WLCG perfSONAR Deployment Information

This page documents setting up perfSONAR-PS for WLCG Sites. Our goal is to support all WLCG sites in deploying, configuring and registering perfSONAR instances so we can gather network metrics about the network paths to other WLCG sites. WLCG has set a deadline of April 1, 2014 for ALL WLCG sites to have deployed two perfSONAR-PS instances: one latency and one bandwidth. NOTE: After the release of perfSONAR-PS v3.4 it should be possible to merge the bandwidth and latency tests onto a single physical node.

Any question concerning perfSONAR deployment in WLCG can be addressed at wlcg-ops-coord-tf-perfsonar@cernNOSPAMPLEASE.ch

A quick note on the physical location for the WLCG perfSONAR-PS instances: Our strong recommendation is to co-locate two perfSONAR-PS nodes with the sites primary grid-storage. The reason is that we want the perfSONAR instances to measure as much of the network path as is possible, end-to-end. The perfSONAR-PS measurements are intended to represent what the network is doing end-to-end and can be used to differentiate network problems from end-host/storage/software problems.

WLCG perfSONAR-PS Requirements

There are a few requirements to deploy perfSONAR-PS toolkit instances at WLCG sites: hosting location, network setup and hardware to run the services

Physical location identified

WLCG sites need to plan to host the recommended two perfSONAR-PS hosts. Hardware details are available at the URL below under "Hardware Guidelines". A typical installation would have two 1U nodes installed at the same "network" location as the site's primary storage. This is to ensure we are measuring as much of the end-to-end network path for the site's storage systems.
Network connectivity

The two perfSONAR-PS instances will both require publicly reachable addresses and DNS names setup. Of course sites must also provide two network ports to connect the systems, appropriate for the type of network connection the systems require (RJ45, fiber, etc.). For DNS names we recommend the following:

- The DNS name should have a two digit number to allow for future additional installations
  - The 'odd' numbers should correspond to latency instances (example ...01 should be the first perfSONAR-PS latency instance)
  - The 'even' number should correspond to bandwidth instances (example ...02 should be the first perfSONAR-PS bandwidth instance)
- The name should start with 'ps' or 'perfsonar'
- The "site" info is assumed to come from the DNS domain but if not it might be included in the DNS name.

Using the above recommendations the perfSONAR-PS installations at the University of Michigan ATLAS Great Lakes Tier-2 (AGLT2) might be:

- psum01.aglt2.org - The first latency node at the AGLT2 UM site
- psum02.aglt2.org - The first bandwidth node at the AGLT2 UM site

Hardware Guidelines

The perfSONAR project has setup a hardware recommendation page at http://psps.perfsonar.net/toolkit/hardware.html. All WLCG sites should be able to purchase Dell R310 or Dell R610 hosts at "LHC" pricing. Update: these 11th generation nodes are no longer available but the R320/R420 or R620 Dell systems should be perfectly suitable. Contact your Dell representative to see about pricing and availability.

Plans for WLCG perfSONAR-PS Installs

For all WLCG sites, we want to configure a "full-mesh" of tests. We plan on having:

- Latency (OWAMP) tests to all members of your Tier-1 cloud and to the Tier-1 sites
- Bandwidth (BWCTL) tests to all members of your Tier-1 cloud (30 second test each way every 6 hours) and to all WLCG sites (30 second test, once / two weeks)
- Traceroute tests to all WLCG sites every hour
- Ping tests (via PINGER) to all WLCG sites 3 times / hour.

These will be controlled by various "Mesh configurations". The primary "configuration" step for sites is now simplified to configuring which "Meshes" the site wishes to participate in. We imagine sites will be participating in intra-regional meshes, inter-regional meshes as well as the complete WLCG mesh. Initial plan is to provide:

- Each Tier-2 site should include a configuration for their Tier-1 cloud
  - One mesh per Tier-1 cloud (possibly experiment specific): Bandwidth and Latency tests
  - One mesh setting up "disjoint" testing to all the Tier-1s (Tier-1 "region" instances will test to all Tier-1s but not to each other): Bandwidth and Latency tests
- Each Tier-1 site will have an LHCOPN mesh
  - One mesh for the LHCOPN: Bandwidth and Latency tests
- WLCG-wide testing in a complete mesh
  - One mesh for ALL sites that only includes Traceroute and Ping testing

Parameters of each scenario will be different.
WLCG perfSONAR-PS Toolkit Installation Instructions

As of this writing the perfSONAR-PS Toolkit v3.3.2 is now available as a final release.

