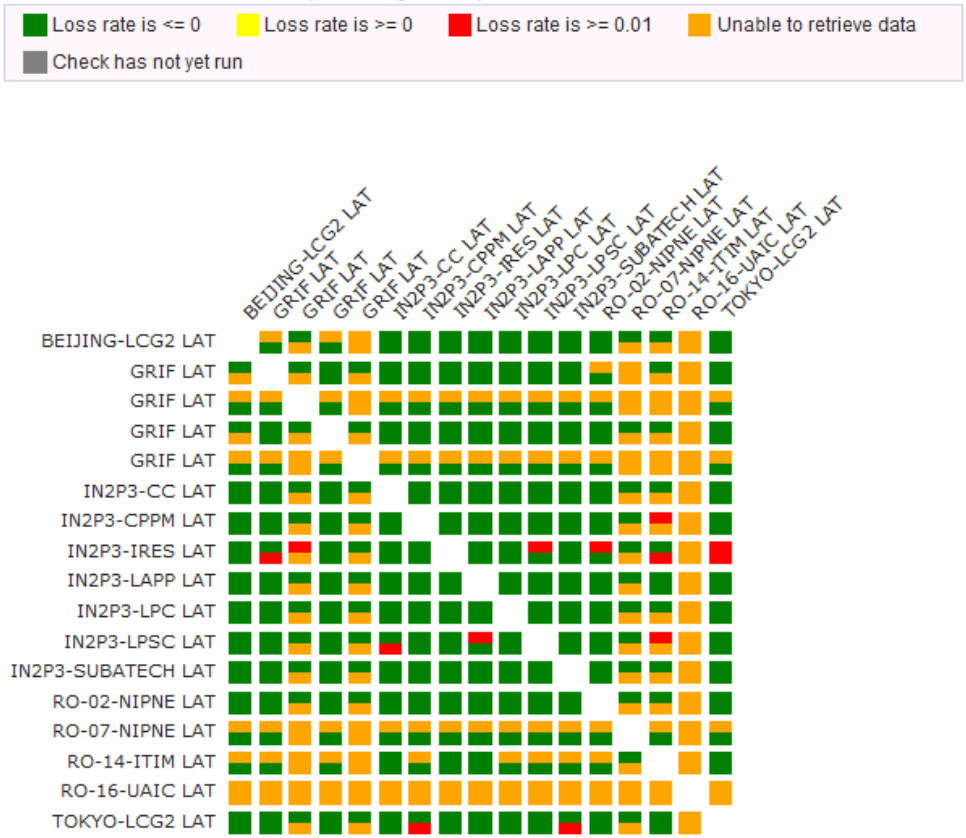
Dashboard MADDash

MaDDash https://maddash.aglt2.org/maddash-webui (prototype URL)

BWCTL (bande passante)



OWAMP (pings/pertes)



Evolutions de l'infrastructure pS

- Evolutions récentes
  - Infrastructure de monitoring des instances pS
  - Interface de configuration des meshs
- En cours : mise en place progressive de tests de latence entre l'ensemble des sites WLCG
  - Il est envisager de désactiver les meshs regionaux pour limiter l'impact des tests
  - https://maddash.aglt2.org/maddash-webui/index.cgi?dashboard=Latency tests between all WLCG hosts
- Encore en développement
  - Central repository (Archive centralisée de l'ensemble des métriques)
  - API pour les applications (FTS…)
  - Outil de mise en correspondance des topologies perfSONAR-réseau Applications-SE pour corréler les métriques perfSONAR avec celles élaborées au niveau applicatif : FTS, PhEDEx, tests Sonar ATLAS

