Intercontinental Multi-Domain Monitoring for LHC with perfSONAR

D Vicinanza¹

Product Manager, DANTE, City House, 126-130 Hills Road, Cambridge CB2 1PQ, UK

E-mail: domenico.vicinanza@dante.net

Abstract. The Large Hadron Collider (LHC) is currently running at CERN in Geneva, Switzerland. Physicists are using LHC to recreate the conditions just after the Big Bang, by colliding two beams of particles and heavy ions head-on at very high energy. The project is generating more than 15 TB of raw data per year, plus 10 TB of "event summary data". This data is sent out from CERN to eleven Tier 1 research centres in Europe, Asia, and North America using a multi-gigabits Optical Private Network (OPN), the LHCOPN. Tier 1 sites are then connected to 100+ academic and research institutions in the world (the Tier 2s) through a Multipoint to Multipoint network, the LHC Open Network Environment (LHCONE). Network monitoring on such complex network architecture to ensure robust and reliable operation is of crucial importance. The chosen approach for monitoring the OPN and ONE is based on the perfSONAR framework, which is designed for multi-domain monitoring environments. perfSONAR (www.perfsonar.net) is an infrastructure for performance monitoring data exchange between networks, making it easier to solve performance problems occurring between network measurement points interconnected through several network domains.

1. Introduction

CERN (European Organization for Nuclear Research) operates the Large Hadron Collider (LHC) in Geneva, Switzerland. LHC is a particle accelerator used by physicists to study the smallest known particles – the fundamental building blocks of matter. Physicists will use the LHC to recreate the conditions just after the Big Bang, by colliding two beams of subatomic particles head-on at very high energy. Teams of physicists from around the world will analyse the particles created in the collisions using sophisticated detectors in 4 major experiments. These experiments will generate large amounts of data, which will be sent to data processing and storage centres around the world. More than 15 petabytes of data (equivalent to over 1.7 million dual-layer DVDs) is expected to be generated on a yearly basis [1].

Given the requirements in terms of amount of data generated to support this experimental challenge, firstly an Optical Private Network (OPN) and then an Open Network Environment (ONE) were designed for LHC by a collaboration of organisations including CERN, the network providers and the research community [2].

The OPN network is structured on different levels, known as Tiers. Each Tier is responsible for generating and/or processing data and returning the results to the upper Tier, eventually arriving at

¹ To whom any correspondence should be addressed.

International Conference on Computing in High Energy and Nuclear Physics 2012 (CHEP2012) IOP Publishing Journal of Physics: Conference Series **396** (2012) 042060 doi:10.1088/1742-6596/396/4/042060

CERN (Tier 0). The Tier 0 at CERN and other 11 Tier 1 centres, spread around Europe, North America and Asia, are connected by the OPN [2]. Subsequent Tiers, mainly composed of universities and/or research centres, are active part of the data processing and/or result analysis of the experiments and they are in the process of being connected through the ONE [3], a L3VPN-based multipoint to multipoint network architecture.

The reliability of the different layers of this complex network structure is of a crucial important for the robustness and reliability of the operations. In particular the monitoring requirements were:

- Scalability
- Interoperability at international level
- Easy to deploy and manage
- Easy to use for network engineers

The solution chosen is based on perfSONAR for both LHCOPN and LHCONE. perfSONAR [3] (Performance focused Service Oriented Network monitoring ARchitecture) is a monitoring framework designed for multi-domain monitoring environments. Its development is based on an international collaboration for network monitoring and its contributors are GÉANT, Internet2, ESnet, and RNP. There are currently two main implementations committed to interoperate:

- perfSONAR MDM within GÉANT: http://perfsonar.geant.net
- perfSONAR PS within I2/ESnet: http://psps.perfsonar.net/

They both use an open OGF protocol to exchange data, they are web-service based and they share the same design goals:

- Flexibility
- Extensibility
- Openness
- Decentralization.

