

HEPiX Fall/Autumn 2017 Workshop

Monday 16 October 2017 - Friday 20 October 2017

KEK

Book of Abstracts

Contents

A user portal at CCINP3	1
AGLT2 Site Report	1
ASGC Site Report	1
Australia site report	2
AutoGOLE bringing high speed data transfers	2
Automatic shutdown of servers in case of A/C failure	2
BNL RACF Site Report	3
CEPH at RAL	3
CERN Site Report	3
CSNS HPC Platform Based on SLURM	4
Cloud deployment at KEK	4
Cloud storage with the Dynafed data federator	4
Configuration automation for CERN's new Wi-Fi infrastructure	5
Continuous Integration for Linux Images	5
Current Status and Future Directions of KEK Computer Security	5
DESY site report	6
Deployment and monitoring for distributed computing sites	6
Deployment of IPv6-only CPU on WLCG - an update from the HEPiX IPv6 Working Group	6
EGI CSIRT: Keeping EGI Secure	7
Fancy Networking	7
Firewall Load-Balancing solution at CERN	8
First experience with SELinux	8
Follow-up about Wi-Fi service enhancement at CERN	8

GSI Site Report	9
HEP Community White Paper	9
HEPiX Benchmarking Working Group - Status Report October 2017	9
Hot days with no mechanical cooling data center	10
IHEP Site Report	10
INFN-T1 site report	10
Integrated Monitoring results at IHEP	11
Integrating HPC and HTC at BNL – a year later	11
KEK Site Report	11
KIT Site Report	12
KR-KISTI-GSDC-01 Tier-1 Site Reports	12
LHCOPN and LHCONE	12
Logistics information	13
Low-Power Wide-Area Network (LPWAN) at CERN	13
Managing cloudy dCache storage pools with Ansible.	13
Migrating a WLCG tier-2 to a Cray XC-50 at CSCS-LCG2	14
Migration from Grid Engine to HTCondor	14
Modernising CERN document conversion service.	15
NDGF Site Report	15
NetSage, a unified network measurement and visualization service	15
Netbench –testing network devices with real-life traffic patterns	16
Network Automation for Intrusion Detection System	17
Network Functions Virtualisation Working Group Proposal	17
Network related updates in IHEP	18
Next Generation Software Defined Services and the Global Research Platform	18
On Server Management Interface (BMC)	19
Optimising the resource needs for the LHC computing: ideas for a common approach	19
PDSF Site Report and Transition	20
PIC report	20
Particulate magnetic tape for data storage and future technologies	20

RAL Site Report	21
Recent network connectivity around KEK	21
Running HEP Payloads on Distributed Clouds	22
Running Jobs Everywhere: an Overview of Batch Services at CERN	22
Securing Elasticsearch for free: integration with SSO and Kerberos at CC-IN2P3	22
Security update	23
Security: case study	23
Singularity at the RACF/SDCC	24
Storage for Science at CERN	24
Swiss National Supercomputing Centre - T2 Site Report	24
Techlab update	25
The HEPiX workshop series	25
The HPSS User Forum (HUF)	25
The LHCONE network	26
The Outlook for Archival Storage at CERN	26
The TransPAC project	26
The birth of HEPiX - at KEK in 1991	27
Tier-1 networking cost and optimizations	27
Tokyo Tier-2 Site Report	27
US ATLAS SWT2 Site Report	28
University of Wisconsin-Madison CMS T2 site report	28
Updates from Database Services at CERN	28
Using Configuration Management to deploy and manage network services	29
Using Docker containers for scientific environments - on-premises and in the cloud	29
WLCG archival storage group	30
WLCG/OSG Networking Update	30
Welcome to KEK	31
Wigner Datacenter's new software defined datacenter architecture	31
Workshop wrap-up	31
printing@gsi	32

End-user services and operating systems / 22**A user portal at CCINP3****Author(s):** Renaud Vernet¹**Co-author(s):** Cyril L'Orphelin² ; Gino Marchetti³¹ *CC-IN2P3 - Centre de Calcul (FR)*² *CNRS/IN2P3*³ *CNRS***Corresponding Author(s):** gino.marchetti@cc.in2p3.fr, cyril.lorphelin@cc.in2p3.fr, renaud.vernet@cern.ch

CCIN2P3 is one of the largest academic data centres in France. Its main mission is to provide the particle, astroparticle and nuclear physics community with IT services, including large-scale compute and storage capacities. We are a partner for dozens of scientific experiments and hundreds of researchers that make a daily use of these resources.

It is essential for users to have at their disposal simple tools to monitor their activity and incidents, in order to use our services efficiently. Although this monitoring information exists today, it is not exposed to users in a fully centralized and easy-to-access way. Therefore, we have been developing since spring a web portal providing useful monitoring displays and pointers to service documentations.

With this presentation, we would like to show what we have implemented so far, share our thoughts on how to display the information to the users, and hopefully exchange ideas with the HEPiX community.

Site reports / 62**AGLT2 Site Report****Author(s):** Shawn Mc Kee¹**Co-author(s):** Philippe Laurens² ; Robert Ball¹¹ *University of Michigan (US)*² *Michigan State University (US)***Corresponding Author(s):** ball@umich.edu, shawn.mckee@cern.ch, philippe.laurens@cern.ch

We will present an update on the ATLAS Great Lakes Tier-2 (AGLT2) site since the Spring 2017 report including changes to our networking, storage and deployed middleware. This will include the status of our transition to CentOS/SL7 for both our servers and worker nodes, our upgrade of VMware from 5.5 to 6.5 and our upgrade of Lustre to 2.10.1 + ZFS 0.7.1 as well as our work to install Open vSwitch on our production dCache instances.

Desired length:

12 minutes

Site reports / 65**ASGC Site Report****Author(s):** Eric YEN¹**Co-author(s):** Felix Lee¹ ; HSIN-YEN Chen

¹ ASGC

Corresponding Author(s): eric@sinica.edu.tw

ASGC site report on facility deployment, recent activities, collaborations and plans.

Desired length:

12

Site reports / 56

Australia site report

Sean Crosby¹ ; Lucien Philip Boland¹

¹ *University of Melbourne (AU)*

Corresponding Author(s): lucien.boland@cern.ch, sean.christopher.crosby@cern.ch

2017 has been a year of change for the Australian HEP site. The loss of a staff member, migration of batch system, and increased use of cloud are just some of the changes happening in Australia. We will provide an update on the happenings in Australia.

Desired length:

12

Networking and security / 18

AutoGOLE bringing high speed data transfers

Gerben van Malenstein¹

¹ *SURFnet*

Corresponding Author(s): j-mambretti@northwestern.edu, gerben.vanmalenstein@surfnet.nl

The Automated GOLE (AutoGOLE) fabric enables research and education networks worldwide to automate their inter-domain service provisioning. By using the AutoGOLE control plane infrastructure, services to other countries can be setup in minutes. Besides automated provisioning we experiment with connecting high-speed Data Transfer Nodes (DTNs) to the AutoGOLE environment. This talk will discuss current possibilities, performance and future plans.

Desired length:

Basic IT services / 86

Automatic shutdown of servers in case of A/C failure

Peter Gronbech¹

¹ *University of Oxford (GB)*

Corresponding Author(s): pete.gronbech@physics.ox.ac.uk

Following various A/C incidents in an Oxford Computer room, we developed a solution to automatically shutdown servers.

The solution has two parts the service which monitors the temperatures and publishes on a web page and the client which runs on the servers, queries the result to determine if shutdown is required. Digttemp software and one wire temperature sensors are used.

