Minutes of LCG France LHCONE PerfSONAR-MDM workshop

Paris 4th April 2013

https://indico.in2p3.fr/conferenceDisplay.py?confId=8240

Attendees:

Roland Karch (DANTE), Domenico Vicinanza (DANTE), Xavier Jeannin (RENATER), Frederique Chollet (IN2P3 LAPP), FERRY Sophie (CEA IRFU), SCHAER Frederic (CEA IRFU), Pascale hennion (IN2P3, LLR), Philippe SERAPHIN (IN2P3 LAPP), Edith Knoops (IN2P3 CPPM), Julien Tugler (IN2P3, LLR), Denis Pugnere (IN2P3, INPL), Jean-Michel Barbet (IN2P3, Subatech), Laurent Caillat-Vallet (CC-IN2P3), Alessandra Forti (Univ. of Manchester)

Acronyms:
PS-PS = perfSONAR-PS
PS-MDM= perfSONAR-MDM
WG = Working Group
TF = Task Force
MP = Measurement Point
MA = Measurement Archive
OWD = One-Way Delay
OWAMP = One-Way Active Measurement Protocol (IETF std RFC4656)
HADES = Hades Active Delay Evaluation System
BWCTL = Bandwidth test Controller, a tool dev. by Internet 2 as wrapper around Iperf

1. **RENATER welcome**   Xavier Jeannin (RENATER)

XJ explained that this meeting gives the opportunity to GEANT and PS-MDM to meet the site administrators of several French LCG sites as the following sites LAPP, IPNL, CC-IN2P3 (Tier 1), LLR, CPPM, CEA-IRFU, Subatech are represented at this meeting.
Objective of RENATER project regarding PerfSONAR-MDM is to test of PS-MDM in operational condition, within LHCONE. Looking at the RENATER action plan, lot actions were already complete especially the deployment of PS-MDM within RENATER LHCONE backbone in Paris, (BWCTL, OWAMP MP).
RENATER supports the deployment of PS-MDM, but PS-MDM must be seen as useful and approved by end users.

2. **Overall Tour of PerfSONAR-MDM** Domenico Vicinanza (DANTE - PerfSONAR-MDM team)

DV gave a tour of all available feature of PS-MDM. There are 2 deployment domains that are separate: GEANT domain and LHCONE domain.
OWAMP module (OWD, jitter, packet loss) can be used only on demand; HADES modules (OWD, jitter, packet loss, traceroute) could provide scheduled test and store (historical data centrally stored) it but is not deployed within LHCONE.

. Final release of PS-MDM will implement OWAMP global scheduling capability. Bandwidth Measurement is provided by the BWCTL MP.

Archive models differ between PS-PS and PS-MDM. PS- PS has a distributed approach; each PS-PS has an archive. PS-MDM has a centralized approach: central archive being needed for scheduled tests.

XJ: A site that will install PS-MDM should benefit of the same services than one that will deploy PS-PS.
Unfortunately, even with LHCONE PS-MDM, user can see the 2 domains but can only access to LHCONE services.
Test on demand is useful but is used rarely, for experiment like ATLAS is important to have regular scheduled tests in order to add network monitoring in general monitoring process.
RK: SQL MA component is missing, so OWAMP test results cannot be stored into a repository. A new version will be available soon (few days).
PS-MDM can collect SNMP data (RRD-MA) to monitor link utilization: input errors, output errors.
Thanks to traceroute available in PerfSONAR-UI, you can check all the data along the path.

3. **Follow-up of the installation of a first instance** Frederic Schaer (IRFU)

This presentation is the first site report from the point of view of a site administrator. Installation works fine. BWCTL MDM and PS can interoperate. No HADES archive is available in LHCONE. Apart RENATER, GEANT, CEA-IRFU and DFN, nobody joined the LHCONE. No documentation is available for OWAMP package (For Debian, you have to use tar ball installation for OWAMP, not user friendly). Upgrade of package leads to change in configuration file. No lookup services available on LHCONE. OWAMP MP does not seem ready for production. Web UI is quite good. We need a One Way delay measurement, scheduling and storage, HADES server would be useful.
There is a need for global scheduling and global archive to go further with PS-MDM. DANTE is ready to help.
PH: What is the scalability of a HADES central server?
RK: 50 sites.

4. **PS deployment : overview of existing instances in France**

Network path to/from the storage (standard latency, bandwidth, traceroute tests) has to be monitored carefully => diagnosis in case of limited performance, degradation.
PerfSONAR-PS was deployed in LHCONE and LHCOPN and is working. PerfSONAR-MDM was deployed in LCHOPN but it is not used anymore.
RK: Now, one PS-MDM machine with 2 Eth interfaces can manage OWD and BWCTL tests.