Their differences are mainly in the software development process, product life cycles, interaction with the users, implementation and deployment models.

2. Dedicated networks for the Large Hadron Collider

The Large Hadron Collider (LHC) data distribution model is based on the same multi-tiered structure, with a main data source, called Tier 0 (T0), some first-level processing and storage sites, called Tier 1 (T1) and many second-level analysing sites (mainly universities and smaller research centres), called Tier 2 (T2). Each site is expected to handle and store about one Petabyte of raw data per month. There are 120+ T2 sites, mainly involved in physics analysis [7]. There is a continuous exchange of high volumes of data between various T0 and T1, between T1 and T2 and among T2.



Figure 2: LHCOPN aggregated traffic (source: cern.ch)

International Conference on Computing in High Energy and Nuclear Physics 2012 (CHEP2012) IOP Publishing Journal of Physics: Conference Series **396** (2012) 042060 doi:10.1088/1742-6596/396/4/042060

Since the wide area data rates into and out of the T1 sites is of the order of many Gigabits per second (Gbps) [2], as visible in Figure 2, this traffic is segregated from the general Internet whenever possible, using an Optical Private Network (LHCOPN). The LHCOPN is a dedicated optical network designed to connect CERN, as T0, with all eleven T1 The high-level network design is shown in Figure 2.



Figure 2: The LHC-OPN network diagram

The twelve T0/1s are listed in Table1:

Table	1.	List	of	Tier	0/1s.
-------	----	------	----	------	-------

T1 name	T1 location
CA-TRIUMF	Vancouver, BC, Canada
CH-CERN	Geneva, Switzerland
DE-KIT	Karlsruhe, Germany
ES-PIC	Barcelona, Spain
FR-CCIN2P3	Lyon, France
IT-CNAF	Bologna, Italy
NDGF	Stockholm, Sweden
NL-T1	Amsterdam, The Netherlands
TW-ASGC	Taipei, Taiwan
UK-T1-RAL	Didcot, United Kingdom
US-FNAL	Batavia, IL, USA
US-T1-BNL	Upton, NY, USA

The LHC-OPN's agreed networking strategy is to use at least one dedicated 10 Gigabit/seconds light path between the T0 and each T1. This 10 Gigabit/seconds light path terminates at a networking

International Conference on Computing in High Energy and Nuclear Physics 2012 (CHEP2012) IOP Publishing Journal of Physics: Conference Series **396** (2012) 042060 doi:10.1088/1742-6596/396/4/042060

equipment at the T0, and inside or as close as possible to the T1. IPv4 has been chosen as the network protocol to provide communications among the upper layer applications during the first stage; other network protocols, such as IPv6, may be considered in the future. Each Tier is encouraged to support a Maximum Transmission Unit (MTU) size of at least 9000 bytes on the entire path between T0 and T1. From an architectural perspective, the T0-T1 links should handle only production LHC data (LHC Network Traffic). The routing among the T0 and T1s sites will be achieved using BGP (Border Gateway Protocol). T1-T1 traffic via the T0 will also be allowed, although T1s are encouraged to provision adequate alternative connectivity, for example using direct light path, tunnel over IP/MPLS networks, transit over IP backbones, or any other suitable technology. Some of the Tier 1s are interconnected using cross-border fibres, to provide a reliable backup as well as direct T1–T1 connectivity.

Most of the analysis will be done in Tier2 research centres and labs, and reliable networking started to be needed to connect Tier2s among themselves and provide the right Tier2 to Tier1 connectivity. Having that in mind, since beginning 2011 a new network started to be designed. From the new requirements

(http://indico.cern.ch/getFile.py/access?contribId=5&resId=4&materialId=0&confId=102716): "Data should be able to be pulled [...] from any Tier-2 or even, if possible, from several Tier-2s simultaneously in order to minimize the time to have the data locally to run the analysis task."

