HEPiX Spring 2024 Workshop

Report of Contributions

https://indico.cern.ch/e/hepix-spring2024

Contribution ID: **1** Type: **not specified**

CEA-IRFU

Monday, 15 April 2024 09:00 (20 minutes)

Speaker release

Desired slot length

Primary author: FORMICA, Andrea (CEA Irfu/Lilas)

Presenters: FORMICA, Andrea (Université Paris-Saclay (FR)); FORMICA, Andrea (CEA Irfu/Lilas)

Session Classification: Welcome

Track Classification: Miscellaneous

Contribution ID: **2** Type: **not specified**

Workshop logistics

Monday, 15 April 2024 09:20 (10 minutes)

Speaker release

Desired slot length

Primary authors: MAGROUNE, Imed (CEA/DRF/IRFU/DEDIP//LIS); SURGET, Joël; Dr HONORE, Pierre-Francois

Presenters: MAGROUNE, Imed (CEA/DRF/IRFU/DEDIP//LIS); SURGET, Joël; Dr HONORE, Pierre-Francois

Session Classification: Welcome

Track Classification: Miscellaneous

Contribution ID: **3** Type: **not specified**

Wrap-up

Friday, 19 April 2024 12:00 (30 minutes)

Presenter: MEINHARD, Helge (CERN)

Session Classification: Wrap-up

Track Classification: Miscellaneous

Contribution ID: **4** Type: **not specified**

The state of AlmaLinux

Monday, 15 April 2024 14:00 (20 minutes)

AlmaLinux has been chosen by many across the world as the replacement for CentOS Linux, but there is still a lot of confusion around AlmaLinux's governance and build pipeline. Given AlmaLinux's newness and the instability in the greater enterprise Linux ecosystem, a strong understanding of how and where AlmaLinux OS started, where AlmaLinux it is today, and where we expect to go in the future is critical to many of our users. This talk will address all of that and more.

Desired slot length

30-45 minutes, as part of the linux ecosystem talk.

Speaker release

Yes

Primary author: VASQUEZ, benny (AlmaLinux OS Foundation)

Presenter: VASQUEZ, benny (AlmaLinux OS Foundation)

Session Classification: Linux future

Track Classification: Operating Systems, Cloud & Virtualisation, Grids

Contribution ID: **5** Type: **not specified**

Streamlining solution for the hyper-converged cloud infrastructure at CSCS

Thursday, 18 April 2024 11:10 (25 minutes)

This presentation provides a detailed overview of the hyper-converged cloud infrastructure implemented at the Swiss National Supercomputing Centre (CSCS). The main objective is to provide a detailed overview of the integration between Kubernetes (RKE2) and ArgoCD, with Rancher acting as a central tool for managing and deploying RKE2 clusters infrastructure-wide.

Rancher is used for direct deployment on MAAS-managed nodes, as well as HPC (High-Performance Computing) nodes designed for high-intensity workloads. In addition, Harvester orchestrates Kubernetes distributions for virtual clusters, improving flexibility and simplifying orchestration on the platform.

ArgoCD plays a key role in automating deployment processes and ensuring consistency between different environments, enabling continuous delivery. The integration of Kubernetes, ArgoCD, Rancher, Harvester and Terraform forms the basis of a hyper-converged, scalable and adaptable cloud infrastructure.

This case study provides information on the architecture, deployment workflows and operational benefits of this approach.

Desired slot length

20

Speaker release

Yes

Primary author: Mr CONCIATORE, Dino (CSCS (Swiss National Supercomputing Centre))

Presenter: Mr CONCIATORE, Dino (CSCS (Swiss National Supercomputing Centre))

Session Classification: Operating systems, clouds, virtualisation, grids

Track Classification: Operating Systems, Cloud & Virtualisation, Grids

Contribution ID: **6** Type: **not specified**

CC-IN2P3 Site Report

Monday, 15 April 2024 09:45 (15 minutes)

We will present a follow-up on the activities on-going at the CC-IN2P3 since the last site report done in Fall 2022.

Desired slot length

Speaker release

Yes

Primary authors: PUEL, Mattieu; Dr GADRAT, Sebastien (CCIN2P3 - Centre de Calcul (FR))

Presenter: PUEL, Mattieu

Session Classification: Site reports

Track Classification: Site Reports

Contribution ID: **7** Type: **not specified**

KEK Site Report

Monday, 15 April 2024 10:00 (15 minutes)

The KEK Central Computer System (KEKCC) provides large-scale computing resources, including Grid computing systems and essential IT services, to support many research activities in KEK. We will report on the current status of KEKCC in this presentation.

Desired slot length

Speaker release

Yes

Primary author: KISHIMOTO, Tomoe

Presenter: KISHIMOTO, Tomoe

Session Classification: Site reports

Track Classification: Site Reports

Contribution ID: **8** Type: **not specified**

SiteReport IRFU/GRIF-IRFU

Monday, 15 April 2024 09:30 (15 minutes)

We present an update of the changes at our site since the last report. Advancements, developments, roadblocks and achievements made concerning various aspects including: WLCG, Unix, Windows, Infrastructure, will be presented.

Desired slot length

10

Speaker release

Yes

Primary author: FERRY, Sophie Catherine (Université Paris-Saclay (FR))

Co-authors: SURGET, Joël; Dr HONORE, Pierre-Francois

Presenter: FERRY, Sophie Catherine (Université Paris-Saclay (FR))

Session Classification: Site reports

Track Classification: Site Reports

Contribution ID: **9** Type: **not specified**

Monitoring the Green IT Cube

Tuesday, 16 April 2024 11:10 (25 minutes)

Managing a data center poses multifaceted challenges, with monitoring emerging as a pivotal aspect for ensuring stability and service quality assurance. Since the first implementation of a monitoring system in the Green IT Cube data center at GSI in 2016, comprising RRDtool and Ganglia, ongoing efforts have been dedicated to enhancing monitoring capabilities. The initial system's tight integration between data storage and visualization layers, along with missing essential features, limited its adaptability to diverse use cases. Consequently, a contemporary system architecture was developed, featuring modular components for data collection, storage, and visualization, leveraging Prometheus, InfluxDB, and Grafana technologies. This talk will address the usage and recent enhancement of the new monitoring components in the Green IT Cube at the GSI.

Desired slot length

Speaker release

Yes

Primary authors: BRUINS, Daniel (GSI); IANNETTI, Gabriele (GSI)

Presenters: BRUINS, Daniel (GSI); IANNETTI, Gabriele (GSI)

Session Classification: Basic and end-user IT services

Track Classification: Basic and End-User IT Services

Contribution ID: **10** Type: **not specified**

GSI Site Report

Monday, 15 April 2024 10:15 (15 minutes)

Recent developments at GSI IT

Desired slot length

Speaker release

Yes

Primary author: Mr HUHN, Christopher

Presenter: Mr HUHN, Christopher

Session Classification: Site reports

Track Classification: Site Reports

Contribution ID: **11** Type: **not specified**

HEPiX Benchmarking Working Group Report

Tuesday, 16 April 2024 15:40 (25 minutes)

The adoption of HEPScore23 as replacement of HS06 in April 2023 marked a significant milestone for the WLCG community. After one year since that change, we conduct a thorough review of the experience, lessons learned, and areas for improvement. In addition, triggered by the community feedback and demand, the Benchmarking WG has started a new development effort to expand the Benchmark Suite with modules that can measure server utilisation metrics (load, frequency, I/O, power consumption) during the execution of the HEPScore benchmark. This activity has open a new potential of study that we will share in this report. Last but not least the work to include GPU workloads in the catalogue of available workloads will be presented.

