

HEPiX Autumn 2020 online Workshop

Monday 12 October 2020 - Friday 16 October 2020

Online Workshop



Book of Abstracts

Contents

CERN Site Report 1	1
PIC report 2	1
Diamond Light Source Site Report 4	1
Welcome & Logistics 5	2
Welcome & Logistics 14	2
DESY Site Report 23	2
Learning-based Approaches to Estimate Job Wait Time in HTC Datacenters 24	2
Benchmarking Working Group: A status report 25	3
Computer Security Update 26	3
Anomaly detection for the Centralized Elasticsearch service at CERN 27	3
IHEP Site Report 28	4
Simulated phishing campaigns at CERN 30	4
WLCG/OSG Network Activities, Status and Plans 31	4
Network Functions Virtualisation Working Group Update 32	5
CERN Cloud Infrastructure status update 33	5
Kubernetes in the CERN Cloud 34	5
EOS storage for Alice O2 35	6
The design of networking and computing system for High Energy Photon Source(HEPS) 36	6
BNL Site Report 37	7
CERN's Business Continuity Working Group 38	7
CERNphone – CERN's upcoming softphone solution 39	7
FTS: Towards tokens, QoS, archive monitoring and beyond 40	7
Roadmap for the DNS Load Balancing Service at CERN 41	8

Federated Identity Management at BNL 42	8
An Evaluation of Podman 43	9
HPSS migration to IBM tape technologies 44	9
Development of online application portal 45	9
IGWN submit node: Designing EU-based gateway nodes for distributed computing for Virgo, LIGO, Kagra 46	10
MySQL High Availability in the Database On Demand service 47	10
Towards a redundant, robust, secure and reliable IoT network 48	11
KEK Site Report 49	11
Operating a production HTCondor cluster: Seamlessly automating maintenance, OS and HTCondor updates with (almost) zero downtimes 50	11
Watching File Storage and Transfers with Elastic Stack at BNL 51	12
Running event-driven workflows with dCache storage events 52	12
A New CMS for a New Decade: Content Management at SDCC 53	13
The BNLBox service at the SDCC 54	13
Cloud computing to support experiment online computing from the data center. 56	13
Scalable High Performance Storage based on Lustre/ZFS over NVMe SSD 57	14
NOTED - identification of data transfer in FTS to understand network traffic. 58	14
ASGC Site Report 59	14
Invenio Based Digital Repositories at BNL 60	15
IPv6-only on WLCG - update from the IPv6 working group 61	15
CVMFS service evolution and infrastructure improvements 62	15
HIFIS backbone transfer service: FTS for everyone 63	16
XCache: exploring cache storage technology 64	17
MALT Project 65	17
CERN Appstore for BYOD devices 66	17
Workshop Wrap-Up 67	17
HEPiX board meeting 68	18
HTCondor 2020 workshop report 69	18
Experiences from online Wokshops and Conferences 70	18

dCache in the cloud environment 71 18

Production deployment of the CERN Tape Archive (CTA) for Atlas 72 18

group photo morning 73 19

group photo evening 74 19

Security Update 76 19

Security update 77 19

Monday morning / 1**CERN Site Report****Author:** Andrei Dumitru¹¹ CERN**Corresponding Author:** andrei.dumitru@cern.ch

News from CERN since the previous HEPiX workshop.

Monday evening / 2**PIC report****Author:** Jose Flix Molina¹**Co-authors:** Gonzalo Merino ²; Maria Del Carmen Porto Fernandez ³; Vanessa Acin Portella ⁴; Esther Accion Garcia ⁵; Carlos Acosta Silva ⁶; Ricard Cruz Josa ⁵; Elena Planas ⁶; Bruno Rodríguez ; Jordi Casals Hernandez ⁷; Antonio Perez-Calero Yzquierdo ¹; Andreu Pacheco Pages ⁸; Carlos Perez Dengra ⁶; Christian Neissner ⁶; Manuel Delfino Reznicek ⁵; Jorge Carretero Palacios ⁹; Jordi Delgado ⁶; Pau Tallada Crespi ¹⁰; Francesc Torradeflot ⁶; Agustín Bruzzese ⁶¹ *Centro de Investigaciones Energéticas Medioambientales y Tecnológicas*² *IFAE - Institute for High Energy Physics*³ *Unknown*⁴ *Institut de Física d'Altes Energies*⁵ *The Barcelona Institute of Science and Technology (BIST) (ES)*⁶ *PIC*⁷ *Port d'Informació Científica*⁸ *Institut de Física d'Altes Energies - Barcelona (ES)*⁹ *Institute for High Energy Physics (IFAE)*¹⁰ *PIC-CIEMAT***Corresponding Authors:** torradeflot@pic.es, eplanas@pic.es, pacheco@ifae.es, vacin@pic.es, neissner@pic.es, tallada@pic.es, bruzzese@pic.es, carretero@pic.es, jose.flix.molina@cern.ch, carlos.acosta.silva@cern.ch, mporto@pic.org.es, jordidem@pic.es, rcruz@pic.es, bruno@pic.es, merino@pic.es, delfino@pic.es, jcasals@pic.es, antonio.perez.calero.yzquierdo@cern.ch, esthera@pic.es, cperez@pic.es

This is the PIC report for HEPiX Autumn 2020 Workshop

Monday evening / 4**Diamond Light Source Site Report****Author:** James Thorne¹¹ *Diamond Light Source***Corresponding Author:** j.i.thorne@gmail.com

Diamond Light Source is a synchrotron light source in Oxfordshire, UK. We will present the developments at Diamond since HEPiX Autumn 2019 and our plans for the future.