The answer was then the LHC Open Network Environment [2]; an L3VPN based multipoint network architecture whose design and implementation is still on-going at the time of writing.

3. perfSONAR

3.1. The perfSONAR Framework

Performance Service-Oriented Network Monitoring Architecture (perfSONAR) is a framework developed by a consortium of organisations (GÉANT, ESnet, Internet 2 and RNP) that enables network performance information to be gathered and exchanged in a multi-domain, federated environment [4]. The primary goals of perfSONAR are to enable ubiquitous gathering and sharing of performance information, simplify management of advanced networks, facilitate cross-domain troubleshooting and allow next-generation applications to tailor their execution to the state of the network (as depicted in Figure 2).

From the user's point of view, perfSONAR can be considered as a set of distributed services, which offer a range of functionalities required by any well-defined and complete multi-domain monitoring system. Examples of functionalities implemented are: storage of network information and measurement results, passive tests, active tests, authentication, authorisation, directory service and visualisation. The distributed services can run in multiple domains and communicate using a single, standard protocol, ensuring that information is shared effectively and securely.



Figure 1: The perfSONAR objective: cross-domain Network Information Availability

One of main advantages of the perfSONAR-based approach is that the network operators do not have to abandon their existing monitoring tools. Instead, perfSONAR can be considered as an extension that provides new features as well as uniform and standardised access to already deployed tools. The perfSONAR framework is built as a set of network performance middleware and visualisation tools, which most a number of low requirements [6]:

tools, which meet a number of key requirements [6]:

- Standardised access to multi-domain network information
- Increased efficiency through on-demand troubleshooting tools
- Easy identification of information sources
- Network performance validation
- Optimised network usage through regular TCP (Transmission Control Protocol) throughput tests
- Faster problem identification through consistent end-to-end service monitoring
- Monitoring of virtual networks
- Secure access.

3.2. perfSONAR Main Components

The modular design and open-source nature of the perfSONAR services (and related code) allows domain administrators to implement, combine and customise tools according to their individual requirements. Domains can thus create *ad-hoc* distributed network monitoring systems with maximum flexibility. perfSONAR integrates all measurement tools by providing common services and visualisation tools. The complex set of perfSONAR services supports access to measurement data and network information across multiple administrative domains. Each individual service is in charge of a specific functionality. The following types of perfSONAR components are available:

3.2.1. Measurement Points

Measurement Points (MPs) collect and publish data obtained (usually locally) using measurement tools. The measurement is carried out either on client demand or automatically at scheduled time intervals. The MP also provides an interface between the perfSONAR data format and the application proprietary data structure.

3.2.2. Measurement Archives

Measurement Archives (MAs) store measurement results in databases (for example SQL) or in file systems (for example RRD files). This data can later be read by clients and further processed or visualised. MAs can also support data aggregation.

3.2.3. Lookup Services

Each network domain that joins the perfSONAR infrastructure should have a Lookup Service (LS). All services in this domain can then register with the Lookup Service. One of the interesting features of this system is that Lookup Services of multiple network domains can communicate with each other to share information about the registered services. In this way any interested user only needs to know the URL of one Lookup Service to be able to find out which other services are available in local and external domains.

3.2.4. Visualisation Tools

The perfSONAR MDM visualisation tools are user interfaces for querying perfSONAR services and presenting the collected data to the user in front-end applications (e.g. perfSONAR UI). The opensource nature of the visualisation tools means that the presentation of performance data can be adapted to the needs of specific user groups.

3.3. GÉANT perfSONAR MDM (Multi Domain Monitoring Service)

There are currently two main implementations used by the LHCONE and LHCOPN communities:

- perfSONAR MDM within GÉANT: http://perfsonar.geant.net
- perfSONAR PS within I2/ESnet: http://psps.perfsonar.net/

LHC sites are typically free to choose the perfSONAR flavour they prefer, according to factors like geographical location, availability of local support, experience and expertise.

perfSONAR MDM is designed to provide a service, with federated deployment, centrally monitored and coordinated, and full support. perfSONAR PS has a distributed support model with the goal of proliferating the number of performance nodes deployed across the community.