Desired slot length

20

Speaker release

Yes

Primary author: GIORDANO, Domenico (CERN)

Presenter: GIORDANO, Domenico (CERN)

Session Classification: Computing and batch services

Track Classification: Computing & Batch Services

Contribution ID: **12** Type: **not specified**

CernVM-FS 2.11: Lessons learnt from large-scale data processing at Jump Trading

Tuesday, 16 April 2024 09:00 (25 minutes)

The CernVM File System (CVMFS) provides the software distribution backbone for High Energy and Nuclear Physics experiments and many other scientific communities in the form of a globally available, shared, read-only filesystem. However recently, CVMFS has found major adoption beyond academia: in particular Jump Trading, an algorithmic and high-frequency trading firm, now uses CVMFS for software and data distribution at scales that surpass any prior usage, pushing the software to its limits. In this talk, we present the latest enhancements that enable this use-case, thanks to the support and contributions of Jump Trading and the close collaboration of their engineers with the CVMFS development team. Concretely, we report on the operational improvements that allow Jump Trading to deploy the latest CVMFS release for their highly parallelized workloads, and current plans for further performance gains.

Desired slot length

Speaker release

Yes

Primary authors: BLOMER, Jakob (CERN); PROMBERGER, Laura (CERN); HARVEY, Matt (Jump Trading); NAGHIBI, Reza (Jump Trading); VOLKL, Valentin (CERN)

Presenters: HARVEY, Matt (Jump Trading); VOLKL, Valentin (CERN)

Session Classification: Storage and file systems

Track Classification: Storage & Filesystems

Contribution ID: **13** Type: **not specified**

Towards a DevSecOps approach in the CICD pipelines for the INFN Information System apps

Tuesday, 16 April 2024 11:35 (25 minutes)

The INFN Information System project was established in 2001 with the aim of computerizing and standardizing the administrative processes of the Institute and gradually moving towards dematerialization and digitization of documents. During these two decades the aim of the project has been accomplished by a series of web applications (what we call sysinfo apps) serving INFN researchers, technologists as well as administratives and human resource teams for activities like business trips, buying computing facilities, managing the recruitment process, accounting.

Those sysinfo apps are developed by a Development team and operated by a Platform team that manages also the underlying infrastructure as well as all the processes to enable the Development team in their activities. In the last four years both teams have been involved in the re-architecting of those apps towards the so-called microservices architecture. One of the main effort has been put in place to rethink the Continuous Integration and Continuous Delivery/Deployment (CICD) pipelines towards a DevSecOps approach based on three guiding principles:

1. Platform and Development teams must agree on a contract (formalized both at project repository and CICD infrastructure levels) on how to develop, build, test, deliver and deploy software.
2. Platform team is the owner to the governance and security of the CICD pipelines.
3. Introduce shift left testing stages as early as possible in the software development life cycle.

In this presentation we will go through the implementation of the aforementioned guiding principles, describing how we leveraged the GitLab-CI pipeline profiles/templates concept to provide end-to-end CICD workflows applying to well defined project's structures and languages. Moreover, focusing on the Continuous Deployment side, we will describe the GitOps approach, driven by the ArgoCD tool, to deploy microservices in our Kubernetes clusters.

Finally we will highlight how moving towards this DevSecOps approach allows us to keep a baseline of governance and security with the agile development while dealing with the challenge of migrating, at first, and evolving the INFN sysinfo apps in microservices architecture and container orchestration contexts.

Desired slot length

20

Speaker release

Yes

Primary author: MISURELLI, Giuseppe (INFN)

Co-authors: BOVINA, Stefano (INFN); GUIZZUNTI, Guido (INFN)

Presenter: MISURELLI, Giuseppe (INFN)

Session Classification: Basic and end-user IT services

Track Classification: Basic and End-User IT Services

Contribution ID: **14** Type: **not specified**

HEPiX Technology Watch Working Group Report

Thursday, 18 April 2024 10:15 (25 minutes)

The Technology Watch Working Group, established in 2018 to take a close look at the evolution of the technology relevant to HEP computing, has resumed its activities after a long pause. In this first official report after such pause, we describe our goals, how the group is organized and the first results of our work.

Desired slot length

20

Speaker release

Yes

Primary authors: CHIERICI, Andrea (INFN-CNAF); Dr SCIABÀ, Andrea (CERN); MISAWA, Shigeki (Brookhaven National Laboratory (US))

Presenter: Dr SCIABÀ, Andrea (CERN)

Session Classification: IT facilities, business continuity and Green IT

Track Classification: IT Facilities, Business Continuity and Green IT

Contribution ID: **15** Type: **not specified**

ASGC Site Report

Monday, 15 April 2024 11:00 (15 minutes)

Update of the WLCG and Scientific computing services, technology and resource at ASGC.

Desired slot length

12

Speaker release

Yes

Primary author: Mr YEN, Eric (ASGC)

Presenter: Mr YEN, Eric (ASGC)

Session Classification: Site reports

Track Classification: Site Reports

Contribution ID: **16** Type: **not specified**

Streaming Science Globally: CERN's Live Streaming Service

Tuesday, 16 April 2024 12:00 (25 minutes)

In the last 20 years, CERN's Live Streaming service [1] has been a pivotal communication tool connecting CERN users and the High Energy Physics (HEP) community in real time. From its initial stages, employing Real Media technologies and Flash, to its present state, integrating cutting-edge technologies like HTTP Live Streaming (HLS), the service has been instrumental in fostering global scientific collaboration.

In this presentation, we will provide a review of the service's evolution, examining its architecture, employed technologies, and more. Our discussion will extend to the future of the service, addressing the challenges we anticipate.

Furthermore, we will provide detailed insight on the integration of the live streaming service with the transcoding and web lecture services. This integration enables users to watch already finished streams, improving their experience with cutting-edge features such as closed captions, composite views, and multi-quality videos.

[1] https://live.cern.ch

Desired slot length

Speaker release

Yes

Primary author: FERNANDEZ SANCHEZ, Rene (CERN)

Presenter: FERNANDEZ SANCHEZ, Rene (CERN)

Session Classification: Basic and end-user IT services

Track Classification: Basic and End-User IT Services

Contribution ID: **17** Type: **not specified**

RAL Site Report

Monday, 15 April 2024 11:15 (15 minutes)

Updates from RAL

Desired slot length

Speaker release

Yes

Primary author: BLY, Martin (STFC-RAL)

Presenter: BLY, Martin (STFC-RAL)

Session Classification: Site reports

Track Classification: Site Reports

Contribution ID: **18** Type: **not specified**

Enabling LHC Run 3 data storage workflows at CERN

Tuesday, 16 April 2024 09:25 (25 minutes)

CERN Storage and Data Management group is responsible for ensuring that all data produced by physics experiments at CERN is safely stored and reliably accessible by the user community. 2023 Run-3 and especially the Heavy Ion Run have pushed further the previous records in terms of data volume and transfer rates delivered by the main LHC experiments. The targets anticipated by the data management coordinators were successfully accommodated by the main storage solutions provided by the Storage group, namely: EOS, CTA (tape storage) and FTS (data transfer orchestration). The EOS service is the main entry point for all data acquisition workflows and has demonstrated reliable operation and excellent peak performance throughout Run 3. In practice, this means that target rates of 10-20 GB/s where regularly surpassed with peaks of 25-30 GB/s for CMS, ATLAS and ALICE experiments.

Distribution of all this data as well as storing the custodial copy on tape required the orchestration capabilities of FTS that successfully met user's expectations and also ensured a good utilization of the tape infrastructure. The most demanding use-case was represented by ALICEO2 which achieved data rates of over 150GB/s using erasure encoded layouts and its own workflow for data distribution. In this presentation, we go over the archived rates and general performance for both the disk and the tape services. Looking towards the restart of Run-3, we detail foreseen challenges and lessons learned during this exceptional period of data taking.