Monday morning / 5

Welcome & Logistics

Corresponding Author: peter.van.der.reest@desy.de

Monday evening / 14

Welcome & Logistics

Corresponding Author: peter.van.der.reest@desy.de

Monday evening / 23

DESY Site Report

Authors: Peter van der Reest¹; Andreas Haupt²

¹ *DESY*

² *Deutsches Elektronen-Synchrotron (DESY)*

Corresponding Authors: peter.van.der.reest@desy.de, andreas.haupt@desy.de

News from the center since Autumn 2019

Computing and Batch Services / 24

Learning-based Approaches to Estimate Job Wait Time in HTC Datacenters

Authors: Luc Gombert¹; Frederic Suter¹

¹ *CC-IN2P3 / CNRS*

Corresponding Authors: frederic.suter@cc.in2p3.fr, luc.gombert@gmail.com

High Throughput Computing (HTC) datacenters are a cornerstone of scientific discoveries in the fields of High Energy Physics and Astroparticles Physics. These datacenters provide thousands of users from dozens scientific collaborations with tens of thousands computing cores and Petabytes of storage. The scheduling algorithm used in such datacenters to handle the millions of (mostly single-core) jobs submitted every month ensures a fair sharing of the computing resources among user groups, but may also cause unpredictably long job wait times for some users. The time a job will wait can be caused by many entangled factors and configuration parameters and is thus very hard to predict. Moreover, batch systems implementing a fair-share scheduling algorithm cannot provide users with any estimation of the job wait time at submission time.

Therefore, in this talk we investigate how learning-based techniques applied to the logs of the batch scheduling system of a large HTC datacenter can be used to get such an estimation of job wait time. After having illustrated the need for users to get an estimation of the time their jobs will wait, we identify some intuitive causes of this wait time based on the analysis of the information found in the batch system logs. Then, we formally analyze the correlation between these intuitive causes and job wait time and propose learning-based estimators of both job wait time and job wait time ranges.

We conclude by presenting the obtained preliminary results and thoughts about how to deploy the proposed estimators in production.

Tuesday evening / 25

Benchmarking Working Group: A status report

Authors: Michele Michelotto¹; Domenico Giordano²; Manfred Alef³

¹ *Universita e INFN, Padova (IT)*

² *CERN*

³ *Karlsruhe Institute of Technology (KIT)*

Corresponding Authors: manfred.alef@kit.edu, michele.michelotto@cern.ch, domenico.giordano@cern.ch

The HEPiX benchmarking working group has been very active in the past months and will present a report of the activities, concentrating on the new developments after the last report. The new candidate benchmark, designed to overcome the HEP-SPEC06 problems is being published. The HEP Benchmarking Suite has been complemented by GPU and Root benchmark.

Tuesday morning / 26

Computer Security Update

Author: Nikolaos Filippakis¹

¹ *CERN*

Corresponding Author: nikolaos.filippakis@cern.ch

This presentation aims to give an update on the global security landscape from the past year. The COVID-19 pandemic has introduced a novel challenge for security teams everywhere by expanding the attack surface to include everyone's personal devices / home networks and causing a shift to new, risky software for a remote-first working environment. It was also a chance for attackers to get creative by taking advantage of the fear and confusion to devise new tactics and techniques. What's more, the worrying trend of data leaks, password dumps, ransomware attacks and new security vulnerabilities does not seem to slow down. We present some interesting cases that CERN and the wider HEP community dealt with in the last year, mitigations to prevent possible attacks in the future and preparations for when inevitably an attacker breaks in.

This talk is based on contributions and input from the CERN Computer Security Team.

Thursday morning / 27

Anomaly detection for the Centralized Elasticsearch service at CERN

Author: Ulrich Schwickerath¹

¹ *CERN*

Corresponding Author: ulrich.schwickerath@cern.ch

For several years CERN offers a centralized service for Elasticsearch. This dynamic infrastructure consists of currently about 30 independent Elasticsearch clusters, covering more than 180 different use cases. Using internal monitoring data, a real time anomaly detection system has been implemented, and is now used in production. This presentation describes how the system works, the experiences gained so far, and the lessons learned from this exercise.

Monday morning / 28

IHEP Site Report

Author: Haibo li¹

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

Corresponding Author: lihaibo@ihep.ac.cn

News and updates from IHEP since the last HEPiX Workshop. In this talk we would like to present the status of IHEP site including computing farm, HPC, Grid, data storage ,network and so on.