Desired length:

10 minutes

Site reports / 31

BNL RACF Site Report

Ofer Rind^{None}

Corresponding Author(s): rind@bnl.gov

BNL's RHIC/ATLAS Computing Facility (RACF) serves the computing needs of experiments at RHIC, while also serving as the US ATLAS Tier-1 facility. In recent years, RACF has been expanding to serve a growing list of scientific communities at BNL. This presentation provides an overview of the RACF, highlighting significant developments since the last HEPiX meeting in Budapest.

Desired length:

12 minutes

Storage and file systems / 83

CEPH at RAL

Ian Collier^{None}

Corresponding Author(s): ian.peter.collier@cern.ch

In March 2017 Echo went in to production at the RAL Tier 1 providing over 7PB of usable storage to WLCG VOs. This talk will present details of the setup and operational experience gained from running the cluster in production.

Desired length:

20 minutes

Site reports / 19

CERN Site Report

Jerome Belleman¹

¹ CERN

Corresponding Author(s): jerome.belleman@cern.ch

News from CERN since the workshop at the Hungarian Academy of Sciences.

Basic IT services / 72

CSNS HPC Platform Based on SLURM

Yakang LI¹¹ *ihep***Corresponding Author(s):** liyk@ihep.ac.cn

China Spallation Neutron Source (CSNS) is a neutron source facility for studying neutron characteristics and exploring microstructure of matter, it will also serve as a high-level scientific research platform oriented to dimensional academic subjects. Scientific research on CSNS requires the support of a high-performance computing environment. So from the research and practice aspects, firstly, specific computational requirements of CSNS will be introduced in this report. Secondly, the design and practice of HPC platform are mainly demonstrated from the aspects of login system, job management system, storage system, network, etc. Finally, some future prospects of CSNS HPC platform are summarized in the ending of this paper.

Desired length:

20

Clouds, virtualisation, grids / 29

Cloud deployment at KEK

Wataru Takase¹¹ *High Energy Accelerator Research Organization (JP)***Corresponding Author(s):** wataru.takase@cern.ch, wataru.takase@kek.jp

Private cloud deployment is on going at KEK. Our cloud will support self-service provisioning, and also will be integrated our batch system in order to provide heterogeneous clusters dynamically. It enables us to support various kinds of data analyses and enabling elastic resource allocation among the various projects supporting at KEK.

In this talk, we will introduce our OpenStack based cloud infrastructure and report on the current status of the deployment. We will also describe our near-future plan of batch system integration of both private and public clouds.

Desired length:

20

Storage and file systems / 41

Cloud storage with the Dynafed data federator

Marcus Ebert¹¹ *University of Victoria***Corresponding Author(s):** mebert@uvic.ca

We describe our use of the Dynafed data federator with cloud computing resources. Dynafed (developed by CERN IT) allows a dynamic data federation, based on the webdav protocol, with the possibility to have a single name space for data distributed over all available sites. It also allows a

failover to another copy of a file in case the connection to the closest file location gets interrupted which makes it very robust and reliable in usage.
Specifically, we report on the challenges and progress on the implementation within the ATLAS and Belle II experiments, the implementation of GRID based authentication and authorization, and the use of S3 storage with Dynafed.

Desired length:

20 minutes

Networking and security / 16**Configuration automation for CERN's new Wi-Fi infrastructure**Quentin Barrand¹¹ CERN**Corresponding Author(s):** quentin.barrand@cern.ch

As presented at HEPiX Fall 2016, CERN is currently in the process of renewing its standalone Wi-Fi Access Points with a new state-of-the-art, controller-based infrastructure. With more than 4000 new Access Points to be installed, it is desirable to keep the existing deployment procedures and tools to avoid repetitive and error-prone actions during configuration and maintenance steps.
This presentation will dive into the new controller setup and will focus on the in-house software and workflows that power configuration automation for the new Wi-Fi infrastructure.

End-user services and operating systems / 26**Continuous Integration for Linux Images**Jerome Belleman¹¹ CERN**Corresponding Author(s):** jerome.belleman@cern.ch

The CERN Linux Support is in charge of providing system images for all Scientific Linux and CentOS CERN users. We currently mostly test new images manually. To streamline their generation towards production, we're designing a continuous integration and testing framework which will automate image production, allow for more tests, running them more thoroughly, with more flexibility.

Desired length:

20

Networking and security / 64**Current Status and Future Directions of KEK Computer Security**

Fukuko Yuasa¹ ; Tadashi Murakami¹ ; Ryouichi Baba¹ ; Teiji Nakamura¹ ; Kiyoharu Hashimoto¹ ; Soh Suzuki¹ ; Mittsuo Nishiguchi¹ ; Toshiaki Kaneko¹

¹ KEK

Corresponding Author(s): kiyoharu.hashimoto@kek.jp, ryouichi.baba@kek.jp, soh.suzuki@kek.jp, fukuko.yuasa@kek.jp, teiji.nakamura@kek.jp, mitsuo.nishiguchi@kek.jp, tadashi.murakami@kek.jp, toshiaki.kaneko@kek.jp

Recently Japanese universities and academic organizations had experienced sever cyber attacks. To mitigate computer security incidents, we are forced to rethink our strategies in aspects of security management and network design.

In this talk, we report current status and present future directions of KEK Computer security.

Desired length:

20 minutes

Site reports / 35

DESY site report

Dirk Jahnke-Zumbusch¹

¹ DESY

Corresponding Author(s): dirk.jahnke-zumbusch@desy.de

news about what happened at DESY during the last months

Desired length:

standard (15min?)

Basic IT services / 81

Deployment and monitoring for distributed computing sites

Wei Zheng¹

¹ IHEP

Corresponding Author(s): zhengw@ihep.ac.cn

Now IHEP can provide maintenance for those distributed computing sites, such as USTC and BUAA. We use both puppet and foreman to achieve these sites' automatic deployment and configuration, OS installation, system configuration and software upgrade. In order to realize unified maintenance, we adopt nagios to monitor this site's healthy status, including network, system, storage, services, etc. Mod-gearman, a module enable nagios to monitor remote sites, integrates remote monitor information into IHEP monitoring system. If sites have any errors, administrators at IHEP can use remote tools to handle these errors.

Desired length:

20 minutes

Networking and security / 34

Deployment of IPv6-only CPU on WLCG - an update from the HEPiX IPv6 Working Group

Dave Kelsey¹

¹ *STFC - Rutherford Appleton Lab. (GB)*

Corresponding Author(s): david.kelsey@stfc.ac.uk

This update from the HEPiX IPv6 Working Group will present the activities of the last 6-12 months. In September 2016, the WLCG Management Board approved the group's plan for the support of IPv6-only CPU, together with the linked requirement for the deployment of production Tier 1 dual-stack storage and other services. A reminder of the requirements for support of IPv6 and the deployment timelines of the plan will be presented. The current status will be reviewed including the deployment of dual-stack storage at the WLCG Tier 1s as well as the status and plans for deployment at Tier 2s.

Desired length:

20 minutes

Networking and security / 45

EGI CSIRT: Keeping EGI Secure

Vincent Brillault¹

¹ *CERN*

Corresponding Author(s): vincent.brillault@cern.ch

The EGI CSIRT main goal is, in collaboration with all resources providers, to keep the EGI e-Infrastructure running and secure. During the past years, under the EGI-Engage project, the EGI CSIRT has been driving the infrastructure in term of incident prevention and response, but also security training. This presentation provides an overview of these activities, focusing on the impact for the community, before concluding on current and future challenges for our infrastructures.

This talk is based on contributions and input from participants in the EGI CSIRT activities.