Desired slot length

15

Speaker release

Yes

Primary author: SINDRILARU, Elvin Alin (CERN)

Presenter: SINDRILARU, Elvin Alin (CERN)

Session Classification: Storage and file systems

Track Classification: Storage & Filesystems

Contribution ID: **19** Type: **not specified**

CERN site report

Monday, 15 April 2024 11:30 (15 minutes)

News from CERN since the last HEPiX workshop. This talk gives a general update from services in the CERN IT department.

Desired slot length

12

Speaker release

Yes

Primary author: SINDRILARU, Elvin Alin (CERN)

Presenter: SINDRILARU, Elvin Alin (CERN)

Session Classification: Site reports

Track Classification: Site Reports

Contribution ID: **20** Type: **not specified**

Prototyping an Analysis Facility at CERN

Friday, 19 April 2024 09:25 (25 minutes)

In recent times, experiment analysis frameworks and physics data formats of the LHC experiments have been evolving in a direction that makes interactive analysis with short turnaround times much more feasible. In parallel, many sites have set up Analysis Facilities to provide users with tools and interfaces to computing and storage that are optimised for interactive analysis. At CERN we conducted detailed performance and scalability measurements using distributed analysis workloads to assess the readiness of the local computing and storage infrastructure, and more recently we launched a pilot for an analysis facility where real users will be able to run, in order to collect information about how CERN could provide such service. This prototype is based on well proven services, like SWAN, DASK, HTCondor and EOS. In this contribution we will show the results obtained so far.

Desired slot length

Speaker release

Yes

Primary authors: Dr SCIABÀ, Andrea (CERN); JONES, Ben (CERN); TEJEDOR SAAVEDRA, Enric (CERN); MCCANCE, Gavin (CERN); SCHULZ, Markus (CERN)

Presenter: Dr SCIABÀ, Andrea (CERN)

Session Classification: Basic and end-user IT services

Track Classification: Basic and End-User IT Services

Contribution ID: **22** Type: **not specified**

Future of Hard Drives

Tuesday, 16 April 2024 10:15 (25 minutes)

Most data center data is stored on hard drives. But can they compete in the future, and what is the role and future of hard drives and the different hard drive technologies available? The challenge is efficiently scaling storage infrastructure while optimizing for write/read performance, TCO, and sustainability goals.

In this session, we will explore how areal density and the latest hard drive technology deliver on Scale, TCO, and Sustainability to manage the data explosion. We will also discuss key technology features and hard drive innovations that mark an inflection point for hard drive storage.

Until now, it was not possible to increase capacity without increasing form factor and using more resources. Culminating in a breakthrough collection of Nobel Prize-winning nanoscale technologies, Mozaic 3+™ is a new hard drive platform that incorporates the unique implementation of Heat Assisted Magnetic Recording (HAMR).
Its high magnetic coercivity media overcomes magnetic instability to deliver unprecedented areal density of 3TB per platter (4TB+ and 5TB+ in the coming years) and capacity points of 30TB and beyond.
Use it like a regular hard disk drive; written data will never fluctuate—it can only be rewritten with its plasmonic writer, ensuring data durability and achievability on hard drives.
This session will also include a live demo of the latest Mozaic 3+ hard drives.

Desired slot length

Speaker release

Yes

Primary author: Mr BERGMANN, Hugo

Presenter: Mr BERGMANN, Hugo

Session Classification: Storage and file systems

Track Classification: Storage & Filesystems

Contribution ID: **23** Type: **not specified**

The Evolution of Storage Media in Data Centers: Assessing the Future of Flash/SSD, Hard Drives, and Tape Storage

Wednesday, 17 April 2024 10:15 (25 minutes)

Data centers are at the forefront of managing vast amounts of data, relying on three primary storage mediums: Flash/SSD, Hard Drives, and Tape. This presentation deals with the special characteristics of the individual technologies and examines the question of whether the forthcoming advances will lead to one technology being displaced by the other. Central to our discussion is the examination of the distinct roles of these storage media and the potential for transformative technological shifts on the horizon.

Key topics to be explored include:

- A high-level overview and future outlook of Flash, Hard Drive, and Tape technologies.
- Understanding the differences, limitations and expected future development of the individual storage media.
- Drawing parallels from the consumer realm, where SSDs have replaced Hard Drives, prompts the question: What is the likelihood of similar transitions occurring within data center environments?

Join us as we navigate the complex landscape of storage media in data centers, offering insights into the dynamic interplay between technology, functionality, and the evolving needs of modern data management.

Desired slot length

Speaker release

Yes

Primary author: BERGMANN, Hugo (Seagate Technology)

Presenter: BERGMANN, Hugo (Seagate Technology)

Session Classification: Storage and file systems

Track Classification: Storage & Filesystems

Contribution ID: **24** Type: **not specified**

Navigating the Intersection of Data Growth, Sustainability, and Innovation in Data Centers

Tuesday, 16 April 2024 16:55 (25 minutes)

In an era defined by the exponential rise of artificial intelligence and data analytics, the value of data is more valuable than ever, propelling the expansion of data centers to accommodate vast datasets. However, this growth comes with a sobering reality: the significant energy consumption of data centers, often rivaling that of entire nations. The need to create a sustainable and scalable storage infrastructure comes up against three critical forces: the unstoppable growth in data generation, the constraints of limited data center space and the scarcity of resources.

This presentation will explore a diverse array of solutions and innovations poised to address these pressing challenges:

- Carbon Emissions: Evaluating different Storage Media (Flash/Hard Drive/Tape) across their full lifecycle, emphasizing the importance of considering not only the usage phase but also the entire lifecycle.
- Areal Density Revolution: advancement in areal density thus the storage efficiency and capacity.
- Unveiling ADAPT + ADR: Advanced Distributed Allocation Protection Technology (Distributed Erasure Code) with Autonomous Drive Repair. Enables a transformative storage architecture offering unprecedented efficiency gains in power consumption, resource utilization, and e-waste reduction. Additionally, it reveals substantial additional savings in compute and networking resources while increasing data durability.
- SED as a Catalyst for Circular Economy: Exploring the pivotal role of Self-Encrypting Drives (SEDs) in fostering a circular hard drive economy, driving sustainability across the entire storage lifecycle.
- Seagate Circularity Program: Highlighting initiatives aimed at advancing circularity and minimizing the embodied carbon emissions underscoring the importance of holistic lifecycle considerations.
- MACH-2 Optimization: Introducing MACH-2, an innovative solution optimized for cost, power efficiency, and density, offering a compelling pathway towards more sustainable data storage infrastructure.

Join us as we navigate the complex terrain of data growth, sustainability imperatives, and technological innovation, forging a path towards a more sustainable and resilient future for data centers.

Desired slot length

Speaker release

Yes

Primary author: BERGMANN, Hugo (Seagate Technology)

Presenter: BERGMANN, Hugo (Seagate Technology)

Session Classification: IT facilities, business continuity and Green IT

Track Classification: IT Facilities, Business Continuity and Green IT

Contribution ID: **26** Type: **not specified**

Cloud Resilience Unleashed: A Dual-Site Approach for HL-LHC era

Thursday, 18 April 2024 11:35 (25 minutes)

In preparation of the increasing computing needs of the HL-LHC, the IT department has built a new datacenter in Prevessin. These additional capacity will also enable IT services to prepare their Business Continuity and Disaster recovery plans.
With those two in mind, the cloud service has prepared a new deployment fully decoupled from the existing in Meyrin. This new setup allows users to operate with both datacentres seamlessly. The solution presented allows them to use resources on any of the sites even when there is a major outage in the other. This presentation will give an overview of the architecture, main differences and synchronization mechanisms between the deployments.