Tuesday morning / 30

Simulated phishing campaigns at CERN

Author: Sebastian Lopienski¹

¹ *CERN*

Corresponding Author: sebastian.lopienski@cern.ch

Since years, e-mail is one of the main attack vectors that organisations and individuals face. Malicious actors use e-mail messages to run phishing attacks, to distribute malware, and to send around various types of scams. While technical solutions exist to filter out most of such messages, no mechanism can guarantee 100% efficiency. E-mail recipients themselves are the next, crucial layer of protection - but unfortunately, they fall for the various tricks used by attackers way too often.

In order to raise awareness and to educate CERN community, CERN Computer Security Team runs regular simulated phishing campaigns. This talk will discuss the motivation behind this activity, various techniques used, as well as the results and lessons learnt. Finally, CERN campaigns will be compared to those run by other organisations, and to available commercial solutions.

Networking and Security / 31

WLCG/OSG Network Activities, Status and Plans

Authors: Marian Babik¹; Shawn Mc Kee²

¹ *CERN*

² *University of Michigan (US)*

Corresponding Authors: 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 routing. 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 primary areas to cover include the status of and plans for the WLCG/OSG perfSONAR infrastructure, the WLCG Network Throughput Working Group and the activities in the IRIS-HEP and SAND projects.

Networking and Security / 32

Network Functions Virtualisation Working Group Update

Authors: Marian Babik¹; Shawn Mc Kee²

¹ CERN

² University of Michigan (US)

Corresponding Authors: marian.babik@cern.ch, shawn.mckee@cern.ch

As the scale and complexity of the current HEP network grows rapidly, new technologies and platforms are being introduced that greatly extend the capabilities of today's networks. 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. In this talk we'll give an update on the working group's recent activities, updates from sites and R&E network providers as well as plans for the near-term future.

In particular we'll focus on the white paper surveying number of existing technologies and tools that was written by NFW WG and we are going to discuss possible future directions as well as give an overview of the already existing efforts that are part of the Research Networking Technical WG.

Grids, Clouds and Virtualisation / 33

CERN Cloud Infrastructure status update

Author: Thomas George Hartland¹

¹ CERN

Corresponding Author: thomas.george.hartland@cern.ch

CERN's private OpenStack cloud offers more than 300,000 cores to over 3,500 users with services for compute, multiple storage types, baremetal, container clusters, and more. The cloud supports CERN's web and administration services and is the tier 0 site in the WLCG. This update will cover the evolution of the cloud over the past year, and the plans for the upcoming year.

Grids, Clouds and Virtualisation / 34

Kubernetes in the CERN Cloud

Author: Thomas George Hartland¹

¹ CERN

Corresponding Author: thomas.george.hartland@cern.ch

The Magnum component of OpenStack is used to provision container orchestration clusters in the CERN cloud, with Kubernetes being by far the most popular cluster type.

This presentation will look at the new features in Magnum and Kubernetes which make it possible to create highly available Kubernetes clusters that are suitable for hosting critical services.

This includes node groups for splitting nodes across multiple availability zones, software-defined networking loadbalancers for cluster ingress and to enable multi-master clusters, pod and cluster autoscaling for dynamically responding to increased load on a service, and other features.

By combining these features the availability and stability that is possible for a cluster is high enough for us to start running parts of the OpenStack control plane on Kubernetes.

Wednesday morning / 35

EOS storage for Alice O2

Authors: Michal Kamil Simon¹; Andreas Joachim Peters¹; Cristian Contescu¹; Luca Mascetti¹

¹ CERN

Corresponding Authors: michal.simon@cern.ch, cristian.contescu@cern.ch, luca.mascetti@cern.ch, andreas.joachim.peters@cern.ch

In this contribution we report on the ongoing R&D activity aiming at preparing the EOS ALICE O2 storage cluster for the very demanding requirements of Run 3. After the planned upgrades of LHC and ALICE detectors, the ALICE experiment is expected to increase the data-taking rate handled by the online system and then recorded into permanent storage by one order of magnitude. In order to accommodate the data sent by the ALICE Data Acquisition system (aggregated throughput of 100GB/s), the EOS ALICE O2 cluster will be equipped with one of the latest generation storage servers and configured to apply erasure coding to the registered data.

During this talk we are going to present the latest tests and experiments we conducted on the storage server cluster in the context of ALICE O2 project. In particular we give an overview of the aggregate throughput tests with EOS native erasure coding, as well as with client side erasure coding. In addition, we discuss the tuning of the storage servers including testing different kernel versions and firewall settings that enables utilizing the hardware at its nominal speed.