The two deployments are committed to interoperate, accepting each-other measurement requests, displaying results and storing metrics into archives available for all the sites part of the LHC computing project. This section will analyse in detail the GÉANT perfSONAR MDM structure and measurement strategy.

3.3.1. GÉANT MDM Service Strategy

The GÉANT Multi-Domain Monitoring (MDM) Service [5] offers a unified view of the network status and information across all sites. The GÉANT MDM infrastructure is usually deployed close to the border routers (administrative boundaries); this sets clear demarcation points and creates a distinction between issues affecting the managed network and issues affecting the sites' network.

The GÉANT MDM service uses perfSONAR as it base service. The main objective of the service is to correctly, efficiently and quickly identify network problems occurring on any part of a network operated in a multi-domain environment. The service aims to provide fast, reliable and uninterrupted network communication, tracking issues across multiple domains [5].

The strategy adopted is to perform network monitoring activities in different network domains, making the information available using a common protocol and cross-domain monitoring. This common infrastructure enables access to network performance metrics from across multiple domains, enabling network problems and performance bottlenecks to be traced and quickly eliminated. Finally, whenever possible, the GÉANT MDM service can be used to proactively identify and prevent problems before service disruption occurs.

3.4. perfSONAR Tools

perfSONAR MDM and PS currently include a set of tools that cover the most important areas of network monitoring and measurements at the IP layer or above. In particular, it includes throughput, delay and passive measurement tools, which are described in the following sections.

International Conference on Computing in High Energy and Nuclear Physics 2012 (CHEP2012) IOP Publishing Journal of Physics: Conference Series **396** (2012) 042060 doi:10.1088/1742-6596/396/4/042060

3.4.1. Throughput Measurement Tools

Throughput measurement is performed using the perfSONAR BWCTL MP, a measurement point based on the Internet2's Bandwidth Controller (BWCTL) software. The client can specify measurement and data flow parameters for the actual bandwidth determination (such as the window size for TCP or the socket buffer for UDP) and the protocol (IPv4 or IPv6), allowing exhaustive checks and effective troubleshooting for different network end-point setups.

3.4.2. Delay Measurement Tools

Delay measurements are based on the Hades Active Delay Evaluation System (HADES) and the One-Way Active Measurement Protocol (OWAMP) infrastructures. HADES devices were initially developed by WiN-Labor at RRZE (Regional Computing Centre Erlangen). They provide Quality of Service (QoS) measurements in DFN's G-WiN infrastructure, based on the metrics developed by the IETF IPPM WG and later extended to numerous GÉANT Points of Presence (PoPs). This system enables the network operators and systems administrators to obtain parameters derived from a timestamped flow of probing packets. This includes one-way delay [RFC2679], delay jitter [RFC3393], packet loss rate [RFC2680] and alternative path routes

OWAMP (http://www.internet2.edu/performance/owamp/) is a command line client application and a policy daemon used to determine one way latencies between hosts. It is an implementation of the OWAMP protocol as defined by http://www.rfc-editor.org/rfc/rfc4656.txt. OWAMP session control uses traditional client-server communication between a control-client and a server, using the OWAMP-Control protocol. The owampd server is implemented using the conventional accept/fork model. The two sides negotiate one-way test parameters using the OWAMP-Control protocol. This OWAMP implementation then forks additional processes to actually send and receive the OWAMP-Test formatted packets that implement the session endpoints for the Sender and Receiver roles.

3.4.3. Passive Measurement Tools

perfSONAR tools based on passive monitoring have recently been developed (e.g. PacketLoss) or are under development (e.g. ABW2). Many network parameters and characteristics can only be obtained using passive measurement, which does not influence the network traffic. This includes classification of flows, link load measurements and exact packet loss of real data flows.