Desired slot length

Speaker release

Yes

Primary author: FAILING, Daniel (CERN)

Presenter: FAILING, Daniel (CERN)

Session Classification: Operating systems, clouds, virtualisation, grids

Track Classification: Operating Systems, Cloud & Virtualisation, Grids

Contribution ID: **27** Type: **not specified**

Wiring the Future: Open Virtual Networking and Beyond

Thursday, 18 April 2024 13:30 (25 minutes)

The cloud service provides resources to the whole CERN community in two datacentres (Meyrin and Prevessin). The deployment of the new datacenter in Prevessin allowed us to reconsider all the design choices we made for Meyrin. In the networking area, we are increasing the flexibility and adding even more options to users by offering Software Defined Networking. In this talk, we will explain the inclusion of Open Virtual Networking (OVN) in the portfolio that allow users to have access to floating IPs, private networks, security groups at scale in the new Prevessin Datacentre.

Desired slot length

Speaker release

Yes

Primary author: FAILING, Daniel (CERN)

Presenter: FAILING, Daniel (CERN)

Session Classification: Operating systems, clouds, virtualisation, grids

Track Classification: Operating Systems, Cloud & Virtualisation, Grids

Contribution ID: **28** Type: **not specified**

Ready for Windows 11 in your endpoint device park?

Thursday, 18 April 2024 13:55 (25 minutes)

Windows 10 is dead, long live Windows 11! Device compatibility, hardware replacement plans, strategies for upgrade campaigns, privacy are all discussed and illustated with examples based on 10000 CERN PCs. Generative AI (Copilot) is also coming to the OS layer. Are you ready?

Desired slot length

Speaker release

Yes

Primary author: DELLABELLA, Sebastien (CERN)

Co-authors: KWIATEK, Michal (CERN); FIROOZBAKHT, Siavas (CERN)

Presenter: DELLABELLA, Sebastien (CERN)

Session Classification: Operating systems, clouds, virtualisation, grids

Track Classification: Operating Systems, Cloud & Virtualisation, Grids

Contribution ID: **29** Type: **not specified**

CERNBox: Community Distribution for Analysis Facilities

Wednesday, 17 April 2024 09:25 (25 minutes)

CERNBox is an innovative scientific collaboration platform, built using solely open-source components to meet the unique requirements of scientific workflows. Used at CERN for the last decade, the service satisfies the 35K users at CERN and seamlessly integrates with batch farms and Jupyter-based services.

Following the presentations given at the CS3 Workshop 2024[1] and CERN Storage Day[2], as well as the BoF session at the WLCG-HSF Workshop 2023[3], there has been a surge in interest from Tier 1 and Tier 2 Data Centres and other scientific institutions, all eager to deploy CERNBox within their own ecosystems. In this talk, we'll delve into the core technology that powers CERNBox —Reva —and demonstrate how you can install your own CERNBox system with either EOS or CephFS as the storage backend. We conclude with remarks on the role of sync and share systems such as CERNBox in the Analysis Facilities landscape.

[1] https://indico.cern.ch/event/1332413/contributions/5740225/
[2] https://indico.cern.ch/event/1353101/contributions/5805537/
[3] https://indico.cern.ch/event/1230126/sessions/492063/#20230506

Desired slot length

Speaker release

Yes

Primary authors: LO PRESTI, Giuseppe (CERN); GONZALEZ LABRADOR, Hugo (CERN); CASTRO, Diogo (CERN)

Presenter: LO PRESTI, Giuseppe (CERN)

Session Classification: Storage and file systems

Track Classification: Storage & Filesystems

Contribution ID: **30** Type: **not specified**

Re-designing the Remote Desktop Service: A solution for remote desktop access management at CERN

Friday, 19 April 2024 09:50 (25 minutes)

More than 2000 daily users use the remote desktop access service from outside CERN, supporting remote work capabilities in the organization.
To improve the operation of the service, a new self-service solution for remote desktop access was launched in 2023 that empower users to manage their devices set up for remote access. The solution includes parallelization, caching, and automated mechanisms of synchronization in order to enhance user experience and improve performance. This presentation will give an overview of the architecture, the mechanisms and the problems found and how they were solved to create a robust system for handling remote desktop connections in the organization.

Desired slot length

20

Speaker release

Yes

Primary author: STOJKOVIC, Petar

Co-authors: REY REGULEZ, Mario (CERN); MARTIN ZAMORA, Pablo (CERN)

Presenter: STOJKOVIC, Petar

Session Classification: Basic and end-user IT services

Track Classification: Basic and End-User IT Services

Contribution ID: **31** Type: **not specified**

Carbon life cycle modelling of scientific computing

Thursday, 18 April 2024 09:00 (25 minutes)

By modelling the life cycle emissions for a given unit of scientific computing under various scenarios of hardware replacement and computing facilities (including the emissions from the local power generation mix), we can find optimal computing hardware replacement cycles in order to minimize carbon emissions.

The majority of this work was presented at ISGC on March 28th: https://indico4.twgrid.org/event/33/contributions/1419/ but we intend to improve it based on the lively audience discussion the ISGC presentation generated.

Desired slot length

Speaker release

Yes

Primary authors: WADENSTEIN, Erik Mattias (University of Umeå (SE)); Prof. VANDERBAUWHEDE, Wim (University of Glasgow)

Presenter: WADENSTEIN, Erik Mattias (University of Umeå (SE))

Session Classification: IT facilities, business continuity and Green IT

Track Classification: IT Facilities, Business Continuity and Green IT

Contribution ID: **32** Type: **not specified**

The intelligent operation and maintenance system of the IHEP computing platform

Friday, 19 April 2024 10:15 (25 minutes)

Based on extensive experience in system maintenance and advanced artificial intelligence technology, we have designed the IHEP computing platform's intelligent operations and maintenance system. Its primary goal is to ensure optimal utilization and efficiency of computing resources.
This system automatically detects user jobs that cause anomalies in computing services and dynamically adjusts their available resources in real time.
Utilizing AI algorithms, it swiftly conducts fast, near real-time analysis of the file system's operational status and logs, identifying potential users and their process names that may be triggering anomalies.
After querying the computing node where the suspected abnormal job is located through the job scheduler, the system utilizes AI algorithms to conduct real-time analysis of the job to determine whether its behavior is causing excessive system load. Once confirmed, the system notifies the job scheduler and file system to limit the number of user job operations and the total I/O volume.
This system is employed for comprehensive monitoring and intelligent operations management of the computing platform. It dynamically adjusts the scale of available resources for users based on the overall situation of the computing platform, ensuring fair and efficient data processing for all users.

Desired slot length

Speaker release

Yes

Primary authors: CHENG, Yaosong (Institute of High Energy Physics Chinese Academy of Sciences, IHEP); SHIJY, 石京燕

Presenter: CHENG, Yaosong (Institute of High Energy Physics Chinese Academy of Sciences, IHEP)

Session Classification: Basic and end-user IT services

Track Classification: Basic and End-User IT Services

Contribution ID: **33** Type: **not specified**

BNL Site Report

Monday, 15 April 2024 11:45 (15 minutes)

An update on recent developments at the Scientific Data & Computing Center (SDCC) at BNL.