Desired length:

20

Networking and security / 78

Fancy Networking

Tristan Suerink¹

¹ *Nikhef National institute for subatomic physics (NL)*

Corresponding Author(s): t.suerink@cern.ch

We've redesigned our HPC/Grid network to be capable of full network function virtualisation, to be prepared for large amounts of 100Gbps connections, and to be 400G ready. In this talk we want to take you through the design considerations for a fully non-blocking 6 Tbps virtual network, and what type of features we have build-in for the cloudification of our clusters using OpenContrail. Through the integration of OpenContrail and OpenStack, we will be able to offer fully virtualised machines (and containers) for both IaaS and the (grid) Platform service, while retaining high-performance access to both local storage and our global network peers at 100+ Gbps.

In support of our wide area plans, in the beginning of this year we've done experiments using novel (and at that time experimental) DWDM equipment from Juniper between Amsterdam and Geneva, in collaboration with SURFnet and CERN.

We want to share the results and the pitfalls of these type of connections and why it is useful to do these tests.

Desired length:

20 minutes

Networking and security / 13

Firewall Load-Balancing solution at CERN

Vincent Ducret¹

¹ *CERN*

Corresponding Author(s): vincent.ducret@cern.ch

The CERN network infrastructure has several links to the outside world. Some are well identified and dedicated for experiments and research traffic (LHCOPN/LHCONE), some are more generics (general internet). For the latter, a specific firewall inspection is required for obvious security reasons, but with tens of gigabits per second of traffic, the firewalls capacities are highly challenged. This presentation will explain how CERN plans to move from a static firewall setup with limited capacity to a more flexible design using a Firewall Load Balancing solution. It will present the current setup, the on-going migration to a temporary firewall load balancing solution, and the long-term plans.

End-user services and operating systems / 84

First experience with SELinux

Michel Jouvin¹

¹ *Université Paris-Saclay (FR)*

Corresponding Author(s): jouvin@lal.in2p3.fr

... at LAL

Desired length:

20 minutes

Networking and security / 12

Follow-up about Wi-Fi service enhancement at CERN

Vincent Ducret¹

¹ *CERN*

Corresponding Author(s): vincent.ducret@cern.ch

As presented during HEPiX Fall 2016, a full renewal of the CERN Wi-Fi network was launched in 2016 in order to provide a state-of-the-art Campus-wide Wi-Fi Infrastructure. This year, the presentation will give a status and feedback about this overall deployment. It will provide information about the technical choices made, the methodology used for such a deployment, the issues we faced and how we solved them, our interaction with the manufacturer, and an overview of the current project status and schedule.

Site reports / 76

GSI Site Report

Jan Trautmann¹

¹ *GSI Darmstadt*

Corresponding Author(s): j.trautmann@gsi.de

News and updates from GSI IT, e.g.:
- status GreenITCube
- new asset management system

Desired length:

10

Computing and batch systems / 88

HEP Community White Paper

Michel Jouvin¹

¹ *Université Paris-Saclay (FR)*

Corresponding Author(s): jouvin@lal.in2p3.fr

A short introduction and status report

Desired length:

Computing and batch systems / 52

HEPiX Benchmarking Working Group - Status Report October 2017

Manfred Alef¹

¹ *Karlsruhe Institute of Technology (KIT)*

Corresponding Author(s): manfred.alef@kit.edu

The HEPiX Benchmarking Working Group has worked on a fast benchmark to estimate the compute power provided job slot or a IaaS VM. The Dirac Benchmark 2012 (DB12) is scaling well with the

performance at least of Alice and LHCb when running within a batch job. Now the group has started the development of a next generation long running benchmark as a successor of the current HS06 metric.

Desired length:

20 minutes

IT facilities / 42

Hot days with no mechanical cooling data center

Cary Whitney¹

¹ *LBNL*

Corresponding Author(s): clwhitney@lbl.gov

I'll talk about how the data collect helped the center get through a heat wave in the Berkeley area. This is significant since Berkeley computing center does not have any mechanical cooling and relies on the external air temperature and water supply. Talking about what data we thought we needed and what data we did need and how the idea of saving all the data and collecting as much as we can actually helped us in the lesson learned sessions after the event.

Desired length:

20

Site reports / 74

IHEP Site Report

Jingyan Shi¹ ; Xiaowei Jiang²

¹ *IHEP*

² *IHEPXXXXXXXXXXXXXXXXXX*

Corresponding Author(s): jiangxw@ihep.ac.cn, shijy@ihep.ac.cn

This report will talk about the current status and recent updates at IHEP Site since the Spring 2017 report, covering computing, network, storage and other related work.

Desired length:

12

Site reports / 40

INFN-T1 site report

Andrea Chierici^{None}

Corresponding Author(s): andrea.chierici@cern.ch

A brief report on Italian T1 activities.

Desired length:

12

Basic IT services / 73**Integrated Monitoring results at IHEP**Qingbao Hu¹ ; Wei Zheng¹ ; xiaowei jiang^{None}¹ *IHEP***Corresponding Author(s):** jiangxw@ihep.ac.cn, zhengw@ihep.ac.cn, huqb@ihep.ac.cn

Various cluster monitoring tools are adapted or developed at IHEP, which show the health status of each device or aspect of IHEP computing platform separately. For example, Ganglia shows the machine load, Nagios monitors the service status, and Job-monitor tool developed by IHEP counts the job success rate and so on. But those monitoring data from different tools are independent and not easy to be analyzed relatively. Integrate and analysis all the monitoring data from multiple sources can provide more valuable information such as health trends and potential errors.

Now, Integrated Monitoring Tools are deployed at IHEP which collects Ganglia, Nagios, Syslog and other monitor metrics. Some cluster monitoring projects based on this Integrated Monitoring Tools have been applied to IHEP.

Desired length:

20 minutes

Computing and batch systems / 53**Integrating HPC and HTC at BNL – a year later**Tony Wong¹¹ *Brookhaven National Laboratory***Corresponding Author(s):** tony@bnl.gov

This presentation discusses the new responsibilities of the Scientific Data & Computing Center (SDCC) in high-performance computing (HPC) and how we are leveraging effort and resources to improve BNL community's access to local and leadership-class facilities (LCF's).

Desired length:

20

Site reports / 55**KEK Site Report****Author(s):** Atsushi Manabe¹**Co-author(s):** Takashi Sasaki ²¹ *High Energy Accelerator Research Organization (KEK)*

² *High Energy Accelerator Research Organization (JP)*

Corresponding Author(s): takashi.sasaki@kek.jp, atsushi.manabe@kek.jp

We would like to introduce the brief history and report the current status of Computing Research Center at KEK. Many activities and near future plans on R&D, for example, networking, computer security, and private cloud deployment, which are submitted to the HEPiX workshop this time, will be summarized.

Desired length:

12 min.

Site reports / 61

KIT Site Report

Andreas Petzold^{None}

Corresponding Author(s): andreas.petzold@cern.ch

News about GridKa Tier-1 and other KIT IT projects and infrastructure. We'll focus on our experiences with our new 20+PB online storage installation.

Desired length:

12

Site reports / 82

KR-KISTI-GSDC-01 Tier-1 Site Reports

Jeongheon Kim¹

¹ *Korea Institute of Science and Technology Information*

Corresponding Author(s): jh.kim@kisti.re.kr

We will present the latest status of the GSDC. And migration plan of administrative system will be presented.