3.4.4. GÉANT MDM Service Operations

Once a GÉANT MDM service is deployed, the service is in operational status and is monitored using the operational components and processes described below, to monitor service availability.

GÉANT MDM service applications are composed of three layers: a service layer, an underlying layer and a hardware layer. The service layer is represented by the software implementing the application interface (typically a web service API) and support software for this layer. The second layer is the resource layer, containing software components such as daemons, measurement tool clients (Net-SNMP client, as an example), data files and/or databases and configuration files. The third layer is related to hardware resources, such as CPU, Load Average, Memory, Network Interfaces and Partitions. This service structure is illustrated in Figure 4.



Figure 4: GÉANT MDM Service General Application Structure

The monitoring of the health of the overall application reflects the status of the components in the three layers. Effective monitoring, performed at different hierarchical levels, allows operators to correctly identify problems, proactively notify parties responsible and take remedial actions.

Service Layer

The service layer is the application layer directly accessible by the end user. It is typically implemented through a web service, so monitoring is focused on checking the availability and performance of this interface.

At this layer, using only some web-service messages on the application side allows the verification of web-service availability and its performance (e.g. loss, delays and uptime). Such an implementation also enables the verification of service downtime, which could trigger any service recovery procedure.

Resource Layer

Monitoring availability and performance at the service layer, as described in the previous section, is not sufficient to ensure that a web service correctly delivers the required functionality to users. In general terms, the resource layer grants the upper layer access to the resources it needs to correctly provide a service to the user. Any problems relating to database access, Round Robin Data (RRD) file existence or measurement tool daemons functionality is monitored at this level. This layer is monitored using an internal special message, sent to the service, which verifies the health of the components used by the service.

Hardware Layer

Hardware layer monitoring has a primary role in the correct diagnosis of problems. Correct monitoring of the hardware availability and performance enables a pro-active response, preventing service disruption to the user. It also permits an understanding of the causes of issues detected at higher levels, such as problems reported by the resource and/or service layers.

Hardware resources are monitored using existing Simple Network Management Protocol (SNMP) [RFC 1157] infrastructure. The SNMP agent installed on each server is queried on a regular basis by a central SNMP manager to gather relevant information from the server. SNMP was chosen because it is a common standard for monitoring and, once running as a separate application on the server, it is not influenced by specific service failures. Also, SNMP agents already have hardware variables present.

3.4.5. Monitoring the perfSONAR service: the GÉANT Monitoring Infrastructure

This infrastructure consists of a fault and capacity monitoring tool, which enables service engineers to pro-actively deal with events, before these become incidents which interrupt the service.

Nagios [8] was chosen as the event management tool, since it includes features such as host and service monitoring, and alerting on different sets of events based on their criticality. It enables the support team to react to the current status of the infrastructure, provides reports of the performance of the service monitored for trend analysis, allows the flagging of scheduled maintenances and, finally, is a good dashboard tool that enables the supporting team to plan any kind of upgrade and/or modification on the monitored infrastructure. Nagios is a flexible and customizable framework, which allows monitoring extensions to be created for any type of service. A screenshot of Nagios is shown in Figure 5.