- Sites will need to plan on upgrading/updating as soon as possible.
- There very well may be bugs or problems (feedback encourage per the release notes at the URL above)

The quick start link for v3.3 is http://code.google.com/p/perfsonar-ps/wiki/pSPerformanceToolkit33

- Unless you have a good reason to choose a different version we are recommending you install (or upgrade-to) the NetInstall CentOS 6 64-bit version which will install to the local system disk. The system can then use yum to update itself.

Note that the above instructions show details about installing and configuring. The configuration notes are in the next section below, so please read ahead before clicking off to follow the quick-start links above.

Upgrading Existing Instances

Now that v3.3.2 (see http://psps.perfsonar.net/toolkit/releasenotes/pspt-3_3_2.html) is available we want all sites to upgrade.

WLCG perfSONAR-PS Configuration

Once the physical boxes are installed and operational, we need to configure the needed set of scheduled tests that the boxes should be running. For sites that are upgrading existing instances you should already have a working configuration (see UpgradePS). For sites that are installing new instances, you should start at https://code.google.com/p/perfsonar-ps/wiki/pSPerformanceToolkit33 and see "Configuring New Installs" below

Configuring New Installs

Some additional information for WLCG sites:

- After installing (either the NetInstall or LiveCD versions) you will need to setup the services running on each type of node. This configuration is handled by including the appropriate "mesh" configuration(s). See below for how to configure your mesh agent_configuration.conf file. FYI, our convention so far has been to make the first node (by name or IP) the Latency node and the second node the Bandwidth node (see networking info above). This is also easy to manually configure for additional tests or site specific tests by using the Web GUI and selecting Enabled Services on the left hand navigation panel under Toolkit Administration. You can select the button at the bottom for enabling only Latency or only Bandwidth services. On the Latency node you should make sure to enable the two Traceroute services (the MA and Scheduler). Note this should be done if you click the "Latency only services" button. **If you want SSH access enabled but sure to click the checkbox next to 'SSH' before you "Save".**
- Each site should fill out the appropriate Administrative Information (under Toolkit Administration on left of Web GUI).
The "Host Location" field should be filled out with a complete address (suitable for locating via GoogleMaps or similar).

- The latitude and longitude can be found using Google Maps (https://www.google.com/maps). Zoom in on Google maps and right-click the location of your perfSONAR-PS installation and select "What's here". The latitude and longitude will then be shown in the Google search box on the page.

- The Communities section (see http://code.google.com/p/perfsonar-ps/wiki/pSPerformanceToolkit322#Communities) should have WLCG added in addition to whatever other communities the site wants to list (LHCONE, ATLAS, CMS, USATLAS, LHC, etc.).

  - The NTP servers need to be setup carefully for the Latency node. Ideally at least 4 good servers should be configured (add local or regional ones if they are not in the distributed setup).
  - Firewalls may be an issue (See comments in table above). If you suspect your site will block ANY of the WLCG sites, can you update your firewalls to allow the ports shown in the FAQ here http://psps.perfsonar.net/toolkit/FAQs.html#Q6

Configuring Sites for Participation in WLCG and Other Meshes

All needed tests for WLCG are configured via so called "Mesh-configurations". All WLCG sites should configure their instances to participate in at least two meshes: the WLCG mesh and a local cloud mesh. The mesh-configurations are centrally managed and define all sites, tests and test parameters.

The mesh configuration for each mesh is stored in a .json file that can be retrieved from a well-known URL. For WLCG we will have a number of meshes available as noted in the previous section. Each site will need to configure is mesh-agent appropriately AND contact their cloud/region mesh coordinator to ensure their hosts are appropriately added.