Desired length:

10 minutes

Miscellaneous / 8

LHCOPN and LHCONE

Edoardo Martelli¹

¹ *CERN*

Corresponding Author(s): edoardo.martelli@cern.ch

Desired length:

Miscellaneous / 5**Logistics information****Author(s):** Tomoaki Nakamura¹**Co-author(s):** Takashi Sasaki ¹¹ *High Energy Accelerator Research Organization (JP)***Corresponding Author(s):** tomoaki.nakamura@cern.ch**Desired length:****Networking and security / 33****Low-Power Wide-Area Network (LPWAN) at CERN**Rodrigo Sierra¹¹ *CERN***Corresponding Author(s):** rodrigo.sierra@cern.ch

The interest in the Internet of Things (IoT) is growing exponentially so multiple technologies and solutions have emerged to connect mostly everything. A ‘thing’ can be a car, a thermometer or a robot that, when equipped with a transceiver, will exchange information over the internet with a defined service. Therefore, IoT comprises a wide variety of user cases with very different requirements.

Low-Power Wide-Area Network (LPWAN) focuses on low-cost devices that need to operate on batteries for long periods and that send small volumes of data. LPWAN offers wireless connectivity for large areas and complements other technologies such as cellular machine-to-machine (M2M), Wi-Fi or personal area networks (PAN).

CERN has studied different options to offer wireless connectivity to non-critical sensors with very low throughput requirements and is preparing the deployment of an LPWAN network campus-wide. Provisioning and identity management, security and confidentiality, surface and underground availability or reliability and QoS have been some of the topics addressed during the network design.

Desired length:

20

Storage and file systems / 36**Managing cloudy dCache storage pools with Ansible.**Ulf Bobson Severin Tigerstedt¹¹ *Helsinki Institute of Physics (FI)***Corresponding Author(s):** ulf.bobson.severin.tigerstedt@cern.ch

NDGF-T1 is transferring the dCache storage to a model where dCache is no longer run by the sysadmin but run as a normal user. This enables centralized management of the software versions and their configs.

This automation is done with 3 roles in Ansible and a playbook to tie them together.
The end result is software running in an environment much like the cloud.

Desired length:

10 minutes

Computing and batch systems / 49

Migrating a WLCG tier-2 to a Cray XC-50 at CSCS-LCG2

Francesco Giovanni Sciacca¹ ; Miguel Gila²

¹ *Universitaet Bern (CH)*

² *ETH Zurich*

Corresponding Author(s): miguel.gila@cscs.ch, gsciacca@mail.cern.ch

Founded in 1991, CSCS, the Swiss National Supercomputing Centre, develops and provides the key supercomputing capabilities required to solve important problems to science and/or society. The centre enables world-class research and provides resources to academia, industry and the business sector. Through an agreement with CHIPP, the Swiss Institute of Particle Physics, CSCS hosts a WLCG tier-2 site, delivering computing and storage services to the ATLAS, CMS and LHCb experiments. In this presentation, we will describe the ongoing efforts in migrating all the tier-2 compute workloads for the served experiments from a dedicated x86_64 cluster that has been in continuous operation and evolution since 2007, to Piz Daint, the current European flagship HPC, which ranks third in the TOP500 at the time of writing.

Desired length:

20

Computing and batch systems / 3

Migration from Grid Engine to HTCondor

Author(s): Thomas Finnern¹

Co-author(s): Christoph Beyer² ; Martin Flemming¹ ; Yves Kemp³

¹ *DESY*

² *DESY HH*

³ *Deutsches Elektronen-Synchrotron (DE)*

Corresponding Author(s): thomas.finnern@desy.de, martin.flemming@desy.de, yves.kemp@cern.ch, christoph.beyer@desy.de

The migration of the local batch system BIRD required the adaptation of different properties like the Kerberos / AFS support, the automation of various operational tasks and the user and project access. The latter includes, inter alia, fairshare, accounting and resource access. For this, some newer features of HTCondor had to be used. We are close to the user release. Building common dynamic resources for local and grid-based batches is still in process. The talk provides some details about Kerberos support, authentication within the operating automation, job timing, and registry integration.

Desired length:

End-user services and operating systems / 2**Modernising CERN document conversion service.**

Ruben Domingo Gaspar Aparicio¹

¹ CERN

Corresponding Author(s): ruben.gaspar.aparicio@cern.ch

The document converter service provides conversion of most office and some engineering applications to PDF, PDF/A or PostScript. The service has been completely rewritten as an OSS [1] and is based on modern IT technology fostered by the CERN IT department. It is implemented as a RESTful API with a containerised approach using the Openshift technology, EOS storage to store documents and jobs, PostgreSQL database, Python3, flask, a Kibana dashboard based on Elastic and its documentation based on gitbook.

The project has been conceived having in mind a multiprocessing design, which allows handling simultaneously several jobs and reducing by more than half the time to process documents, compared to the old service incarnation. It allows adding different converter software, presently using Neevia [2]. The design allows to easily scale up thanks to the technology used, which is HAProxy + Openshift as web interface and Openstack VM's, Windows 2012R2 servers, as worker nodes in the backend.

Currently, the document converter service is mainly used by services like Indico [3] or EDMS [4] to automate conversion of thousands of documents.

[1] <https://github.com/CERNCDAIC/doconverter>

[2] <https://neevia.com/>

[3] <https://indico.cern.ch/>

[4] <https://edms.cern.ch/ui/>

Site reports / 23**NDGF Site Report**

Erik Mattias Wadenstein^{None}

Corresponding Author(s): mattias.wadenstein@cern.ch

News and updates from the NDGF Tier-1 site.

Focus on this report will be new disk and tape resources and some performance numbers from both.

Also some site news from our distributed sites.

Desired length:

10

Networking and security / 58**NetSage, a unified network measurement and visualization service**

Andrew Lee¹

¹ *Indiana University*

Corresponding Author(s): leea@iu.edu

Modern science is increasingly data-driven and collaborative in nature, producing petabytes of data that can be shared by tens to thousands of scientists all over the world. NetSage is a project to develop a unified open, privacy-aware network measurement, and visualization service to better understand network usage in support of these large scale applications. New capabilities to measure and analyze the utilization of international wide-area networks are essential to ensure end-users are able to take full advantage of such infrastructure. NetSage was developed to support the US National Science Foundation international networking program, but can be deployed in other settings. This talk will offer an overview of the project and emphasize recent developments within the project.

Desired length:

20

Networking and security / 66

Netbench –testing network devices with real-life traffic patterns

Stefan Nicolae Stancu¹ ; Adam Lukasz Krajewski¹

¹ *CERN*

Corresponding Author(s): stefan.stancu@cern.ch, adam.krajewski@cern.ch

Network performance is key to the correct operation of any modern datacentre or campus infrastructure. Hence, it is crucial to ensure the devices employed in the network are carefully selected to meet the required needs.

The established benchmarking methodology [1,2] consists of various tests that create perfectly reproducible traffic patterns. This has the advantage of being able to consistently assess the performance differences between various devices, but comes at the disadvantage of always using known, pre-defined traffic patterns (frame sizes and traffic distribution) that do not stress the buffering capabilities of the devices to the same extent as real-life traffic would.

Netbench is a network-testing framework, based on commodity servers and NICs, that aims at overcoming the previously mentioned shortcoming. While not providing identical conditions for every test, netbench enables assessing the devices' behaviour when handling multiple TCP flows, which closely resembles real-life usage.

Furthermore, due to the prohibitive cost of specialized hardware equipment that implements RFC tests [1,2], few companies/organisations can afford a large scale test setup. The compromise that is often employed is to use two hardware tester ports and feed the same traffic to the device multiple times through loop-back cables (the so called "snake-test"). This test fully exercises the per-port throughput capabilities, but barely stresses the switching engine of the device in comparison to a full-mesh test [3]. The per-port price of a netbench test setup is significantly smaller than that of a testbed made using specialized hardware, especially if we take into account the fact that generic datacentre servers can be time-shared between netbench and day-to-day usage. Thus, a large-scale multi-port netbench setup is easily affordable, and enables organisations/companies to complement the snake test with benchmarks that stress test the switching fabric of network devices.