<u>N</u> agios'	Service Status Details For Host 'DE-KIT- DE-KARLSRUHE'								
General			a		- · · •	•··· • •			
Home	Host TV	Service TV	Status TN	Last Check TV	Duration TV	Attempt TN	Status Information		
Documentation	DE-KIT- DE-KARLSRUHE	PING	ок	03-04-2010 19:13:27	1d 19h 10m 16s	1/4	OK - mdm-de-kit.gridka.de: rta 39.320ms, lost 0%		
Monitoring		Torncat response test	ок	03-04-2010 19:13:08	15d 2h 40m 4s	1/4	mdm-de-kit.gridka.de Torncat OK - Median elapsed time: .080		
Tactical Querview		checking CPU load v3	ок	03-04-2010 19:14:16	3d 12h 18m 22s	1/4	CPU used 3.0% (<10) : OK		
Service Detail Host Detail		checking DE-KIT-DE- KARLSRUHE cacti	ок	03-04-2010 19:13:26	1d 7h 7m 11s	1/4	OK - mdm-de-kit.gridka.de: Time: 2010-03-04 19:10:02 CACTI Polling Time: 03/04/2010 07:10:02 PM CACTI Poller Performance: 1.1389		
Hostgroup Overview Hostgroup Summary		checking loginattempts	ок	03-04-2010 19:16:26	1d 13h 4m 31s	1./4	0K - mdm-de-kit.gridka.de - Intrusion detection check (from 03/04/2010 18:00 to 03/04/2010 19:00) - Successful login attempts: 0 - Failed login attempts: 0		
Hostgroup Grid		checking management interface	оĸ	03-04-2010 19:14:24	9d 11h 33m 49s	164	OK - 192 108 46 109: rta 12 106ms. Inst 0%		
Servicegroup Overview		checking memory v3	ok	03.04.2010 19:13:27	14d 6h 49m 15e	1.4	Rem : 33% Swen : 8% : : 0K		
Servicegroup Summary		okooking mund ning	OK OK	03-04-2010 13:13:27	04 6h 62m 27o	1.74	OK MuSOL is LID and suppling. Automatic grant participances is 0.520 accords		
Status Map		checking mysgi ping	ok	03-04-2010 13.17.17	00 0H 33H 37S	4.04	We also up 4 up ov		
3-D Status Map		checking network internace etric vs	ok	03-04-2010 19.17.49	700 15ri 15ri 35s	124	ethologi, i de, ok		
Service Problems		checking network interface ethil v3	UK	03-04-2010 19:13:15	15d 2h 39m 11s	1.4	eth1:UP:1 UP: UR OK: synchropised to NTP server (192:108:46:124) at stratum 2 time correct to within 19		
© Unhandled		checking ntp_client	ок	03-04-2010 19:13:46	1d 8h 16m 45s	1/4	ms polling server every 1024 s		
Outhandled		checking raid client	ок	03-04-2010 19:12:57	0d 6h 31m 15s	1/4	OK: RAID is running properly - log_id 2 OPTIMAL phys_id 1 ONLINE phys_id 0 ONLINE		
Network Outages		checking sensors client	ок	03-04-2010 19:17:12	1d 20h 14m 23s	1.4	OK: All temperature, voltage and fan sensors for all components are working fine.		
Show Host:		checking storage / v3	ок	03-04-2010 19:14:25	78d 15h 10m 58s	1/4	/: 27%used(530MB/1938MB) (<80%): OK		
Chainmost		checking storage Aboot v3	ок	03-04-2010 19:14:16	7d 4h 54m 57s	1/4	Aboot: 13%used(24MB/190MB) (<80%): OK		
		checking storage /home_v3	ок	03-04-2010 19:15:34	15d 2h 36m 44s	1/4	/home: 1%used(172MB/29045MB) (<80%) : OK		
ê		checking storage /opt v3	OK	03-04-2010 19:13:36	15d 2h 38m 51s	1/4	/opt: 1%used(219MB/36887MB) (<80%) : OK		