The first step for a new site joining the WLCG perfSONAR infrastructure is to make sure they are included in the appropriate mesh-configuration. For now, this needs to be done "manually" but our goal is to automate this process by using the OIM/GOCDB registration information to create/update meshes as needed.
To make sure your hosts are appropriately included, please see the contact details on MeshUpdates.

Next, sites should configure the agent on their perfSONAR-PS toolkit instations to use the appropriate mesh(es) as follows:

- Make sure your perfSONAR-PS node is up-to-date: `yum update`
- Configure the mesh agent to point to the WLCG json file `test-wlcg-all.json` by editing the perfSONAR-PS host's `/opt/perfsonar_ps/mesh_config/etc/agent_configuration.conf` and adding the following URL in its own `<mesh>...</mesh>` block. See https://twiki.cern.ch/twiki/bin/view/LCG/MeshRegionList for information on the configuration:

  ```xml
  <mesh>
  validate_certificate         0
  </mesh>
  ```

  This line replaced the default `configuration_url` near the beginning. Sites should also add additional `configuration_url` lines corresponding to other regional meshes as appropriate.

- Configure the mesh agent to point to the appropriate Tier-1 region JSON file by adding the following URL in its own `<mesh>...</mesh>` block:

  ```xml
  <mesh>
  validate_certificate         0
  </mesh>
  ```

  Where `REGION` is the two character country code for the Tier-1 region and `VO` is either the "experiment" ('atlas','cms','lhcb', etc.) or 'all'. The complete list by Tier-1 region is shown in MeshRegionList.

- **NOTE:** every mesh URL needs to be added in its own `<mesh>...</mesh>` block!
- Also sites need to edit the `agent_configuration.conf` to customize the `admin_email` setting appropriate for each site. For AGLT2 the example is:

  ```
  admin_email     smckee@umich.edu
  admin_email     laurens@pa.msu.edu
  ```

- You can configure your `agent_configuration.conf` as above, even if you have not yet contacted your cloud/region mesh coordinator but you will get an error when the generation step is performed.

  - **We recommend you set** `use_toolkit` 1 **to allow the pS-Performance Toolkit's configuration daemon to save the configuration and restart the services**
  - **Last step is to make sure** `skip_redundant_tests` 1 **is set (uncomment this line at the end of the `agent_configuration.conf` file)**

Once you update the `agent_configuration.conf` file a crontab entry

```
/etc/cron.d/cron-mesh_config_agent
```

will generate your configuration nightly just after midnight. It is recommended to generate a new configuration manually to test things via:

```bash
sudo -u perfsonar /opt/perfsonar_ps/mesh_config/bin/generate_configuration
```

*NOTE: In perfSONAR-PS version 3.3 and 3.3.1 there is a known issue with the `generate_configuration` being very slow. This was addressed in 3.3.2*
Network Access (Firewalls)

As of release 3.3.2, the perfSONAR-PS Toolkit instances come preconfigured with an `iptables` setup based upon the requirements shown at http://fasterdata.es.net/performance-testing/perfsonar/ps-howto/perfsonar-firewall-requirements/ However there may be additional firewall appliances or network ACLs in the local area network that could interfere with the proper operation of perfSONAR-PS. Please review the requirements at the FasterData link above and ensure the proper access to your nodes is provided.

The following subnets may need some additional access:

- **CERN/WIGNER**
  - 188.184.0.0/15
  - 128.142.0.0/16
  - 137.138.0.0/17
- **maddash.aglt2.org** (a prototype WLCG monitoring instance)
  - 192.41.231.110/32
- **OSG monitoring subnet** (eventual "production" monitoring host-subnet) *(Added July 2014)*
  - 129.79.53.0/24

You can quickly setup iptables on your perfSONAR-PS Toolkit hosts for the above networks using the following lines just before #NTP rules

```
# Allow maddash.aglt2.org
-A INPUT -s 192.41.231.110/32 -j ACCEPT

# Allow OSG monitoring subnet
-A INPUT -s 129.79.53.0/24 -j ACCEPT

# Allow CERN nets
-A INPUT -s 188.184.0.0/15 -j ACCEPT
-A INPUT -s 128.142.0.0/16 -j ACCEPT
-A INPUT -s 137.138.0.0/17 -j ACCEPT
```