The presentation will cover the design of the netbench software platform and subsequently present results from a recent evaluation of high-end routers conducted by CERN.

Netbench has a layered architecture and uses standard technologies. At its core, it relies on iperf3 [4] as an engine to drive TCP flows between servers. The orchestration platform that sets up multiple iperf3 sessions is written in Python and relies on XML-RPC for fast provisioning of flows. Per-flow statistics are gathered into a PostgreSQL database, and the results visualisation is based on a Python REST API and a web page using JavaScript and the D3.js library for displaying graphs. Statistics are

presented at different levels of detail allowing the human tester to quickly assess the overall state of a test from both per-node and per-pair (source-destination) statistics.

During its last call for tender for high-end routers, CERN has employed netbench for evaluating the behaviour of network devices when exposed to meshed TCP traffic. We will present results from several devices. Furthermore, during the evaluation it became apparent that, due to the temporary congestion caused by competing TCP flows, netbench provides a good estimation of the devices' buffering capabilities.

To summarize, we present netbench, a tool that allows provisioning TCP flows with various traffic distributions (pairs, partial and full-mesh). We consider netbench an essential complement to synthetic RFC tests [1][2], as it enables affordable, large-scale testing of network devices with traffic patterns that closely resemble real-life conditions.

[1] RFC 2544, Bradner, S. and McQuaid J., "Benchmarking Methodology for Network Interconnect Devices"

[2] RFC 2889, Mandeville, R. and Perser J., "Benchmarking Methodology for LAN Switching Devices"

[3] RFC 2285, Mandeville, R., "Benchmarking Terminology for LAN Switching Devices"

[4] iperf3 <http://software.es.net/iperf/>

Desired length:

20 minutes

Networking and security / 79

Network Automation for Intrusion Detection System

Author(s): Adam Lukasz Krajewski¹

Co-author(s): Stefan Nicolae Stancu¹ ; Liviu Valsan¹

¹ CERN

Corresponding Author(s): liviu.valsan@cern.ch, adam.krajewski@cern.ch, stefan.stancu@cern.ch

CERN networks are dealing with an ever-increasing volume of network traffic. The traffic leaving and entering CERN must be precisely monitored and analysed to properly protect the networks from potential security breaches. To provide the required monitoring capabilities, the Computer Security team and the Networking team at CERN have joined efforts in designing and deploying a scalable Intrusion Detection System (IDS). The initial setup, presented at the HEPiX Fall 2016 Workshop in Berkeley, featured a Brocade MLXe configured with OpenFlow to provide dynamic offload and selecting mirroring capabilities. Due to technical requirements, the setup has been evolved and currently leverages the Brocade SLX platform with network automation software (StackStorm / Brocade Workflow Composer) deployed for additional programmability and flexibility. The new technology stack is under testing with a promising perspective of production deployment in 2018.

Desired length:

20

Networking and security / 32

Network Functions Virtualisation Working Group Proposal

Marian Babik¹ ; Shawn Mc Kee²

¹ CERN

² University of Michigan (US)

Corresponding Author(s): shawn.mckee@cern.ch, marian.babik@cern.ch

High Energy Physics (HEP) experiments have greatly benefited from a strong relationship with Research and Education Network (REN) providers and thanks to the projects such as LHCOPN/LHCONE and REN contributions, have enjoyed significant capacities and high performance networks for some time. RENs have been able to continually expand their capacities to over-provision the networks relative to the experiments needs and were thus able to cope with the recent rapid growth of the traffic between sites, both in terms of achievable peak transfer rates as well as in total amount of data transferred. For some HEP experiments this has led to designs that favour remote data access where network is considered an appliance with almost infinite capacity. There are reasons to believe that the network situation will change due to both technological and non-technological reasons starting already in the next few years. Various non-technological factors that are in play are for example anticipated growth of the non-HEP network usage with other large data volume sciences coming online; introduction of the cloud and commercial networking and their respective impact on usage policies and securities as well as technological limitations of the optical interfaces and switching equipment.

As the scale and complexity of the current HEP network grows rapidly, new technologies and platforms are being introduced, collectively called Network Functions Virtualisation (NFV), ranging from software-based switches such as OpenVSwitch, Software Defined Network (SDN) controllers such as OpenDaylight up to full platform based open solutions such as Cumulus Linux. With many of these technologies becoming available, it's important to understand how we can design, test and develop systems that could enter existing production workflows while at the same time changing something as fundamental as the network that all sites and experiments rely upon. As this is not the only activity in the area and there are many different projects already running, it's important to find the effort to contribute and coordinate various different networking activities within HEP.

In this talk we'll describe these challenges and propose the formation of a NFV working group that would evaluate the existing technologies and provide guidance on their performance and adoption to the sites and experiments.

Desired length:

25 minutes

Networking and security / 75

Network related updates in IHEP

zhihui sun¹

¹ *Institute of High Energy Physics Chinese Academy of Sciences*

Corresponding Author(s): sunzh@ihep.ac.cn

We give the design and plan of network architecture updates in IHEP at HEPiX Spring 2017, and it has been finished in August 2017. This report talks about the network architecture upgrades, Dual stack ipv6 test, network measurement and monitoring at IHEP and network security upgrades.

Desired length:

15

Networking and security / 54

Next Generation Software Defined Services and the Global Research Platform

Joe Mambretti¹

¹ *Northwestern University*

Corresponding Author(s): j-mambretti@northwestern.edu

The Global Research Platform is a world-wide software defined distributed environment designed specifically for data intensive science. The talk will show how this environment could be used for experiments like the LHC

Desired length:

20

Basic IT services / 50

On Server Management Interface (BMC)

Alexandru Grigore¹

¹ *CERN*

Corresponding Author(s): alexandru.grigore@cern.ch

In this presentation, I will go over CERN's efforts in improving the security and usability of the management interfaces for various server manufacturers.

Desired length:

20

Computing and batch systems / 44

Optimising the resource needs for the LHC computing: ideas for a common approach

Jose Flix Molina¹ ; Markus Schulz² ; Andrea Sciaba²

¹ *Centro de Investigaciones Energéticas Medioambientales y Tecnológicas*

² *CERN*

Corresponding Author(s): andrea.sciaba@cern.ch, markus.schulz@cern.ch, jose.flix.molina@cern.ch

The increase of the scale of LHC computing expected for Run 3 and even more so for Run 4 (HL-LHC) over the course of the next 10 years will most certainly require radical changes to the computing models and the data processing of the LHC experiments. Translating the requirements of the physics programme into resource needs is an extremely complicated process and subject to significant uncertainties. Currently there is no way to do that without using complex tools and procedures developed internally by each LHC collaboration. Recently there has been much interest in developing a common model for estimating resource costs, which would be beneficial for experiments, WLCG and sites and in particular to understand and optimise the path towards HL-LHC. For example, it could be used to estimate the impact of changes in the computing models or to optimise the resource allocation at the site level. In this presentation we expose some preliminary ideas on how this could be achieved, with a special focus on the site perspective and provide some real world examples.