5. **Overal picture of LCG context, needs and deployment strategy** Alessandra Forti (University of Manchester)

Data access is carefully managed by the experiments management. The WLCG Ops&Tools TEG (Technical Evolution Groups) provides recommendations regarding network monitoring and there is WLCG perfSONAR Task Force.

Recommendations: WLCG Network Monitoring: deploy a WLCG-wide and experiment independent monitoring system for network connectivity

- It is suggested that the PerfSONAR network monitoring system is deployed at all WLCG sites (two boxes, one for throughput and one for latency tests).

No recommendation was made by the TEG regarding the flavor of PS (PS-PS or PS-MDM) but PS-PS is widely deployed and pushed by ATLAS.

For Atlas experiment, originally star topology with a hierarchical structure:

- Evolved towards a flatter data distribution with the introduction of T2D. T2 which can distributed data to other T2

We don't want to have a network monitoring system per VO! There is a will to unify the network monitoring tool. PerfSONAR will be chosen as network monitoring tool for ATLAS, CMS, LHcB. ALICE use MonAlisa to collect network monitoring data (network topology is used to find the closest storage) and could swap to perfSONAR but this is not a priority..

Future already talking about network on demand application for both CMS and Atlas

- Network needs to be monitored
- Applications need to be instrumented

Twiki of this WG (TEG): https://twiki.cern.ch/twiki/bin/view/LCG/WLCGTegOperationsWG1

Recommendations of the WG (TEG) :

https://twiki.cern.ch/twiki/pub/LCG/WLCGTEGOperations/Ops_TEG_finalreport-05042012.docx

WLCG PS deployment Task Force: https://twiki.cern.ch/twiki/bin/view/LCG/PerfsonarDeployment

LHC-FR in the modular dashboard developed by BNL :

http://perfsonar.racf.bnl.gov:8080/exda/?page=25&cloudName=LHC-FR

XJ: Does PS-MDM team plan to join this working group?

DV: There is a room to set-up an active collaboration with this TF.

6. **Interoperability with existing PS-PS instances** Roland Karch (PerfSONAR-team)

Interoperability with PS-PS is possible for bandwidth test can be launch toward Esnet and Internet 2 running PS-PS. One way delay measurement interoperates as it achieved thanks to the same software OWAMP. Schedule tests, store tests and historic access to those tests do not interoperate between PS-MDM and PS-PS.

PS-PS and PS-MDM use the same protocol and same tools but PS-MDM deploys also measurement point.

Interoperability at data exchange between PS-PS and PS-MDM

- You access to data in PS-PS domain at the end point or through the PS-PS dashboard. Only the PS-PS dashboard can retrieve information from PS-PS end point so SQL MA or Hades MA could request to the PS-PS dashboard in order to retrieve the PS-PS data. In the opposite way the dashboard could make request the data to SQL MA or Hades MA.

For the moment, there is no possibility to exchange data between PS-MDM and PS-PS.

7. **Discussion on the requirement of LHC community, Network Monitoring with PS**

- Reference state:

The idea is to have reference state for the number of HOP, OWD average and others network metrics so when an important change happens, we can set-off an alarm.

RK - Better solution is to use Nagios plugin rather than PS-MDM

- Authorization

There is no authorization control in PS-MDM, but we think that PS-PS has no authorization either. The problem is still pending.

LHCONE is a VPN so there is already a security achieved.

- What are the showstoppers for the deployment
  - LHCONE requires a OWD delay services
    - OWAMP is available no archive possible since SQL MA is not been released
      - DANTE will talk with the developer to know the date of the next release
    - no scheduling test available → require an HADES server. DANTE will do the necessary to make this possible within LHCONE PS-MDM setup.

- **Next step** :

  - Demonstration of the following points:
    - Central configuration of the test
      - Require a SSH access and restart daemon by a specific user
      - Is acceptable by the site regarding their security policy?
    - Schedule test
      - Same question concerning ssh access
    - OWD service
    - Store the result of the scheduled tests → an archive should be set-up
    - Register CEA nodes into LHCONE PS-PS (lookup service)
      - Interoperability for BWCTL (already done) and OWD

DANTE: trying to address at least some of the above points could be achieved for end of May. There is an opportunity to report on progress made at the next LHG-France spring 2013 meeting, May 28-30 : See https://indico.in2p3.fr/conferenceDisplay.py?confId=8140