If you wish you can enable more specific rules. if you are participating in more detailed monitoring of your instances:

- **6556/tcp** If you are running the `check_mk-agent` and plugins (see WLCGperfSONARMonitoring link)
- **161/udp** If you are running `snmpd` for monitoring

**IMPORTANT:** As noted above, you also need to ensure any firewall appliances or network ACLs allow the ports listed in http://fasterdata.es.net/performance-testing/perfsonar/ps-howto/perfsonar-firewall-requirements/ for your perfSONAR-PS Toolkit instances

For sites willing to participate in more extensive monitoring we have detailed instructions at WLCGperfSONARMonitoring.
Registering Your perfSONAR-PS Instances

Once sites have completed the installation and configuration steps above, we need them to register their instances.

If your site is an OSG site, please follow the directions at http://www.opensciencegrid.org/bin/view/Documentation/RegisterPSinOIM

For non-OSG sites, please register with GOCDB by following the directions at PerfSONARInGOCDB

perfSONAR-PS Maintenance and Troubleshooting

Jason Zurawski/ESnet has provided a PDF file which documents some basic maintenance, troubleshooting and repair steps to address some issues in perfSONAR-PS. Have a look at 20120204-USATLAS-pSPT.pdf. It should be noted that many of the maintenance items have been incorporated into the 3.3 release and no longer need to be done manually. NOTE: All WLCG sites need to make sure they have provided a sufficient number of ports for testing...see section 6 in the PDF file. This was updated in June 2013 to recommend at least 1200 ports for OWAMP testing compared to the original 200. See http://psps.perfsonar.net/toolkit/FAQs.html#Q6

We have seen some issues with the config_daemon not running. If you see something like this:

2013/07/10 17:24:08 (1453) ERROR> ConfigClient.pm:89 perfSONAR_PS::NPToolkit::ConfigManager::ConfigClient::saveFile - ... RPC::XML::Client::send_request: HTTP server error: Can't connect to localhost:9000 (connect: Connection refused)

You should restart the config_daemon:

# /etc/init.d/config_daemon start

You should then verify something is listening on port 9000:

# netstat -apn | grep 9000

Note we are trying to maintain a list of tips, maintenance items and troubleshooting at http://www.usatlas.bnl.gov/twiki/bin/view/Projects/LHCperfSONAR so please check there for new items. In addition if you find other items we should add to this Wiki, please email Shawn McKee <smckee@umich.edu> and Simone Campana <simone.campana@cern.ch>.

Modular Dashboard for perfSONAR

The original modular dashboard work is now on hold. A new version operated by OSG is under development. The original dashboard development effort is still hosted in GitHub at https://github.com/PerfModDash Please contact the WLCG perfSONAR Deployment task-force if you are interested in working on modular dashboard development.

The new version targeting deployment in OSG for use by WLCG is based upon MadDash by Andy Lake of ESnet. You can see ESnet’s deployment at http://ps-dashboard.es.net/ See MadDashWLCG for details on deployment and testing.

Comments and Suggestions

Please send along any comments or suggestions about this information and planning via email to Shawn McKee <smckee@umich.edu> and Simone Campana <simone.campana@cern.ch>.
Background Documentation and perfSonar presentations

Experiments Data Access Evolution

- Daniele Bonacorsi ATLAS and CMS Data Management tools and Federated Data Store Implementations
- Daniele Bonacorsi LHCONE ATLAS/CMS point of view
- Costin Grigoras ALICE Data Access Model

LHC OPN and LHC ONE

- Artur Barczyk LHC Open Network Environment LHCONE
- Marco Marletta Stefano Zani LHC OPN & LHC ONE

PerfSonar presentations

- Shawn McKee LHCONE Monitoring: Where do we go from here?
- Shawn McKee GDB 12th September 2012
- Shawn McKee GDB 18th April 2012

-- ShawnMcKee - 15-Jan-2013