Desired length:

20

Site reports / 14**PDSF Site Report and Transition****Author(s):** Tony Quan¹**Co-author(s):** James Botts² ; Jeff Porter³ ; Jan Balewski⁴ ; Douglas Jacobsen⁴ ; Lisa Gerhardt⁴ ; Tina Declerck⁴¹ *LBL*² *LBNL*³ *Lawrence Berkeley National Lab. (US)*⁴ *NERSC***Corresponding Author(s):** jfbotts@lbl.gov, dmjacobsen@lbl.gov, rjporter@lbl.gov, twquan@lbl.gov, lgerhardt@lbl.gov, tmdeclerck@lbl.gov, balewski@lbl.gov

PDSF, the Parallel Distributed Systems Facility, was moved to Lawrence Berkeley National Lab from Oakland CA in 2016. The cluster has been in continuous operation since 1996 serving high energy physics research. The cluster is a tier-1 site for Star, a tier-2 site for Alice and a tier-3 site for Atlas.

The PDSF cluster is in transition this year, moving the batch system from UGE to SLURM and to delivering computing environments using Shifter, a NERSC software package for deploying docker-like containers. In the near two years, the Linux cluster hosting PDSF will be retired and its workloads will be moved to a Cray XC-40 system. This site report will describe recent updates to the system, upcoming transitions, and the future of High Energy Physics workloads at NERSC.

Site reports / 30**PIC report**Jose Flix Molina¹¹ *Centro de Investigaciones Energéticas Medioambientales y Tecnológicas***Corresponding Author(s):** jose.flix.molina@cern.ch

This is the PIC report to HEPiX Fall 2017.

Desired length:

12

Storage and file systems / 1**Particulate magnetic tape for data storage and future technologies**Masahito Oyanagi¹¹ *Recording Media Research Laboratories, Fujifilm Corporation*

Enterprise tape drives are widely used at major laboratories in the world, such as CERN, US DoE Labs, KEK and so on as well as data centers in commercial companies. Demands on capacity and speed

of I/O inflate infinitely in the tape market. Not only drive technology but also media technology is the key for answering such future requirements. Fujifilm is the world-leading company in the market of magnetic tape media, and has played a major role for evolution of tape technologies in decades.

Secure digital data storage sustainable in very long duration with lower costs is the concern of data centers. Particulate magnetic tapes on linear tape storage systems have been widely used for data backup and archival because of their low TCO, long-term stability, and high reliability. However, in order to meet the further demand, expanding the recording capacity of tape storage is essential. Currently, tape cartridges with a capacity of up to 10 TB are available commercially, and the future roadmap of tape storage systems shows to expand the cartridge capacity to 120 TB. To realize this, technologies to increase the areal density of magnetic tape are required. For this purpose, high-density recording studies using barium ferrite (BaFe) magnetic particles have been carried out. The latest study demonstrated an areal density of 123 Gb/in², corresponding to a 220 TB cartridge capacity in 2015. Furthermore, fine strontium ferrite (SrFe) magnetic particles, which are almost half the size of the current BaFe particles, were also developed; therefore, magnetic tapes are the most effective data storage media to continue enhancing the recording capacity at a low TCO in long-term future.

Desired length:

Site reports / 70

RAL Site Report

Martin Bly¹

¹ *STFC-RAL*

Corresponding Author(s): martin.bly@stfc.ac.uk

An update on activities at the UK Tier1 @ RAL

Desired length:

12

Networking and security / 63

Recent network connectivity around KEK

Soh Suzuki^{None}; Fukuko YUASA¹; Tomoaki Nakamura²; Go Iwai¹; Takanori Hara²; Ryosuke ITOH^{None}; Tomoyuki Konno³

¹ *KEK*

² *High Energy Accelerator Research Organization (JP)*

³ *Tokyo Institute of Technology*

Corresponding Author(s): soh.suzuki@kek.jp, go.iwai@kek.jp, fukuko.yuasa@kek.jp, takanori.hara@kek.jp, ryosuke.itoh@kek.jp, konno@hep.phys.titech.ac.jp, tomoaki.nakamura@cern.ch

Last year, KEK had upgraded the upstream link to 100Gbps in Apr. then officially started the peer with LHCONE since Sep. Then KEK can distribute huge data to WLCG sites by adequate throughput although this upgrade didn't made large impact on the firewalls for the ordinary internet usage from the campus network.

We will report changes by the LHCONE peer and how we connect our campus network and computing resources, and the data acquisition system for Belle II that will be a major source of rawdata on our computing resource.

Desired length:

15

Clouds, virtualisation, grids / 43

Running HEP Payloads on Distributed Clouds

Rolf Seuster¹ ; Randy Sobie¹ ; Frank Berghaus¹ ; Marcus Ebert² ; Michael Paterson³ ; Colin Roy Leavett-Brown¹ ; Kevin Casteels¹

¹ *University of Victoria (CA)*

² *University of Victoria*

³ *U*

Corresponding Author(s): kevin.casteels@cern.ch, mebert@uvic.ca, colin.roy.leavett-brown@cern.ch, frank.berghaus@cern.ch, mhph@uvic.ca, rolf.seuster@cern.ch, rsobie@cern.ch

The University of Victoria HEP group has been successfully running on distributed clouds for several years using the CloudScheduler/HTCondor framework. The system uses clouds in North America and Europe including commercial clouds. Over the last years, the operation has been very reliably, we are regularly running several thousands of jobs concurrently for the ATLAS and Belle II experiments. Currently we are writing a new version of CloudScheduler (version 2) that aims at further increasing the scale of number of jobs to levels in excess of 10,000. Further, it will be easier to configure existing and new clouds, as well as automate more of the operation of the clouds. We describe our operation experience and review the planned changes to the system.

Desired length:

15

Computing and batch systems / 85

Running Jobs Everywhere: an Overview of Batch Services at CERN

Jerome Belleman¹

¹ *CERN*

Corresponding Author(s): jerome.belleman@cern.ch

Batch services at CERN have diversified such that computing jobs can be run everywhere, from traditional batch farms, to disk servers, to people's laptops, to commercial clouds. This talk offers an overview of the technologies and tools involved.

Desired length:

10 minutes

Basic IT services / 39**Securing Elasticsearch for free: integration with SSO and Kerberos at CC-IN2P3**Fabien Wernli¹¹ CCIN2P3**Corresponding Author(s):** wernli@in2p3.fr

It is now a well-known fact in the *HEPiX* community that the *Elastic* stack (FKA *ELK*) is an extremely useful tool to dive into huge log data entries. It has also been presented multiple times as lacking the security features so often needed in multi-user environments. Although it now provides

a plugin addressing some of those concerns, it requires the acquisition of a commercial license.

We present *floragunn's Searchguard*: an *Elasticsearch* plugin that provides authentication, authorization

and encryption. It also bundles a *Kibana* plugin that offers multi-tenant views and dashboards.

We then focus on its integration with *Kerberos*, *CAS* (SSO) and *syslog-ng* at *CC-IN2P3*.

If time permits we'll present gotchas and performance considerations.

Desired length:

20

Networking and security / 46**Security update**Romain Wartel¹¹ CERN**Corresponding Author(s):** romain.wartel@cern.ch

This presentation gives an overview of the current computer security landscape. It describes the main vectors of compromises in the academic community including lessons learnt, and reveal inner mechanisms of the underground economy to expose how our resources are exploited by organised crime groups, as well as recommendations to protect ourselves. By showing how these attacks are both sophisticated and profitable, the presentation concludes that the only mean to adopt and appropriate response is to build a tight international collaboration and trusted information sharing mechanisms within the community.