Miscellaneous / 36

The design of networking and computing system for High Energy Photon Source(HEPS)

Author: Fazhi Qi¹

Co-authors: Haifeng Zhao²; Hao Hu²; Haolai Tian²; Hongmei Zhang²; Lu Wang²; Qiulan Huang²; Shan Zeng²

¹ Chinese Academy of Sciences (CN)

² IHEP

Corresponding Authors: huangql@ihep.ac.cn, zengshan@ihep.ac.cn, fazhi.qi@cern.ch, hmzhang@ihep.ac.cn, zhaohf@ihep.ac.cn, tianhl@ihep.ac.cn, wanglu@ihep.ac.cn, huhao@ihep.ac.cn

High Energy Photon Source (HEPS) is the first national high-energy synchrotron radiation light source in Beijing China, and will be ready for users and scientists in 2025. According to the estimated data rates, we predict 30 PB raw experimental data will be produced per month from 14 beamlines at the first stage of HEPS, and the data volume will be even greater after over 90 beamlines are completed at the second stage in the near future.

This report will introduce the design of networking, computing and data service system for HEPS.

Monday evening / 37

BNL Site Report

Author: Tejas Rao^{None}

Corresponding Author: raot@bnl.gov

An update on BNL activities since the Fall 2019 workshop

Thursday morning / 38

CERN's Business Continuity Working Group

Author: Helge Meinhard¹

¹ CERN

Corresponding Author: helge.meinhard@cern.ch

In April 2020 CERN has formed a working group on business continuity. In this presentation, we will describe the mandate and the direction of the group, and will discuss some of the tools used. Areas for potential collaboration with other labs will be discussed as well.

Basic IT Services / 39

CERNphone – CERN's upcoming softphone solution

Author: German Cancio¹

¹ CERN

Corresponding Author: german.cancio.melia@cern.ch

CERNphone is the new softphone-based solution that will gradually be deployed across CERN. Based on open-source components and developments, CERNphone provides mobile and desktop clients and back-end services as a replacement for legacy hard phones and commercial PBX systems. In this contribution, we will describe the architecture and main components of CERNphone, discuss the main challenges for delivering reliable mobile and secure SIP clients, and provide an outlook for its deployment at CERN.

Wednesday morning / 40

FTS: Towards tokens, QoS, archive monitoring and beyond

Author: Mihai Patrascoiu¹**Co-authors:** Edward Karavakis¹; Cristina Voineag²; Carles Garcia Cabot¹; Andrea Manzi¹ *CERN*² *Polytechnic University of Bucharest***Corresponding Authors:** cristina.voineag@cern.ch, carles.garcia.cabot@cern.ch, andrea.manzi@cern.ch, mihai.patrascoiu@cern.ch, edward.karavakis@cern.ch

The File Transfer Service (FTS) is a fundamental component for the LHC experiments, distributing the majority of the LHC data across the WLCG infrastructure. Tightly integrated with experiment frameworks, it has transferred more than 1 billion files and a total of 950 petabytes of data in 2019 alone. With more than 30 experiments using FTS at CERN and outside, it has steadily gained popularity in data-intensive sciences.

Playing a crucial role in data distribution, FTS is constantly evolving in preparation for LHC RUN-3 and forward. With the 2018 participation in the EU-funded project eXtreme Data Cloud (XDC) and continuous involvement within the WLCG DOMA TPC and QoS working groups, a series of developments have been performed in order to meet the requirements of the LHC experiments and community alike.

This presentation will provide a detailed overview of activities carried out for the upcoming 3.10 release, focusing on OpenID Connect (OIDC) token support, QoS functionality, Third Party Copy (TPC) support for XRootD and HTTP protocols, archive monitoring for the new CERN Tape Archive (CTA) system, service scalability improvements and the future direction of FTS.

Tuesday morning / 41

Roadmap for the DNS Load Balancing Service at CERN

Author: Kristian Kouros¹**Co-authors:** Pablo Saiz²; Ignacio Reguero²¹ *VI Trainee*² *CERN***Corresponding Authors:** ignacio.reguero@cern.ch, kristian.kouros@cern.ch, pablo.saiz@cern.ch

This presentation delivers a holistic view of the current and future state of the DNS Load Balancing service at CERN. CERN runs this service to provide the necessary tools for managing the nodes that aliases should present. The service contains three main components:

- 1.An administrative interface, where users can define aliases and their policies.
- 2.A client that runs on the nodes to assess their healthiness.
- 3.An arbiter which, based on the information from the clients, selects the best nodes according to the aliases' policies.

Currently, the service is under migration from Python to Golang. The aim is to improve the parallelism of the service and to decrease the latency of the checks. At the same time, the service is also under migration from a VM-oriented implementation to a Cloud-Native one. The final goal is for this setup to provide a more scalable and reliable system.

Networking and Security / 42

Federated Identity Management at BNL

Author: Shigeki Misawa¹

¹ *Brookhaven National Laboratory (US)*

Corresponding Author: misawa@bnl.gov

We will report on recent activities on integrating Federated Identity Management at the Scientific Data & Computing Center (SDCC).

Thursday evening / 43

An Evaluation of Podman

Authors: Christopher Henry Hollowell¹; Costin Caramarcu¹

¹ *Brookhaven National Laboratory (US)*

Corresponding Authors: costin.caramarcu@cern.ch, hollowec@bnl.gov

Podman is the default container execution platform shipped with RHEL/CentOS 8, and is now also available in RHEL/CentOS/SL 7. It provides a user command line interface that is effectively identical to Docker's, and supports rootless container execution. In this talk we'll give an overview of Podman, our experiences with this software, and a comparison with Singularity and Docker.