Desired slot length

12 min

Speaker release

Yes

Primary authors: RIND, Ofer (Brookhaven National Laboratory); SMITH, Thomas; WONG, Tony

Presenter: SMITH, Thomas

Session Classification: Site reports

Track Classification: Site Reports

Contribution ID: **34** Type: **not specified**

Exascale challenges

Wednesday, 17 April 2024 15:40 (25 minutes)

High-performance digital technology has entered a new era in recent years with the arrival of exascale. Anticipated by experts for more than ten years, exascale is supposed to respond to increasingly varied uses that go far beyond traditional numerical simulation. Data processing and AI are shaking up the HPC landscape and have implications at all levels, from hardware to software, at the application level and the role of the supercomputer, which is losing its central role in the data processing chain.
During this presentation, we will present the current and future supercomputing landscape and we will illustrate with a few examples the issues linked to exascale.

Desired slot length

Speaker release

Yes

Primary author: CALVIN, Christophe

Presenter: CALVIN, Christophe

Session Classification: Computing and batch services

Track Classification: Computing & Batch Services

Contribution ID: **36** Type: **not specified**

Carbon Costs in IRIS & UKRI NetZero DRI

Thursday, 18 April 2024 09:25 (25 minutes)

As the UK's journey towards NetZero accelerates, we need robust information to inform both strategic and operational decisions, from policy development and funding allocation to hardware procurement, code optimisation and job scheduling.

The UKRI Digital Research Infrastructure NetZero Scoping Project published its technical report and recommendations in August 2023 [1] and funded the IRISCAST project which took a learning-by-doing approach to conduct a proof of concept 24-hour carbon audit snapshot across a multi-site heterogeneous research infrastructure [2].

IRIS [3] has taken this a step further by funding a Carbon Mapping Project (IRIS-CMP) to develop practical carbon models to apportion carbon costs and to deliver an outline delivery roadmap. These models have been tested with real world data from both the QMUL GridPP T2 and from STFC SCD-Cloud.

We present our key IRISCAST and IRIS-CMP findings, recommendations, and lessons learned, in the context of the UKRI Net Zero DRI journey.

[1] https://doi.org/10.5281/zenodo.8199984
[2] https://doi.org/10.5281/zenodo.7692451
[3] https://www.iris.ac.uk/

Desired slot length

20

Speaker release

Yes

Primary author: Dr OWEN, Richard Alexander (Queen Mary University of London)

Co-author: OWEN, Alex (University of London (GB))

Presenter: Dr OWEN, Richard Alexander (Queen Mary University of London)

Session Classification: IT facilities, business continuity and Green IT

Track Classification: IT Facilities, Business Continuity and Green IT

Contribution ID: **37** Type: **not specified**

IHEP Site Report

Monday, 15 April 2024 12:00 (15 minutes)

IHEP operates a comprehensive infrastructure comprising an HTC cluster, HPC cluster, and WLCG grid site, dedicated to facilitating data processing for over 20 experiments. Additionally, ongoing research in AI and QC is actively pursued.

Recently, we expanded our local HTC cluster by integrating it with 250 worker nodes from a remote Slurm cluster through gliding job slots. This enhancement significantly boosts our computational capabilities.

Furthermore, IHEP established the LHCb Tier 1, poised to commence operations within the current year.

Moreover, the construction of new machine room for the High Energy Physics Synchrotron (HEPS) has been completed.

Desired slot length

Speaker release

Yes

Primary author: Dr SHI, Jingyan

Presenter: Dr SHI, Jingyan

Session Classification: Site reports

Track Classification: Site Reports

Contribution ID: **38** Type: **not specified**

From Entanglement to Duality: Preparing a multi-ecosystem Linux strategy at CERN

Monday, 15 April 2024 15:30 (20 minutes)

The recent turmoils in the Red Hat ecosystem and the corresponding uncertainties they created in the HEP community have triggered the CERN Linux team to review their options for a multi-year Linux strategy. This presentation will summarise the state of Linux at CERN and discuss options moving forward as input to the Linux-themed discussion at this HEPiX meetup.

Desired slot length

10 mins

Speaker release

Yes

Primary author: Dr WIEBALCK, Arne (CERN)

Presenter: Dr WIEBALCK, Arne (CERN)

Session Classification: Linux future

Track Classification: Operating Systems, Cloud & Virtualisation, Grids

Contribution ID: **39** Type: **not specified**

A Lightweight Analysis Facility for the DARWIN Collaboration

Wednesday, 17 April 2024 14:20 (25 minutes)

A robust computing infrastructure is essential for the success of scientific collaborations. However, smaller collaborations often lack the resources to establish and maintain such an infrastructure, resulting in a fragmented analysis environment with varying solutions for different members. This fragmentation can lead to inefficiencies, hinder reproducibility, and create collaboration challenges.

We present an analysis facility for the DARWIN collaboration, a new dark matter experiment, designed to be lightweight with minimal administrative overhead while providing a common entry point for all DARWIN collaboration members. The facility setup serves as a blueprint for other collaborations, that want to provide a common analysis facility for their members. Grid computing and storage resources are integrated into the facility, allowing for distributed computing and a common entry point for storage. The authentication and authorization infrastructure for all services is token-based, using an Indigo IAM instance.

This talk will discuss the architecture of the facility, its provided services, the DARWIN collaboration's experience with it, and how it can serve as a sustainable blueprint for other collaborations.

Desired slot length

15

Speaker release

Yes

Primary author: BROMMER, Sebastian (KIT - Karlsruhe Institute of Technology (DE))

Presenter: BROMMER, Sebastian (KIT - Karlsruhe Institute of Technology (DE))

Session Classification: Basic and end-user IT services

Track Classification: Basic and End-User IT Services

Contribution ID: **40** Type: **not specified**

Institute Computing Infrastructure with Dynamic Extension

Tuesday, 16 April 2024 13:30 (25 minutes)

The Institute for Experimental Particle Physics (ETP) at the Karlsruhe Institute of Technology has access to several computing and storage resources. Besides the local resources such as worker nodes and storage, the ETP has access to the HPC cluster NEMO in Freiburg and to the Throughput Optimized Analysis System (TOpAS) cluster and Grid storage at the WLCG-Tier1 GridKa.
Hence, we use a pilot-like concept and the HTCondor flocking mechanism to make these additional resources transparent and dynamically available to users. This system provides users from ETP with up to several thousand CPU cores and several dozen data center GPUs in a homogeneous software environment.
This talk will show how to set up and use that computing infrastructure and its dynamic extensions. In addition to the admin point of view, the user point of view will also be discussed.

Desired slot length

15

Speaker release

Yes

Primary author: SCHNEPF, Matthias Jochen

Presenter: SCHNEPF, Matthias Jochen

Session Classification: Computing and batch services

Track Classification: Computing & Batch Services

Contribution ID: **41** Type: **not specified**

Micro-services for ATLAS muon alignment monitoring system

Wednesday, 17 April 2024 14:45 (25 minutes)

After a brief introduction on the Muon Alignment optical system and the dataflow for the optical lines, I will illustrate our infrastructure which is based on Micro-services in Java (mainly for the access to Oracle DB) and C++ (for the alignment algorithm itself) and deployed on a dedicated K8s cluster at CERN.

Desired slot length

10 minutes (possibly not on the 18th of April)

Speaker release

Yes

Primary authors: FORMICA, Andrea (Université Paris-Saclay (FR)); GIRAUD, Pierre-Francois (Université Paris-Saclay (FR))

Presenter: FORMICA, Andrea (Université Paris-Saclay (FR))

Session Classification: Basic and end-user IT services

Track Classification: Basic and End-User IT Services

Contribution ID: **42** Type: **not specified**

Packaging and system administration with Nix and NixOS

Friday, 19 April 2024 11:10 (25 minutes)

Nix is a tool for packaging software with a heavy focus on reproducibility. NixOS is a Linux distribution based on the Nix package manager.