Comments		checking storage Amp v3	ок	03-04-2010 19:15:35	82d 15h 14m 0s	1.4	/mp: 3%used(138MB/4836MB) (<80%): OK		
- Downcine		checking storage Ausr v3	ок	03-04-2010 19:14:45	7d 4h 54m 57s	1./4	/usr: 8%used(2200MB/29045MB) (<80%): OK		
Process Info		checking storage /var v3	ок	03-04-2010 19:14:25	15d 2h 36m Os	1./4	/var: 2%used(720MB/29045MB) (<80%) : OK		
Performance Info		checking top connections	ок	03-04-2010 19:14:07	3d 22h 46m 39s	1/4	TCP connection attemps check OK		
Scheduling Queue		checking top currestab	ок	03-04-2010 19:17:57	1d 8h 18m 38s	1/4	OK - TCP current established connections: 3		
Reporting		checking top segments	ок	03-04-2010 19:13:49	1d 0h 52m 4s	1/4	TCP segments check OK		
Trends		checking udp datagrams	ок	03-04-2010 19:14:37	5d 20h 19m 53s	1/4	UDP datagram check OK		
Availability		exist reachable	ок	03-04-2010 19:14:24	15d 2h 38m 31s	1/4	HTTP OK HTTP/1.1 200 OK - 12341 bytes in 0.242 seconds		
Alert Histogram		perfSONAR AS echo	ок	03-04-2010 19:13:15	16d Oh 43m 49s	1/4	mdm-de-kit.oridka.de AS OK - Elapsed time: .0893 - Time corrected for http delay:0217		
 Alert History Alert Summary Notifications Event Log 		pertSONAR AS selfcheck	ок	03-04-2010 18:52:54	15d 1h 43m 47s	1,4	OK - 2 self-tests: 2 successful, 0 failed (as_validsfile_test ::: success - Valids compone file /usrlinibperfsonar/services/geart2/av-as/MEB-NF/classes/perfsonar /conf/ASValdComponentsFile ready: as_truststore_test :: success - Truststore file /usrlinbjerfsonar/services/geart2/ava-as/MEB-NF/classes/perfsonar		
Configuration			01/	00.01.0010.1010.10	454.05.0000-	4.04	/conf/perfSONARtruststore.jks' ready).		
View Config		pertSUNAR LS echo	UK	03-04-2010 19:13:15	150 Zh 39m 20s	1.4	mam-de-kit.gridka.de LS OK - clapsed time: .uo12 - Time corrected for http delay: .uuo2		
		perfSONAR_LS_selfcheck	ок	03-04-2010 18:58:04	113d 17h 33m 43s	1/4	OK - 3 self-4ests: 3 successful, U tailed (is selftest collection isstore :: success - collection LSStore exists and is accessible; is selftest disconnectivity :: success - eXist DB connection established; is selftest collection isstore-control :: success - collection LSStore-control exists and is accessible).		
		perfSONAR_RRD-MA_echo	ок	03-04-2010 19:13:10	15d 2h 39m 13s	1/4	mdm-de-kit.gridka.de RRD-MA OK - Elapsed time: .0874 - Time corrected for http delay: .0254		
		perfSONAR RRD-MA selfcheck	ок	03-04-2010 18:53:12	15d 1h 43m 13s	1/4	OK - 3 self-tests: 3 successful, 0 failed (xmidb-content-test :: success - Number of all metadata and data elements in the xmi database is 32; xmidb-access-test :: success - Ym Database urs 1, 3 ; is accessful, random facts hart :: access - Fatting data for the content of the state of the test of the state of the		