Wednesday evening / 44

HPSS migration to IBM tape technologies

Author: Tim Chou¹

¹ *Brookhaven National Laboratory (US)*

Corresponding Author: tchou@bnl.gov

The Scientific Data and Computing Center is migrating part of its 200 PB tape-resident data archive to new high-density robotic libraries. The talk will focus on the product/vendor evaluation process and the complex of decision criteria observed. We will also discuss the potential implications of moving from high-performance to high-density robotic systems for hosting active data archives.

End-user Services, Operating Systems / 45

Development of online application portal

Authors: Wataru Takase¹; Shogo Okada^{None}; Sari Kaneko²; Koichi Murakami²

¹ *High Energy Accelerator Research Organization (JP)*

² *KEK*

Corresponding Authors: kmura@post.kek.jp, wataru.takase@kek.jp, shogo@post.kek.jp

The KEK Computing Research Center provides various IT services, such as email, WiFi network, data analysis system, etc. A user submits paper application forms to apply those services. Moreover, some services require endorsement by a KEK staff. From the Computing Research Center point of view, we receive applications from many users every day. These forms are checked, processed, and handed out among the responsible persons to complete applications.

To improve these circumstances, we have developed an online application portal system named “ccPortal.” This portal allows users to apply new services and modify account information online. The portal also supports the center staff’s online workflow, checking validity, creating an account, DB registration, and sending a completion notice. The ccPortal service was rolled out to KEK staff from the middle of September and gradually opened to all others by November.

We will describe the portal service and share our experience of improving work efficiency in IT service operations.

Miscellaneous / 46

IGWN submit node: Designing EU-based gateway nodes for distributed computing for Virgo, LIGO, Kagra

Author: Mary Hester¹

¹ *Nikhef*

Corresponding Author: maryh@nikhef.nl

While there are a handful of International Gravitational-Wave Observatory Network (IGWN) submit nodes deployed in the US for LIGO, Virgo, Kagra (aka IGWN) data-processing pipelines, Nikhef has worked with LIGO and Virgo collaborators to design a submit node that can be deployed at EU sites. More EU-based submit nodes will allow additional points of entry for IGWN computing resources and make the resources more readily accessible for the scientists. This talk will discuss some of the design choices that made the Nikhef IGWN submit node different from the US nodes and how this fits in with the data lifecycle for IGWN.

End-user Services, Operating Systems / 47

MySQL High Availability in the Database On Demand service

Author: Abel Cabezas Alonso^{None}

Corresponding Author: abel.cabezas.alonso@cern.ch

The DBOD Service is a Database as a Service platform that provides MySQL, PostgreSQL, and InfluxDB database instances to CERN users.

During the last few years, more and more critical services have moved to open source database solutions and are now making use of the DBOD service. As a consequence, high availability is expected for some of these services.

This presentation describes how to achieve High Availability for MySQL databases using ProxySQL. I will also present other open source alternatives for MySQL HA.

Basic IT Services / 48**Towards a redundant, robust, secure and reliable IoT network****Author:** Christoph Merscher¹¹ CERN**Corresponding Author:** christoph.merscher@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.

After having studied various Low-Power Wide Area Network (LPWAN) protocols CERN finally selected the Long Range Wide-Area (LoRa) network as base protocol which marked the start of the establishment of an IoT network at CERN. In order to build a functioning basic infrastructure LoRa gateways and a network server have been selected. With these basic components it was possible to build a simple IoT network. Unfortunately this configuration is not suited for a productive environment as it is not reliable enough.

To improve the reliability and security of the network it is necessary to identify the weaknesses of the technology used as well of those of the current infrastructure. Furthermore, it must be determined which of these vulnerabilities can be improved and which cannot.

CERN is currently running two projects that comprise several thousand end devices and thus put the IoT infrastructure to the test.

Monday morning / 49**KEK Site Report****Author:** Tomoaki Nakamura¹**Co-authors:** Go Iwai²; Koichi Murakami; Takashi Sasaki¹; Soh Suzuki; Wataru Takase¹¹ High Energy Accelerator Research Organization (JP)² KEK**Corresponding Authors:** soh.suzuki@kek.jp, tomoaki.nakamura@kek.jp, wataru.takase@kek.jp, go.iwai@kek.jp, koichi.murakami@kek.jp, takashi.sasaki@kek.jp

KEK Central Computer System (KEKCC) has been upgraded by the full-scale hardware replacement and started the operation in September 2020. In this report, we present the specifications of the new KEKCC with the usage achievement of the previous system last year.

Computing and Batch Services / 50**Operating a production HTCondor cluster: Seamlessly automating maintenance, OS and HTCondor updates with (almost) zero downtimes****Author:** Oliver Freyermuth¹

Co-author: Peter Wienemann¹

¹ *University of Bonn (DE)*

Corresponding Authors: oliver.freyermuth@cern.ch, peter.wienemann@cern.ch

Our HTC cluster using HTCondor has been set up at Bonn University in 2017/2018. All infrastructure is fully puppetised, including the HTCondor configuration.