This talk is a series of demonstrations of what Nix and NixOS can do for you. Depending on the time, here is what I'm going to show off:

- A presentation of the Nix model
- Its advantages in terms of supply-chain security
- Building minimal, reproducible Docker images
- Declarative Linux distribution with NixOS
- Remote administration of an offline NixOS system
- Rollbacks of a whole NixOS system
- Building NixOS Docker images
- Building VMs, ISOs, Containers
- Integration tests
- Cross-compilation and generating images for embedded systems

Desired slot length

Speaker release

Yes

Primary author: NICOLE, Remi (CEA IRFU/DIS)

Presenter: NICOLE, Remi (CEA IRFU/DIS)

Session Classification: Basic and end-user IT services

Track Classification: Basic and End-User IT Services

Contribution ID: **43** Type: **not specified**

Grand Unified Token (GUT) profile working group

Wednesday, 17 April 2024 11:10 (25 minutes)

The Grand Unified Token (GUT)-profile working group is trying to create a single OAuth2 token profile to replace the main token profiles: SciTokens, WLCG and AARC. These token profiles are being used by infrastructures and collaborations such as LIGO, HTCondor, WLCG, EGI etc. for a "new" authentication method, replacing the current X.509-based authentication. All these profiles share various characteristics, such as being based on JSON Web Tokens (JWT) and designed for typical R&E distributed infrastructures. On the other hand there are important differences making unification non-trivial.
With such a diverse group of people, organisations and timezones, the unification design is not the only non-trivial task for the group. We will go over how we overcome these challenges, discussing the way we work, and go over some of the details for how we plan to achieve unification.

Desired slot length

Speaker release

Yes

Primary author: Dr SALLÉ, Mischa (NWO-I Nikhef)

Presenter: Dr SALLÉ, Mischa (NWO-I Nikhef)

Session Classification: Network and security

Track Classification: Networking & Security

Contribution ID: **44** Type: **not specified**

Infrastructure as Code and DevOps approach for the SVOM mission

Thursday, 18 April 2024 12:00 (25 minutes)

The Space-based multi-band astronomical Variable Objects Monitor (SVOM) is a French-Chinese mission dedicated to the study of the most distant explosions of stars, the gamma-ray bursts. This talk will cover a brief overview of the whole mission infrastructure before focusing on the French Scientific Ground Segment (FSGS) infrastructure. The FSGS relies on a micro-services architecture with a full container based approach. The development of these micro-services are under the responsibility of several French and Chinese laboratories. Thus to ensure control and homogeneity among the different actors during the integration / delivery / deployment process we make an intensive use of the GitLab CI/CD by adopting a common git workflow and common job CI templates automatising the build, test, packages delivery and deployment steps. The micro-services orchestration is performed with Docker Swarms, we are currently migrating towards a Kubernetes cluster in high availability mode.
The FSGS uses 3 environments sites for the integration, pre-production stages and production. Each environment consists of an OpenStack project hosted at the IN2P3 Centre de Calculs and at IJCLab. The deployment of our infrastructure on each of these environments is fully automatised using a combination of Infrastructure as Code (IaC) tools, namely Terraform and Ansible. The former for provisioning the immutable OpenStack cloud infrastructure (Network, VMs, Volumes, Security Groups) and the later for configuring the VMs. This IaC approach drastically improved the immutability and idempotency of our infrastructure with reasonable effort which is valuable when manpower is limited.

Desired slot length

20

Speaker release

Yes

Primary author: Dr CORRE, David (CEA Irfu/Lilas)

Co-authors: Dr FORMICA, Andrea (CEA Irfu/Lilas); Dr LOUVIN, Henri (CEA Irfu/Lilas)

Presenter: Dr CORRE, David (CEA Irfu/Lilas)

Session Classification: Operating systems, clouds, virtualisation, grids

Track Classification: Operating Systems, Cloud & Virtualisation, Grids

Contribution ID: **45** Type: **not specified**

Power and Performance Impact of Inlet temperature differences

Thursday, 18 April 2024 09:50 (25 minutes)

At our site we have varied the datacenter inlet temperature between 23-25C while monitoring the effects on the total system power usage and temperature. In this talk i will give a overview of the results and findings of this. And how we collected all the relevant information, and how to visualize this in a useful format.

Desired slot length

15

Speaker release

Yes

Primary authors: Mr VAN DOK, Dennis (Nikhef); KOOISTRA, Erik

Presenter: KOOISTRA, Erik

Session Classification: IT facilities, business continuity and Green IT

Track Classification: IT Facilities, Business Continuity and Green IT

Contribution ID: **46** Type: **not specified**

News from the DESY clusters

Tuesday, 16 April 2024 13:55 (25 minutes)

We will give an overview and status, what over the past year is new and where we plan to go with our compute clusters. The migration to EL9 will be used for an overall update of Condor & Jupyter including a renovation & rewrite of the current configuration and some enhancements concluding from the past experience running the NAF.

Desired slot length

Speaker release

Yes

Primary authors: BEYER, Christoph; HARTMANN, Thomas (Deutsches Elektronen-Synchrotron (DE))

Presenter: BEYER, Christoph

Session Classification: Computing and batch services

Track Classification: Computing & Batch Services

Contribution ID: **47** Type: **not specified**

Performance of Cluster File System Backed by an HDD-Based Data Storage System Under True Concurrent Read-Write Load

Tuesday, 16 April 2024 09:50 (25 minutes)

Despite of the growing number of flash-based data storage systems the usage of spinning disks (HDDs) for large on-line data storage systems is still advantageous. Measurements of the read-write behaviour of a cluster file system using external storage controllers backed by HDDs are presented. Contrary to commonly expected balanced read and write rates, resp., or even read rates slightly outbalancing write rates by far prevailing write rates were seen. Starting point was the test procedure required in a Call for Tenders which turned out to be totally inadequate to characterize the system behaviour. A more thorough approach showed that in true parallel read-write traffic attempting to maximise both data streams the read rate is about one order smaller than the write rate. Possible Explanations are considered and some discussion of the results is given.

Desired slot length

12

Speaker release

Yes

Primary author: FALKE, Uwe (KIT SCC , GridKa)

Presenter: FALKE, Uwe (KIT SCC , GridKa)

Session Classification: Storage and file systems

Track Classification: Storage & Filesystems

Contribution ID: **48** Type: **not specified**

AGLT2 Site Update

Monday, 15 April 2024 13:45 (15 minutes)

We will present an update on our site since the Spring 2023 report, covering our changes in software, tools and operations.

The three primary areas to report on our work on performance evaluation with ZFS vs. Dell RAID systems, our plans and status for our transition from EL7 to RHEL9 and the work to deploy an operational WLCG Security operations Center implementation.

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

Desired slot length

12

Speaker release

Yes

Primary authors: HAYDEN, Daniel (Michigan State University (US)); LAURENS, Philippe (Michigan State University (US)); MC KEE, Shawn (University of Michigan (US)); Dr WU, Wendy (University of Michigan)

Presenter: MC KEE, Shawn (University of Michigan (US))

Session Classification: Site reports

Track Classification: Site Reports

Contribution ID: **49** Type: **not specified**

EPICS archiver infrastructure at ESS

Wednesday, 17 April 2024 09:50 (25 minutes)

ESS is getting ready for its next major milestone, which we call 'beam on dump,' where we will commission the full LINAC at the end of this year.
Although ESS is not yet completed, we have already built most of the IT infrastructure to support the control system for the accelerator, target, and neutron instruments.

System experts, operators, and beam physicists are already requesting to archive a lot of signals to understand, operate, and optimize their systems.
The control system being built with EPICS, we have deployed the open-source EPICS archiver appliance deployed on the technical network computing infrastructure to archive more than 700k PVs at 14 Hz and store up to 6TB of data per day.
We use CEPH for both block storage for those VMs and shared filesystem for the archiver appliance storage backend.