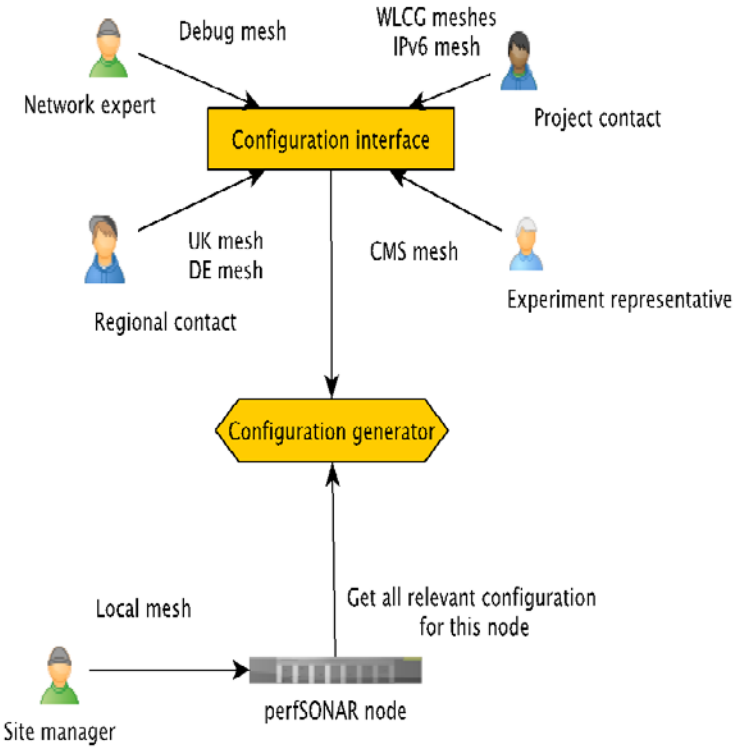
Interface de configuration des meshs

https://oim.grid.iu.edu/oim/meshconfig

Accès autorisé pour Frédéric Schaer, Laurent Caillat, FC

Prise en main aisée lors l'ajout des instances du LLR

- Chaque instance persSONAR peut rejoindre différentes meshs
- Une seul configuration du service à partir un auto-url : http://myosg.grid.iu.edu/pfmesh/*hostname*
- Interface de configuration permet de reconfigurer dynamiquement le maillage des perfSONARs



CHEP 2015
Shawn McKee/University of Michigan

Infrastructure de Monitoring

https://maddash.aglt2.org/WLCGperfSONAR/check_mk (prototype URL)