OS updates are fully automated, and necessary reboots for security patches are scheduled in a staggered fashion backfilling all draining nodes with short jobs to maximize throughput. Additionally, draining can also be scheduled for planned maintenance periods (with optional backfilling) and tasks to be executed before a machine is rebooted or shutdown can be queued. This is combined with a series of automated health checks with large coverage of temporary and long-term machines failures or overloads, and monitoring performed using Zabbix.

In the last year, heterogeneous resources with different I/O capabilities have been integrated and MPI support has been added. All jobs run inside Singularity containers allowing also for interactive, graphical sessions with GPU access.

Combining increasingly heterogeneous resources and different data centre locations in one cluster allows operations with almost zero (full) downtime. During this talk, some examples will be presented on how the automations can be leveraged for different interventions and how the cluster the impact on users and cluster CPU efficiency is minimized.

Basic IT Services / 51

Watching File Storage and Transfers with Elastic Stack at BNL

Author: Matthew Snyder¹

¹ *Brookhaven National Laboratory*

Corresponding Author: msnyder@bnl.gov

The Elastic stack monitors several systems at BNL. At the SDCC two of these are BNLbox and Globus Connect Server. BNLBox- an implementation of nextcloud- is a service for file storing and sharing while Globus is a platform for file transfers. In this talk I will present how we configured Elastic and its components and how they give us insights into client app usage via log ingestion (filebeats) and system monitoring (metricbeats).

Grids, Clouds and Virtualisation / 52

Running event-driven workflows with dCache storage events

Author: Bastien Gounon¹

¹ *CC-IN2P3 / CNRS*

Corresponding Author: bastien.gounon@cc.in2p3.fr

dCache is an open-source distributed storage system for scientific use cases, actively used by large-scale experiments, including within the WLCG community. In 2018, developers started to introduce the concept of storage events within dCache. We are going to present this concept and see how it can be used to trigger automated workflows, with the example of a proof-of-concept implemented at CC-IN2P3 to help with the LSST image simulation campaign. Finally, we will see how events

can be used to transform the way we currently manage infrastructure and workloads when paired with function-as-a-service computing, a new way to deploy highly flexible, automated and smart services.

End-user Services, Operating Systems / 53

A New CMS for a New Decade: Content Management at SDCC

Authors: John Steven De Stefano Jr¹; Christian Lepore²; Louis Pelosi³; Jason Alexander Smith¹

¹ *Brookhaven National Laboratory (US)*

² *Brookhaven National Laboratory*

³ *Brookhaven National Lab*

Corresponding Authors: clepore@bnl.gov, jd@bnl.gov, lpelosi@bnl.gov, smithj4@bnl.gov

The Scientific Data & Computing Center (SDCC) at BNL is migrating its web content management system from Plone to Drupal. This presentation provides a status update on the project. Several technologies were evaluated and tested according to facility and user requirements and specifications.

End-user Services, Operating Systems / 54

The BNLBox service at the SDCC

Author: Ofer Rind^{None}

Corresponding Author: rind@bnl.gov

This presentation covers the BNLBox file sharing service recently introduced at the Scientific Data & Computing Center (SDCC).

Thursday evening / 56

Cloud computing to support experiment online computing from the data center.

Author: Shigeki Misawa¹

¹ *Brookhaven National Laboratory (US)*

Corresponding Author: misawa@bnl.gov

Next generation nuclear and high energy physics experiments are moving filtering and processing tasks, previously done with online resources, to the data center. High resolution imaging systems at light sources and electron microscopes require significant amounts of “online” computing resources to rapidly reconstruct images to allow researchers to make “on the fly” adjustments to running experiments, but physical infrastructure prevents these resources from being co-located with the instrument. Financial constraints will also force these resources to be shared. These trends are blurring the line between online and offline computing. Cloud computing technologies, e.g. SDN, VM/Container provisioning, are possible solutions to the myriad of problems that arise from supporting these experiments from the data center.

Wednesday evening / 57

Scalable High Performance Storage based on Lustre/ZFS over NVMe SSD

Author: Riccardo Veraldi¹

Co-author: Federico Fornari

¹ INFN

Corresponding Authors: federico.fornari@cnafr.infn.it, riccardo.veraldi@cnafr.infn.it

The presentation shows the Storage test bed set up that I implemented in 2018/2019 when working at SLAC

to meet the 2020 LCLS2 data reduction pipeline requirements reaching 100GB/s.

This set up is primarily oriented on high performance rather than high availability and is based on Lustre/ZFS with NVMe SSD storage and EDR Infiniband connection between Object Storage Servers and clients generating traffic.

Tuesday morning / 58

NOTED - identification of data transfer in FTS to understand network traffic.