Desired length:

30

Networking and security / 47**Security: case study**Romain Wartel¹ ; Vincent Brillault¹¹ CERN

Corresponding Author(s): vincent.brillault@cern.ch, romain.wartel@cern.ch

This is a TLP:RED presentation of a case study. Slides and details will not be made publicly available, and attendees have to agree to treat all information presented as confidential and refrain from sharing details on social media or blog. The presentation focuses on an insider attack and concentrates on the technical aspects of the investigation, in particular the network and file system forensics led by EGI and WLCG security experts, as well as key lessons learnt.

Desired length:

30

Clouds, virtualisation, grids / 20

Singularity at the RACF/SDCC

Author(s): Christopher Hollowell¹

Co-author(s): Xin Zhao²

¹ Brookhaven National Laboratory

² Brookhaven National Laboratory (US)

Corresponding Author(s): hollowec@bnl.gov, xzhao@bnl.gov

In this presentation, we'll give an overview of the Singularity container system, and our experience with it at the RACF/SDCC at Brookhaven National Laboratory. We'll also discuss Singularity's advantages over virtualization and other Linux namespace-based container solutions in the context of HTC and HPC applications. Finally, we'll detail our future plans for this software at our facility.

Storage and file systems / 68

Storage for Science at CERN

Giuseppe Lo Presti¹

¹ CERN

Corresponding Author(s): giuseppe.lopresti@cern.ch

In this contribution the vision for the CERN storage services and their applications will be presented. Traditionally, the CERN IT Storage group has been focusing on storage for Physics data. A status update will be given about CASTOR and EOS, with the recent addition of the Ceph-based storage for High-Performance Computing. More recently, the evolution has focused on providing higher-level tools to access, share and interact with the data. CERNBox is at the center of this strategy, as it provides the broader scientific community with high level applications (Office, SWAN) to interact with the data. Examples of ongoing collaborations with other non-HEP institutes will be given, where scientists are directly enabled to perform their data processing, leveraging this approach.

Desired length:

20

Site reports / 21**Swiss National Supercomputing Centre - T2 Site Report**

Dino Conciatore¹ ; Dario Petrusic¹ ; Gianni Mario Ricciardi¹ ; Miguel Gila²

¹ *Eidgenoessische Technische Hochschule Zuerich (CH)*

² *ETH Zurich*

Corresponding Author(s): dino.conciatore@cscs.ch, miguel.gila@cscs.ch, dario.petrusic@cern.ch, gianni.mario.ricciardi@cern.ch

Site report, news and ongoing activities at the Swiss National Supercomputing Centre (CSCS-LCG2), running ATLAS, CMS and LHCb.

Computing and batch systems / 51**Techlab update**

Romain Wartel¹ ; Aritz Brosa Iartza² ; Maxime Reis¹

¹ *CERN*

² *Universidad de Oviedo (ES)*

Corresponding Author(s): maxime.reis@cern.ch, aritz.brosa.iartza@cern.ch, romain.wartel@cern.ch

Techlab is a CERN IT activity aimed at providing facilities for studies improving the efficiency of the computing architecture and making better utilisation of the processors available today.

It enables HEP experiments, communities and project to gain access to machines of modern architectures, for example Power 8, GPUs and ARM64 systems.

The hardware is periodically updated based on community feedback and industry trends. All results can be published; studies requiring non-disclosure agreements (NDAs) are not allowed.

Techlab also runs extensive evaluations of all its systems and makes them available on its benchmarking portal, which is open for all HEP participants to view and share benchmarking results.

This presentation gives a status update of Techlab, including service status, current hardware and plans, and gives an overview of the benchmarking portal.

Desired length:

20

Miscellaneous / 6**The HEPiX workshop series**

Helge Meinhard¹ ; Tony Wong²

¹ *CERN*

² *Brookhaven National Laboratory*

Corresponding Author(s): tony@bnl.gov, helge.meinhard@cern.ch

Desired length:

Miscellaneous / 7

The HPSS User Forum (HUF)

Desired length:

Networking and security / 15

The LHCONE network

Vincenzo Capone¹

¹ *GÉANT*

Corresponding Author(s): vincenzo.capone@geant.org

LHCONE is a worldwide network dedicated to the data transfers of HEP experiments. The presentation will explain the origin and the architecture of the network, the services and advantages it provides, the benefits achieved so far. It will also include an update with the latest achievements

Storage and file systems / 69

The Outlook for Archival Storage at CERN

Michael Davis¹ ; German Cancio Melia¹

¹ *CERN*

Corresponding Author(s): german.cancio.melia@cern.ch, michael.davis@cern.ch

The CERN Physics Archive is projected to reach 1 Exabyte during LHC Run 3. As the custodial copy of the data archive is stored on magnetic tape, it is very important to CERN to predict the future of tape as a storage medium.

This talk will give an overview of recent developments in tape storage, and a look forward to how the archival storage market may develop over the next decade. The presentation will include a status update on the new CERN Tape Archive software.

Desired length:

20 minutes

Networking and security / 57

The TransPAC project

Andrew Lee¹

¹ *Indiana University*

Corresponding Author(s): leea@iu.edu

The TransPAC project has a long history of supporting R&E networking, connecting the Asia Pacific region to the United States to facilitate research. This talk will give an overview of the project for those who may not be familiar with it or its activities and a brief sketch of future plans. Then the

talk will cover LHCONE connectivity from our perspective and lay out options for how TransPAC can help the LHCONE community along with our colleagues from ASGC.

Desired length:

20

Miscellaneous / 60

The birth of HEPiX - at KEK in 1991

Helge Meinhard¹

¹ *CERN*

Corresponding Author(s): helge.meinhard@cern.ch

I will report on material describing how HEPiX was born at KEK in 1991.

Desired length:

20

Networking and security / 24

Tier-1 networking cost and optimizations

Erik Mattias Wadenstein^{None}

Corresponding Author(s): mattias.wadenstein@cern.ch

As the WLCG data sets grow ever bigger, so will network usage. For those of us with limited budgets, it would be nice if network costs won't get ever bigger too.

As NDGF is one of the few tier-1 sites in WLCG required to pay full networking costs, including transit, we'll look at the cost breakdown of networking for a tier-1 site and talk about where optimizations might be found.

Desired length:

20

Site reports / 28

Tokyo Tier-2 Site Report

Tomoe Kishimoto¹

¹ *University of Tokyo (JP)*

Corresponding Author(s): tomoe.kishimoto@cern.ch

The Tokyo Tier-2 site, which is located in International Center for Elementary Particle Physics (ICEPP) at the University of Tokyo, is providing computing resources for the ATLAS experiment in the WLCG.

Updates on the site since the Spring 2017 meeting and a migration plan for the next system upgrade will be presented.

Desired length:

12

Site reports / 48

US ATLAS SWT2 Site Report

Horst Severini¹

¹ *University of Oklahoma (US)*

Corresponding Author(s): hs@nhn.ou.edu

We will present an update on our sites and cover our work with various efforts like xrootd storage elements, opportunistic usage of general HPC resources, and containerization.

We will also report on our latest hardware purchases, as well as the status of network updates.

We conclude with a summary of successes and problems we encountered and indicate directions for future work.

Desired length:

12 min

Site reports / 37

University of Wisconsin-Madison CMS T2 site report

Ajit Kumar Mohapatra¹ ; Sridhara Dasu¹ ; Daniel Charles Bradley² ; Carl Vuosalo¹ ; Chad W Seys³

¹ *University of Wisconsin Madison (US)*

² *University of Wisconsin-Madison (US)*

³ *University of Wisconsin-Madison*

Corresponding Author(s): seys@wisc.edu, daniel.bradley@cern.ch, carl.vuosalo@cern.ch, dasu@hep.wisc.edu, ajit.kumar.mohapatra@cern.ch

As a major WLCG/OSG T2 site, the University of Wisconsin-Madison CMS T2 has consistently been delivering highly reliable and productive services towards large scale CMS MC production/processing, data storage, and physics analysis for last 11 years. The site utilises high throughput computing (HT-Condor), highly available storage system (Hadoop), scalable distributed software systems (CVMFS), and provides efficient data access using xrootd/AAA. The site fully supports IPv6 networking, and is a member of the LHCONE community with 100Gb WAN connectivity. An update on the activities and developments at the T2 facility over the last year (since the LBNL meeting) will be presented.