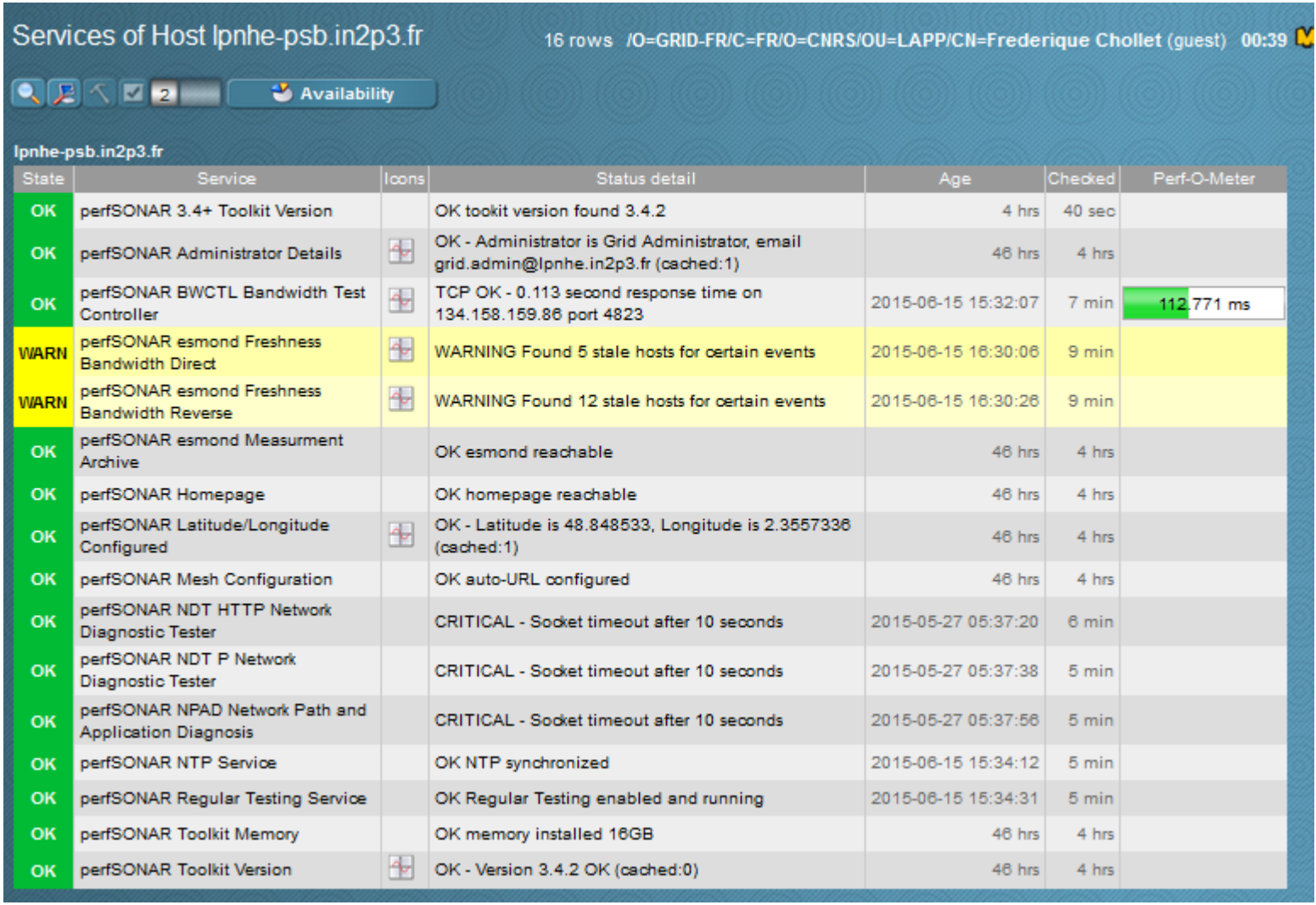
Outil : OMD/Check_MK

Host group FR

state	Host	Icons	Alias	OK	Wa	Un	Cr	Pd
UP	atrogr007.nipne.ro		atrogr007.nipne.ro	14	2	0	0	0
UP	atrogr009.nipne.ro		atrogr009.nipne.ro	13	2	0	1	0
UP	ccperfsonar1.in2p3.fr		ccperfsonar1.in2p3.fr	14	2	0	0	0
UP	ccperfsonar2.in2p3.fr		ccperfsonar2.in2p3.fr	14	2	0	0	0
UP	clrperf-bwctl.in2p3.fr		clrperf-bwctl.in2p3.fr	14	2	0	0	0
UP	clrperf-owamp.in2p3.fr		clrperf-owamp.in2p3.fr	14	2	0	0	0
UP	lapp-ps01.in2p3.fr		lapp-ps01.in2p3.fr	14	2	0	0	0
UP	lapp-ps02.in2p3.fr		lapp-ps02.in2p3.fr	14	2	0	0	0
UP	llrpsonar1.in2p3.fr		llrpsonar1.in2p3.fr	14	2	0	0	0
UP	llrpsonar2.in2p3.fr		llrpsonar2.in2p3.fr	14	2	0	0	0
UP	lpnhe-psb.in2p3.fr		lpnhe-psb.in2p3.fr	14	2	0	0	0
UP	lpnhe-psl.in2p3.fr		lpnhe-psl.in2p3.fr	14	2	0	0	0
UP	lpsc-perfsonar.in2p3.fr		lpsc-perfsonar.in2p3.fr	14	2	0	0	0
UP	lpsc-perfsonar2.in2p3.fr		lpsc-perfsonar2.in2p3.fr	14	2	0	0	0
UP	marperf01.in2p3.fr		marperf01.in2p3.fr	13	2	0	1	0
UP	marperf02.in2p3.fr		marperf02.in2p3.fr	13	2	0	1	0
UP	nanperfs01.in2p3.fr		nanperfs01.in2p3.fr	14	2	0	0	0
UP	nanperfs02.in2p3.fr		nanperfs02.in2p3.fr	14	2	0	0	0
DOWN	perfsonar-grid.uaic.ro		perfsonar-grid.uaic.ro	3	0	6	7	0
UP	perfsonar.ihep.ac.cn		perfsonar.ihep.ac.cn	14	2	0	0	0
UP	perfsonar01.datagrid.cea.fr		perfsonar01.datagrid.cea.fr	10	2	0	4	0
UP	perfsonar02.datagrid.cea.fr		perfsonar02.datagrid.cea.fr	10	2	0	4	0