Figure 5: Event Management Infrastructure developed for LHCOPN in Nagios

Since Nagios only supports event management, Cacti [9] was selected for capacity management. Cacti offers a complete network graphing solution, has a fast poller, advanced graph templating, multiple data acquisition methods and standard user management features [10]. Cacti was integrated with Nagios via the Nagios plug-in for Cacti, enabling regular data imports, which provide a historical view of service performance metrics. A screenshot of Cacti is shown in Figure 6.



Figure 6: Capacity Management Infrastructure in Cacti

4. GÉANT visualization tools for LHCOPN and LHCONE

In the past two years GÉANT developed a series of visualization tools for the LHC project aiming at collecting data from both perfSONAR MDM and perfSONAR PS sites and show them in a coherent way. This section contains some screenshots taken from the user interfaces currently developed for the T0 and T1 sites.

4.1. perfSONAR web user interface

Since 2011 perfSONAR MDM has a new web interface [11] and a special version of this tool has been made available for the LHC community [12].

The interface offers a comprehensive access to the following metrics:

- Link utilisation, with possibility to compare two links
- One-way delay
- One-way delay variation
- Packet loss
- Hopcount (traceroute monitoring) with route comparison
- Regularly scheduled bandwidth measurement
- On-demand bandwidth measurement

from a single URL. The interface has been designed in collaboration with a panel of real users and it is currently used by GÉANT Network Operation Centre (NOC) engineers and gradually adopted by European NOCs. Its design has been professionally reviewed by a usability expert and it is customisable in terms of security and access rights. It is easy to manage and Debian and Redhat packages have been made available. A live virtual machine is also available for download from the perfSONAR MDM website [11].

Figures 7-10 show how different metrics can be successfully accessed and visualised coherently using the web interface.



Figure 7: Link utilisation on the perfSONAR MDM user interface



Figure 8: One-Way Delay, Jitter, Packet Loss and Traceroute on the perfSONAR MDM user interface



Figure 9: Accessing BWCTL Historic Measurements on the perfSONAR MDM user interface



Figure 10: Making an on-demand BWCTL measurement on the perfSONAR MDM user interface

4.2. perfSONAR MDM weathermap for LHCOPN

At the end of 2011 a weather map reading live perfSONAR MDM data was prepared for the LHCOPN community (and one is currently under design for the LHCONE). The map gives access to 19 different metrics, offering a colour-coding view according to the metric chosen. The available metrics include:

- OWD
- Jitter
- Packet Loss
- Bandwidth
- Link Utilisation / Input errors / Discarded packets

Figures 11 shows how it is possible to select one of metrics and Figure 12 shows the final result: a coloured, live, customised LHCOPN map.



Figure 11: Choosing metric on the weathermap



Figure 12: The weathermap coloured according the One-Way Delay

5. Conclusions

Supporting the monitoring of the LHCOPN and LHCONE infrastructure is an extremely challenging task, particularly exciting as the LHC went into its data taking era. The new web interfaces designed and customised for the LHC community are creating a big difference in the way monitoring information is gathered, visualised and correlated across three continents.

The authors of this paper would like to thank all the perfSONAR community for their support, including the GÉANT partners, who were involved on the service deployment for the LHC-OPN. We are grateful for the support of DANTE in providing the infrastructure to accomplish the work required and to the perfSONAR development team who provided the required tools to build up the service. Finally, thanks to CERN and its partners for deploying the service.

References

[1] CERN - The Large Hadron Collider. Website: http://public.web.cern.ch/public/en/LHC/.

[2] LHCOPN - Large Hadron Collider Optical Private Network. Website:

https://twiki.cern.ch/twiki/bin/view/LHCOPN/WebHome. LHCONE - Large Hadron Collider Open Network Environment Website: http://lhcone.net/

[3] perfSONAR - Performance focused Service Oriented Network monitoring Architecture. Website: http://www.perfsonar.net/.

[4] perfSONAR datasheet:

http://www.geant.net/Media_Centre/Media_Library/Media%20Library/PUB%2012%20053%20perfsonar%20datasheet%2005.12.pdf

[5] perfSONAR MDM Website: http://perfsonar.geant.net

[6] Dijkstra F., Fried A. *et al.*, GÉANT Deliverable DJ2.3.1: "Specification of Advanced Features for a Multi-Domain Monitoring Infrastructure"

[7] Tierney B., Metzger J., *et al*, "perfSONAR: Instantiating a Global Network Measurement Framework"

[8] Nagios. Website: http://www.nagios.org/.

[9] CACTI. Website: http://www.cacti.net/.

[10] Nagios Plugin for Cacti (NPC). Website: http://trac2.assembla.com/npc.

[11] perfSONAR MDM web interface: http://psui.geant.net

[12] perfSONAR MDM web interface for the LHC community: http://psui-lhc.grid.aau.dk/perfsonarui/