Author: Joanna Waczynska¹

Co-author: Edoardo Martelli²

¹ Wroclaw University of Science and Technology (PL)

² CERN

Corresponding Authors: joanna.katarzyna.waczynska@cern.ch, edoardo.martelli@cern.ch

NOTED is a project that aims to better exploit WAN bandwidth needed by FTS data transfers.

The main component is the Transfer Broker, which interpreters information coming from FTS to identify large data transfer which could benefit of network optimization. The Transfer Broker then enriches the FTS transfers with network information coming from CRIC, the resource database used by ATLAS and CMS.

In this presentation we describe the work done to identify the large transfers in FTS, how these information are enhanced with the data from CRIC, how we try to estimate the duration of the queue, caused by network overloaded. In the nutshell we will show you how we can explain network traffic by using information from FTS.

Monday morning / 59

ASGC Site Report

Authors: Han-Wei Yen^{None}; Felix Lee¹

¹ ASGC

Corresponding Authors: felix@twgrid.org, han-wei.yen@cern.ch

Report of development at ASGC recently.

End-user Services, Operating Systems / 60**Invenio Based Digital Repositories at BNL****Author:** Carlos Fernando Gamboa¹¹ *Brookhaven National Laboratory (US)***Corresponding Author:** cgamboa@bnl.gov

A Research Digital Management (RDM) repository is a web based service that would provide BNL's scientific community a means to share and preserve their scientific results while making them Findable, Accessible, Interoperable, and Reusable (FAIR). Towards this goal Invenio, an open-source software framework for building large-scale digital repositories, is being used to create a research data management platform called InvenioRDM. An overview of the invenio service at BNL and the status of invenioRDM development will be provided in this talk.

Networking and Security / 61**IPv6-only on WLCG - update from the IPv6 working group****Author:** David Kelsey¹¹ *Science and Technology Facilities Council STFC (GB)***Corresponding Author:** david.kelsey@stfc.ac.uk

The transition of WLCG central and storage services to dual-stack IPv4/IPv6 has gone well, thus enabling the use of IPv6-only CPU resources as agreed by the WLCG Management Board. More and more WLCG data transfers now take place over IPv6. The dual-stack deployment does however result in a networking environment which is much more complex than when using just IPv6. During the last year the HEPiX IPv6 working group has been investigating the removal of the IPv4 protocol in more places. We will present our recent work and future plans.

Wednesday morning / 62**CVMFS service evolution and infrastructure improvements****Authors:** Enrico Bocchi¹; Fabrizio Furano¹; Dan van der Ster¹; Jakub Moscicki¹¹ *CERN***Corresponding Authors:** fabrizio.furano@cern.ch, jakub.moscicki@cern.ch, daniel.vanderster@cern.ch, enrico.bocchi@cern.ch

The Cern VM File System (CVMFS) is a service for fast and reliable software distribution on a global scale. It is capable of delivering scientific software onto physical nodes, virtual machines, and HPC clusters by providing POSIX read-only file system access. Files and metadata are downloaded on-demand by means of HTTP requests and take advantage of aggressive caching on intermediate caches and clients. The choice of the HTTP protocol also enables the exploitation of standard web servers and web caches, including commercially-provided content delivery networks.

CVMFS is widely adopted in the HEP community for the distribution of production software, integration builds, auxiliary datasets, and has recently introduced new capabilities to broaden its scope of application. As a prime example, it implements extensive support for container images with DUCC (Daemon that Unpacks Container images into CVMFS), a specialized component that unpacks container images and publishes their extracted form on a repository, and tight integration

with container runtimes, making published container images usable by widely-adopted container platforms (e.g., Singularity, Docker, Kubernetes). Such functionality provides an alternative to traditional container registries (e.g., Docker Hub, GitLab Container Registry) and makes the distribution of container images more efficient by leveraging on file-based deduplication and on-demand caching provided by CVMFS.

CVMFS at CERN has also been subject to several infrastructural updates. The repository storage for Stratum Zero servers is now hosted on the Ceph-based S3 service, which provides a relevant performance improvement with respect to block storage provided via Cinder volumes. Also, content distribution to clients has been made more resilient by deploying dedicated caches for sets of repositories, which greatly reduces the problem of interference across repositories and cache thrashing phenomena.

Wednesday morning / 63

HIFIS backbone transfer service: FTS for everyone

Authors: Tim Wetzel¹; Paul Millar^{None}; Uwe Jandt²; Patrick Fuhrmann^{None}

Co-authors: Stefan Helmert³; Tobias Huste³; Guido Juckeland³

¹ *Deutsches Elektronen-Synchrotron DESY*

² *Deutsches Elektronensynchrotron (DESY)*

³ *Helmholtz-Zentrum Dresden-Rossendorf*

Corresponding Authors: paul.millar@desy.de, tim.wetzel@desy.de, uwe.jandt@desy.de, s.helmert@hzdr.de, patrick.fuhrmann@desy.de, g.juckeland@hzdr.de, t.huste@hzdr.de

The German Helmholtz Association (HGF) encompasses 19 research institutes distributed all over Germany, covering a wide variety of research topics ranging from particle and material physics over cancer research to marine biology. In order to stimulate collaborations between different centres, the HGF established so-called incubator platforms. Two of those platforms, relevant for this presentation, are the Helmholtz Artificial Intelligence Cooperation Unit (Helmholtz AI) and Helmholtz Federated IT Services (HIFIS). While Helmholtz AI was established to connect domain scientists and AI experts for a stronger adoption of AI solutions for increasingly complex research tasks, HIFIS targets the exploitation of synergy effects in federated IT services offered by the different HGF centres.