OMD/Check_MK

https://maddash.aglt2.org/WLCGperfSONAR/check_mk/view.py?view_name=host&host=lpnhe-psb.in2p3.fr



Services of Host lpnhe-psb.in2p3.fr — 16 rows — /O=GRID-FR/C=FR/O=CNRS/OU=LAPP/CN=Frederique Chollet (guest) — 00:39

Availability

lpnhe-psb.in2p3.fr

State	Service	Icons	Status detail	Age	Checked	Perf-O-Meter
OK	perfSONAR 3.4+ Toolkit Version		OK tookit version found 3.4.2	4 hrs	40 sec	
OK	perfSONAR Administrator Details		OK - Administrator is Grid Administrator, email grid.admin@lpnhe.in2p3.fr (cached:1)	46 hrs	4 hrs	
OK	perfSONAR BWCTL Bandwidth Test Controller		TCP OK - 0.113 second response time on 134.158.159.86 port 4823	2015-06-15 15:32:07	7 min	112.771 ms
WARN	perfSONAR esmond Freshness Bandwidth Direct		WARNING Found 5 stale hosts for certain events	2015-06-15 16:30:06	9 min	
WARN	perfSONAR esmond Freshness Bandwidth Reverse		WARNING Found 12 stale hosts for certain events	2015-06-15 16:30:26	9 min	
OK	perfSONAR esmond Measurment Archive		OK esmond reachable	46 hrs	4 hrs	
OK	perfSONAR Homepage		OK homepage reachable	46 hrs	4 hrs	
OK	perfSONAR Latitude/Longitude Configured		OK - Latitude is 48.848533, Longitude is 2.3557336 (cached:1)	46 hrs	4 hrs	
OK	perfSONAR Mesh Configuration		OK auto-URL configured	46 hrs	4 hrs	
OK	perfSONAR NDT HTTP Network Diagnostic Tester		CRITICAL - Socket timeout after 10 seconds	2015-05-27 05:37:20	6 min	
OK	perfSONAR NDT P Network Diagnostic Tester		CRITICAL - Socket timeout after 10 seconds	2015-05-27 05:37:38	5 min	
OK	perfSONAR NPAD Network Path and Application Diagnosis		CRITICAL - Socket timeout after 10 seconds	2015-05-27 05:37:56	5 min	
OK	perfSONAR NTP Service		OK NTP synchronized	2015-06-15 15:34:12	5 min	
OK	perfSONAR Regular Testing Service		OK Regular Testing enabled and running	2015-06-15 15:34:31	5 min	
OK	perfSONAR Toolkit Memory		OK memory installed 16GB	46 hrs	4 hrs	
OK	perfSONAR Toolkit Version		OK - Version 3.4.2 OK (cached:0)	46 hrs	4 hrs	

Conclusions

- L'infrastructure perfSONAR existe ; elle permet un monitoring réseau de bout-en-bout
- Souciis de déploiement derrière nous ; la maintenance de l'infrastructure pS a certes un coût
- La ressource réseau est particulièrement fiable
  - D'éventuels incidents réseau ne justifient pas tout cela
- Support Unit GGUS pour traiter les cas de baisses performances observées au niveau applicatif (transfert de données) particulièrement sensibles pour les expériences
- A court terme, les expériences devraient également disposer de métriques réseau pour rendre leurs applications « network-aware »
- Au-delà, il y a sans doute d'autres perspectives d'utilisation.