Desired length:

12

Basic IT services / 71

Updates from Database Services at CERN

Andrei Dumitru¹

¹ *CERN*

Corresponding Author(s): andrei.dumitru@cern.ch

CERN has a great number of applications that rely on a database for their daily operations and the IT Database Services group is responsible for current and future databases and their platform for accelerators, experiments and administrative services as well as for scale-out analytics services including Hadoop, Spark and Kafka. This presentation aims to give a summary of the current state of the database services at CERN, the recent migration to new hardware and some insights into the evolution of our services.

Desired length:

20

Networking and security / 17

Using Configuration Management to deploy and manage network services

Quentin Barrand¹

¹ *CERN*

Corresponding Author(s): quentin.barrand@cern.ch

Configuration Release Management (CRM) is rapidly gaining popularity among service managers, as it brings version control, automation and lifecycle management to system administrators. At CERN, most of the virtual and physical machines are managed through the Puppet framework, and the networking team is now starting to use it for some of its services.

This presentation will focus on the specificities of using CRM for network services, and will list a few pitfalls that can be avoided.

Clouds, virtualisation, grids / 27

Using Docker containers for scientific environments - on-premises and in the cloud

Sergey Yakubov¹ ; Martin Gasthuber¹ ; Birgit Lewendel²

¹ *DESY*

² *Deutsches Elektronen-Synchrotron (DE)*

Corresponding Author(s): birgit.lewendel@desy.de, martin.gasthuber@desy.de, sergey.yakubov@desy.de

Docker container virtualization provides an efficient way to create isolated scientific environments, adjusted and optimized for a specific problem or a specific group of users. It allows to efficiently separate responsibilities - with IT focusing on infrastructure for image repositories, preparation of basic images, container deployment and scaling, and physicists focusing on application development in environment of their choice.

Depending on demand, compute resources can be dynamically provisioned and a containerised scientific environment can be deployed in a matter of seconds on a user laptop, a batch farm, an HPC cluster or a cloud without need for a user to learn new environment, install additional libraries, resolve dependencies, recompile applications.

The present talk will describe DESY's experience with providing Docker service on our HPC cluster and report progress in using cloud to transparently and elastically extend containerised scientific environments - a work being done within the HELIX NEBULA Science Cloud project.

Desired length:

15

Storage and file systems / 87

WLCG archival storage group

Oliver Keeble¹ ; Helge Meinhard¹

¹ *CERN*

Corresponding Author(s): helge.meinhard@cern.ch, oliver.keeble@cern.ch

Brief introduction, and call for contributions, to a working group on archival storage at WLCG sites

Desired length:

10 minutes

Networking and security / 25

WLCG/OSG Networking Update

Marian Babik¹ ; Shawn Mc Kee²

¹ *CERN*

² *University of Michigan (US)*

Corresponding Author(s): shawn.mckee@cern.ch, marian.babik@cern.ch

WLCG relies on the network as a critical part of its infrastructure and therefore needs to guarantee effective network usage and prompt detection and resolution of any network issues, including connection failures, congestion and traffic outing. The OSG Networking Area is a partner of the WLCG effort and is focused on being the primary source of networking information for its partners and constituents. We will report on the changes and updates that have occurred since the last HEPiX meeting.

The WLCG Throughput working group was established to ensure sites and experiments can better understand and fix networking issues. In addition, it aims to integrate and combine all network-related monitoring data collected by the OSG/WLCG infrastructure from both network and transfer systems. This has been facilitated by the already existing network of the perfSONAR instances that is being commissioned to operate in full production.

We will provide a status update on the LHCOPN/LHCONE perfSONAR infrastructure as well as cover recent changes in the higher level services that were developed to help bring perfSONAR network to its full potential. This includes a new set of dashboards based on Grafana that offer a combined view on network utilisation and network performance as measured by RENs, WLCG

transfers and perfSONAR; updates and changes to the Web-based mesh configuration system and OSG network datastore (esmond), which collects, stores and provides interfaces to access all the network monitoring information from a single place.

In addition, we will provide an overview of the recent major network incidents that were investigated with the help of perfSONAR infrastructure and provide information on changes that are included in the recent release of version 4.0.1 of the perfSONAR Toolkit. We will also cover the status of our WLCG/OSG deployment and provide some information on our future plans.

Desired length:

20 minutes

Miscellaneous / 4

Welcome to KEK

Author(s): Yasuhiro Okada¹

Co-author(s): Takashi Sasaki²; Tomoaki Nakamura²

¹ KEK

² High Energy Accelerator Research Organization (JP)

Corresponding Author(s): takashi.sasaki@kek.jp, tomoaki.nakamura@cern.ch, yasuihiro.okada@kek.jp

Desired length:

Basic IT services / 67

Wigner Datacenter's new software defined datacenter architecture

Author(s): Zoltan Szelezky¹

Co-author(s): Balazs Bago²

¹ Wigner Datacenter

² Hungarian Academy of Sciences (HU)

Corresponding Author(s): balazs.bago@cern.ch, szelezky.zoltan@wigner.mta.hu

Our cloud deployment at Wigner Datacenter (WDC) is undergoing significant changes. We are adapting a new infrastructure, an automated OpenStack deployment using TripleO and configuration management tools like Puppet and Ansible. Over the past few months, our team at WDC have been testing TripleO as the base of our OpenStack deployment. We are also planning a centralized monitoring and logging system and open-source firewall solution.

The goal is to create an highly scalable and secure open-source SDDC (software defined datacenter) for the scientific community. In this presentation we are going to show the software and the toolset we use, and our progress and challenges so far.

Desired length:

20

Miscellaneous / 80

Workshop wrap-up

Helge Meinhard¹ ; Tony Wong²

¹ CERN

² Brookhaven National Laboratory

Corresponding Author(s): tony@bnl.gov, helge.meinhard@cern.ch

Desired length:

End-user services and operating systems / 77

printing@gsi

Stefan Haller¹

¹ GSI Darmstadt

Corresponding Author(s): s.haller@gsi.de

Some remarks to current design with printer subnets and managing CUPS configuration via CHEF data bags.

Desired length:

10

Basic IT services / 38

riemann: a different stream processor

Fabien Wernli¹

¹ CCIN2P3

Corresponding Author(s): wernli@in2p3.fr

We present *riemann*: a low-latency transient shared state stream processor.

This opensource monitoring tool is written by Kyle Kingsbury and maintained by the community. Its unique design makes it as flexible as it gets by melting the walls between configuration and code. Whenever its rich API doesn't fit the use-case, it's as simple as using any library in the *clojure* or *java* ecosystem and importing it into the configuration.

We present *riemann*'s basic concepts, as well as some key elements of its API.

Moreover, we illustrate its usefulness in our Tier-1 by showing a few examples of use-cases at CC-IN2P3.

For instance, its ability to aggregate thousands of metrics along high-level metadata keys using straightforward configuration entries is illustrated. Its integration with other monitoring tools like *Nagios*, *InfluxDB*, *Elasticsearch*, etc. is presented.

Desired length:

20