During the ongoing ramp-up phase of both platforms, specific use cases of interdisciplinary research are arising and showing that there is a definitive need to transfer a significant amount of large data sets between centres. This primarily results from the fact that the currently used AI solutions are trained on specific data sets and that the processing of that data is sensitive to network latencies. Consequently, remote data access is less efficient in those cases and consequently data needs to be transferred from the domain scientists' home institutions to the AI experts' location, where the model training is taking place.

In order to cater to those needs, a file transfer service is being established by HIFIS for convenient and automated data transfer between the sites of those interdisciplinary research groups. After evaluating competing solutions like Globus Online and Onedata we agreed to go for FTS3 for reasons we will elaborate on during the presentation. FTS3 is a file transfer service that can commission data transfers between storage endpoints and has been developed at CERN for the transfer of WLCG research data between CERN and several hundred LHC Tier centres. Those endpoints need to be able to communicate with each other as well as with FTS, using a third-party copy (TPC) extension of the HTTP protocol. In order to facilitate an easy installation of endpoints, which are not the known WLCG storage systems (like dCache, DPM and EOS), we provide an Apache web server extension that complies with the needs of FTS3 and can thus act as a storage endpoint for data transfers via HTTP-TPC.

We will present the necessary prerequisites for such an endpoint, its configuration details as well as the modifications applied to the Apache modules. Adding to that, we will present insights into the performance and reliability of the TPC data transfers measured with test data sets.

Wednesday morning / 64

XCache: exploring cache storage technology

Author: Paul Musset¹

¹ *IN2P3*

Corresponding Author: paul.musset@cern.ch

Distributed computing involves processing data from far remote site. We explore a new type of cache, XCache, developed within XROOTD, and explore its capability to improve the access of the jobs to remote storages. The proof-of-concept is realized within the ESCAPE european project and measurement of the performances are done using remote sites. The presentation will develop the configuration of the cache and the different tests done to evaluate the efficiency of the cache for the use cases we have at CC-IN2P3.

Basic IT Services / 65

MALT Project

Corresponding Author: maria.barroso.lopez@cern.ch

End-user Services, Operating Systems / 66

CERN Appstore for BYOD devices

Corresponding Author: tamas.bato@cern.ch

The number of BYOD devices at CERN is growing and there is interest in moving from a centrally-managed model to a distributed model where users are responsible for their own devices. Following this strategy, new tools are needed to distribute and - in case of licensed software - also manage licences for applications provided by CERN. Available open source and commercial solutions were previously analyzed and none of them proved to be a good fit for CERN academic environment. Therefore, we started to develop a system that can integrate existing open source solutions and provide desired functionality for multiple platforms, both mobile and desktop.

The functionality of the tool has been growing since the project started. The list of supported platforms has been growing and very recently the integration with Flatpak has been added in order to provide applications not just on Windows, but on Linux as well. The number of features provided is also growing and the UX is enhanced along the way.

This talk will describe the architecture and design decisions made to develop a platform-independent, modern, maintainable and extensible system for future software distribution at CERN.

Miscellaneous / 67

Workshop Wrap-Up

Corresponding Author: peter.van.der.reest@desy.de

Board Meeting / 68

HEPiX board meeting

Corresponding Authors: tony@bnl.gov, peter.van.der.reest@desy.de

Tuesday evening / 69

HTCondor 2020 workshop report

Author: Jose Flix Molina¹

¹ *Centro de Investigaciones Energéticas Medioambientales y Tecnológicas*

Corresponding Author: jose.flix.molina@cern.ch

The “European” HTCondor workshop was held from 21 to 25 September as a purely virtual event. We will report on topics covered and observations made.

End-user Services, Operating Systems / 70

Experiences from online Workshops and Conferences

Author: Graeme A Stewart¹

¹ *CERN*

Corresponding Author: graeme.andrew.stewart@cern.ch

The talk will meld the experience of the HSF Workshop, PyHEP and ICHEP and discuss what proved to work best and things that didn't work as well as hoped.

Wednesday evening / 71

dCache in the cloud environment

Corresponding Author: tigran.mkrtchyan@desy.de

Wednesday morning / 72

Production deployment of the CERN Tape Archive (CTA) for Atlas

Corresponding Author: julien.leduc@cern.ch

Tuesday morning / 73

group photo morning

Networking and Security / 74

group photo evening

Thursday morning / 76

Security Update

Corresponding Author: romain.wartel@cern.ch

HEPiX might have been hacked - stand by for breaking news.

Thursday evening / 77

Security update

Corresponding Author: romain.wartel@cern.ch