This presentation will walk you through the technical details of this implementation, some challenges we have faced, improvements we have made, as well as upcoming challenges

Desired slot length

20 min

Speaker release

Yes

Primary author: ARMANET, Stephane (European Spallation Source)

Presenter: ARMANET, Stephane (European Spallation Source)

Session Classification: Storage and file systems

Track Classification: Storage & Filesystems

Contribution ID: **50** Type: **not specified**

Computing for the KM3NeT research infrastructure

Tuesday, 16 April 2024 14:45 (25 minutes)

KM3NeT is a research infrastructure currently under construction in the Mediterranean Sea.
It consists of two neutrino detectors: ARCA for studying astrophysical sources and ORCA for studying neutrino properties.
Currently 15% of the infrastructure is operational.
The output of the entire infrastructure will eventually amount to a data rate of 100 Gbps, and a data volume of 500 TB per year.
In view of the final infrastructure configuration, the KM3NeT collaboration is transitioning to the use of standards and services in the e-Infrastructure commons.
This contribution focuses on the data processing and data management that KM3NeT envisions in this context.

Desired slot length

15min

Speaker release

No

Primary author: BOUWHUIS, Mieke (NIKHEF)

Presenter: BOUWHUIS, Mieke (NIKHEF)

Session Classification: Computing and batch services

Track Classification: Computing & Batch Services

Contribution ID: **51** Type: **not specified**

Secret management with HashiCorp Vault at DESY

Friday, 19 April 2024 11:35 (25 minutes)

An important aspect of IT security is the management, controlled sharing and storage of sensitive data such as passwords or API tokens. In this talk we present how HashiCorp Vault is used at DESY to address this challenge and how the system is integrated into workflows like certificate management and the existing IT infrastructure such as Puppet and GitLab. As secret management is a critical component for site operations, we describe how we aim for a fault tolerant and hardened setup.

Desired slot length

20 minutes

Speaker release

Yes

Primary author: WIEMANN, Kai (DESY Hamburg)

Co-authors: SEVER, Krunoslav (Deutsches Elektronen-Synchrotron DESY); Mr KÖLPIN, Maximilian (DESY Hamburg); STERNBERGER, Sven (DESY); HARTMANN, Thomas (Deutsches Elektronen-Synchrotron (DE))

Presenter: WIEMANN, Kai (DESY Hamburg)

Session Classification: Basic and end-user IT services

Track Classification: Basic and End-User IT Services

Contribution ID: **52** Type: **not specified**

WLCG Network Monitoring and Analytics Updates

Wednesday, 17 April 2024 13:30 (25 minutes)

Given the importance of the network to WLCG, it is important to guarantee effective network usage and prompt detection and resolution of any network issues, including connection failures, congestion and traffic routing. This talk will focus on the status and plans for the joint WLCG and IRIS-HEP/OSG-LHC effort to operate a global perfSONAR deployment and develop associated network metric analytics. We will report on the changes and updates that have occurred since the last HEPiX meeting, including recent updates to alerting and alarming and proactive problem identification.

Desired slot length

20

Speaker release

Yes

Primary authors: VUKOTIC, Ilija (University of Chicago (US)); BABIK, Marian (CERN); VASILEVA, Petya (University of Michigan (US)); MC KEE, Shawn (University of Michigan (US))

Presenter: MC KEE, Shawn (University of Michigan (US))

Session Classification: Network and security

Track Classification: Networking & Security

Contribution ID: **53** Type: **not specified**

Research Networking Technical Working Group Status and Plans

Wednesday, 17 April 2024 13:55 (25 minutes)

The high-energy physics community, along with the WLCG sites and Research and Education (R&E) networks have been collaborating on network technology development, prototyping and implementation via the Research Networking Technical working group (RNTWG) since early 2020.

In this talk we'll give an update on the Research Networking Technical working group activities, challenges and recent updates, emphasizing recent work related to DC24, Scitags and network use optimizations. We will also discuss near to long term plans for the group.

Desired slot length

20

Speaker release

Yes

Primary authors: BABIK, Marian (CERN); MC KEE, Shawn (University of Michigan (US))

Presenter: MC KEE, Shawn (University of Michigan (US))

Session Classification: Network and security

Track Classification: Networking & Security

Contribution ID: **54** Type: **not specified**

DESY site report

Monday, 15 April 2024 13:30 (15 minutes)

DESY site report

Desired slot length

12

Speaker release

Yes

Primary author: KEMP, Yves (Deutsches Elektronen-Synchrotron (DE))

Co-authors: HAUPT, Andreas (DESY); REEST, Peter van der

Presenter: KEMP, Yves (Deutsches Elektronen-Synchrotron (DE))

Session Classification: Site reports

Track Classification: Site Reports

Contribution ID: **55** Type: **not specified**

ARM Compute Testing and Provision at Glasgow's Tier2

Tuesday, 16 April 2024 14:20 (25 minutes)

With the recent developments in ARM technology and ongoing efforts by experiments in the integration of it into their workflows, there is increasing interest in getting Tier2 sites to obtain ARM kit in future procurements for testing and potential pledging. Here we present tests conducted by Glasgow on a variety of next-generation CPU to strengthen this case of future heterogenous computing facilities, and to share our experiences creating an ARM farm and running ARM work for LHC experiments.

Desired slot length

20

Speaker release

Yes

Primary author: SPITERI, Dwayne (University of Glasgow)

Co-author: SIMILI, Emanuele (University of Glasgow (GB))

Presenter: SPITERI, Dwayne (University of Glasgow)

Session Classification: Computing and batch services

Track Classification: Computing & Batch Services

Contribution ID: **56** Type: **not specified**

Simulating Carbon Opportunity Cost at Grid Site

Tuesday, 16 April 2024 16:30 (25 minutes)

With the advent of new species of ARM architecture on the market, and increasing developments by Intel/AMD to match the power-savings by ARM, it can be difficult for Grid sites to decide which machines to target in future procurements. While cost is an important factor, sites are increasingly able to make at least part of their choices on sustainability grounds. Obtaining test machines and running HEPScore and power measurements is only part of the story when it comes to making these decisions, and one machine does not make a farm. It can also be difficult to both have an active site, and perform the large-scale tests ideally required to make the most informed decision on both the equipment you want to buy and the way you can run the site. We present work done by Glasgow to simulate our active site, with the aim of testing different ways of running the site, and the potential savings in carbon from running different types of machines in the future.

Desired slot length

20

Speaker release

Yes

Primary author: SPITERI, Dwayne (University of Glasgow)

Presenter: SPITERI, Dwayne (University of Glasgow)

Session Classification: IT facilities, business continuity and Green IT

Track Classification: IT Facilities, Business Continuity and Green IT

Contribution ID: **57** Type: **not specified**

Collaborative operational security and Security Operations Centre deployment models

Wednesday, 17 April 2024 11:35 (25 minutes)

In previous HEPiX meetings we have presented on the strategic direction of the Security Operations Centre working group, focused on building reference designs for sites to deploy the capability to actively use threat intelligence with fine-grained network monitoring and other tools. This work continues in an environment where the cybersecurity risk faced by research and education, notably from ransomware attacks, remains persistent.

In this report we discuss recent developments in the working group, including a summary of our most recent hackathon, with a particular focus on potential methodologies for different types of facilities wishing to deploy this kind of capability.

Desired slot length

20 minutes

Speaker release

Yes

Primary authors: Dr CROOKS, David (UKRI STFC); VALSAN, Liviu (CERN)

Presenter: Dr CROOKS, David (UKRI STFC)

Session Classification: Network and security

Track Classification: Networking & Security

Contribution ID: **58** Type: **not specified**

Computer security update

Wednesday, 17 April 2024 12:00 (25 minutes)

This presentation provides an update on the global security landscape since the last HEPiX meeting. It describes the main vectors of risks and compromises in the academic community including lessons learnt, presents interesting recent attacks while providing recommendations on how to best protect ourselves.

Desired slot length

Speaker release

No

Primary authors: SUMAILOV, Roman (CERN); LUEDERS, Stefan (CERN)

Presenter: SUMAILOV, Roman (CERN)

Session Classification: Network and security

Track Classification: Networking & Security

Contribution ID: **59** Type: **not specified**

Update from the HEPiX IPv6 Working Group

Thursday, 18 April 2024 14:20 (25 minutes)

The HEPiX IPv6 Working Group has been encouraging the deployment of IPv6 in WLCG for many years. At the last HEPiX meeting in Canada we reported that more than 97% of all LHC experiment Tier-2 storage services are IPv6-capable. Since then, we have turned our attention to compute services and have launched a GGUS ticket campaign for WLCG sites to deploy dual-stack computing elements and worker nodes. The working group also monitored the use of IPv6 during the recent WLCG Data Challenge DC24. As before we continue to identify uses of legacy IPv4 data transfers and strive to move these to IPv6.

This talk will present the activities of the working group since October 2023 and our future plans.

Desired slot length

Speaker release

Yes

Primary author: KELSEY, David (Science and Technology Facilities Council STFC (GB))

Presenter: KELSEY, David (Science and Technology Facilities Council STFC (GB))

Session Classification: Network and security

Track Classification: Networking & Security

Contribution ID: **60** Type: **not specified**

From VOMS to INDIGO-IAM

Friday, 19 April 2024 09:00 (25 minutes)

The end of life of CentOS 7 accelerates the transition from VOMS proxies to OAuth tokens as the means to convey authorization information on a Grid/Cloud infrastructure. As a consequence, the VOMS and VOMS-Admin services will be abandoned in favor of INDIGO-IAM (or equivalent products) for the management of VO membership and the issuance of proxies and tokens.
In this contribution we present the current state of affairs for the transition from VOMS to IAM, in terms of development, deployment and usage, with a peek into the future.

Desired slot length

Speaker release

Yes

Primary author: GIACOMINI, Francesco (INFN CNAF)

Presenter: GIACOMINI, Francesco (INFN CNAF)

Session Classification: Basic and end-user IT services

Track Classification: Basic and End-User IT Services

Contribution ID: **61** Type: **not specified**

Backing up ET EMR Seismic data

Wednesday, 17 April 2024 09:00 (25 minutes)

The Euregio Meuse-Rhine border region between Belgium, the Netherlands and Germany is a potential site for the Einstein Telescope. In late October of 2023 Nikhef was asked to organise a backup and archiving of seismic survey data. This talk covers how the seismic data was then being shared and not backed up. The quick fix to backup the existing data; some custom python code being used as a medium term solution and longer term archiving.

Desired slot length

15

Speaker release

Yes

Primary author: Dr PICKFORD, Andrew (Nikhef)

Presenter: Dr PICKFORD, Andrew (Nikhef)

Session Classification: Storage and file systems

Track Classification: Storage & Filesystems

Contribution ID: **62** Type: **not specified**

Measuring Machine Metrics and Performance

Tuesday, 16 April 2024 16:05 (25 minutes)

The HEP Benchmark suite has been expanded beyond assessing only the CPU execution speed of a server via HEPScore23. In fact the suite incorporates metrics such as machine load, memory usage, memory swap, and notably, power consumption. In this report we detail the ongoing studies enabled by these new features.

Desired slot length

20

Speaker release

Yes

Primary author: GIORDANO, Domenico (CERN)

Presenter: GIORDANO, Domenico (CERN)

Session Classification: Computing and batch services

Track Classification: Computing & Batch Services

Contribution ID: **64** Type: **not specified**

Debian: What it is and what it can do for science

Monday, 15 April 2024 14:20 (20 minutes)

Presenter: DANDRIMONT, Nicolas

Session Classification: Linux future

Contribution ID: **65** Type: **not specified**

Discussion

Monday, 15 April 2024 14:55 (5 minutes)

Session Classification: Linux future

Contribution ID: **66** Type: **not specified**

Experience with migrating to Debian

Monday, 15 April 2024 14:40 (15 minutes)

Presenter: Mr VAN DOK, Dennis (Nikhef)

Session Classification: Linux future

Contribution ID: **67** Type: **not specified**

Status and plans at DESY

Monday, 15 April 2024 15:50 (10 minutes)

Presenter: KEMP, Yves (Deutsches Elektronen-Synchrotron (DE))

Session Classification: Linux future

Contribution ID: **68** Type: **not specified**

Dealing with CentOS 7 EOL

Monday, 15 April 2024 16:00 (30 minutes)

Round-table of sites

Session Classification: Linux future

Contribution ID: **69** Type: **not specified**

Discussion

Monday, 15 April 2024 16:30 (30 minutes)

Session Classification: Linux future

Contribution ID: **70** Type: **not specified**

Puppet module for Network Manager

Thursday, 18 April 2024 15:40 (15 minutes)

Speaker release

Desired slot length

Primary author: MIKULA, Alexandr (Czech Academy of Sciences (CZ))

Presenter: MIKULA, Alexandr (Czech Academy of Sciences (CZ))

Session Classification: Deploying dual-stack IPv6/IPv4 connectivity to site computing services

Contribution ID: **71** Type: **not specified**

IPv6 deployment on containers at RAL Tier1

Thursday, 18 April 2024 15:55 (20 minutes)

Primary author: BLY, Martin (STFC-RAL)

Presenter: BLY, Martin (STFC-RAL)

Session Classification: Deploying dual-stack IPv6/IPv4 connectivity to site computing services

Contribution ID: **72** Type: **not specified**

IPv6 deployment at KIT and removing IPv4

Thursday, 18 April 2024 16:15 (30 minutes)

Primary author: HOEFT, Bruno Heinrich (KIT - Karlsruhe Institute of Technology (DE))

Presenter: HOEFT, Bruno Heinrich (KIT - Karlsruhe Institute of Technology (DE))

Session Classification: Deploying dual-stack IPv6/IPv4 connectivity to site computing services

Contribution ID: **73** Type: **not specified**

DIscussion and further information

Thursday, 18 April 2024 16:45 (15 minutes)

Primary author: KELSEY, David (Science and Technology Facilities Council STFC (GB))

Presenter: KELSEY, David (Science and Technology Facilities Council STFC (GB))

Session Classification: Deploying dual-stack IPv6/IPv4 connectivity to site computing services

Contribution ID: **74** Type: **not specified**

WLCG GGUS ticket campaign for dual-stack CPU services and worker nodes

Thursday, 18 April 2024 14:45 (25 minutes)

Speaker release

Desired slot length

Primary author: Dr SCIABÀ, Andrea (CERN)

Presenter: Dr SCIABÀ, Andrea (CERN)

Session Classification: Deploying dual-stack IPv6/IPv4 connectivity to site computing services

Contribution ID: **75** Type: **not specified**

Intro

Wednesday, 17 April 2024 16:05 (5 minutes)

Session Classification: Show us your toolbox

Contribution ID: **76** Type: **not specified**

Alex's Toolbox

Wednesday, 17 April 2024 16:50 (10 minutes)

Session Classification: Show us your toolbox

Contribution ID: **77** Type: **not specified**

Dennis' Toolbox

Wednesday, 17 April 2024 16:15 (10 minutes)

Session Classification: Show us your toolbox