HEPIX Spring 2025 Workshop Wrap-up

Lugano, Switzerland

Joint SDCC-NPPS Meeting
04/17/2025

O. Rind, adapted from slides by J. Flix

With contributions from: Andreas Petzold, Sébastien Gadrat, Shawn McKee, Andrew Pickford, Elvin Sindrilaru, Elia Oggian, Dino Conciatore, Michele Michelotto, Max Fischer, Matthias Schnepf, Peter Gronbech, David Britton, Jingyan Shi, Dennis van Dok, Mattias Wadenstein, Alessandro Di Girolamo, James Letts, Helge Meinhard, Ofer Rind, Tomoaki Nakamura

What is HEPiX?

The HEPiX forum brings together worldwide Information Technology staff, including system administrators, system engineers, and managers from the High Energy Physics and Nuclear Physics laboratories and institutes, to foster a learning and sharing experience between sites facing scientific computing and data challenges. Participating sites include ASGC, BNL, CERN, DESY, FNAL, IHEP, IN2P3, INFN, JLAB, KEK, KIT, Nikhef, PIC, RAL, SLAC, TRIUMF and many others. The HEPiX organization was formed in 1991, and its semi-annual meetings are an excellent source of information and sharing for IT experts in scientific computing.

- Active Working Groups: TechWatch, Benchmarking, IPv6
- Co-Chairs: J. Flix (Europe), T. Nakamura (Asia), O. Rind (N. Amer.)

More info: https://www.hepix.org/

Spring 2025 Workshop

HEPiX Spring 2025 Workshop

31 March 2025 to 4 April 2025
Hotel De La Paix
US/Eastern timezone

Enter your search term

- Overview
- Scientific Program
- Workshop Chairs and Organizers
- Payment, Visas and Invitation Letters
- Call for Abstracts
  - └ Reviewing Area
- Timetable
- Contribution List
- My Conference
  - └ My Contributions
- Book of Abstracts
- Registration
- Participant List
- Welcome reception and Social dinner
- Accommodation
- Transportation and Workshop Venue

HEPiX Spring 2025 at the Swiss National Supercomputing Centre in Lugano

#weareALPS

Scientific Program

Site Reports

Networking & Security

Storage & data management

Cloud Technologies, Virtualization & Orchestration, Operating Systems

Computing & Batch Services

Environmental sustainability, business continuity, and Facility improvement

Software and Services for Operation

Mid-long term evolution of facilities (Joint Topical Session with WLCG OTF)

3

Special thanks to our Sponsors

HEPiX Spring 2025 - Lugano

31 March - 4 April 2025

124 participants

40 institutions

18 countries

8 sponsors

75 contributions

10 tracks

1 topical session

26.6 hours

HEPiX Spring 2025

Swiss National Supercomputing Centre, Lugano, Switzerland
31 March - 4 April 2025

Photo credit: Dino Conciatore

Academic Collaborations (Sponsors)

The **Cherenkov Telescope Array Observatory (CTAO) talk** (by Andrii Neronov) described the next-generation gamma-ray observatory. Andrii highlighted the CTAO role in studying cosmic particle accelerators

- Generating and processing hundreds of petabytes of data, leveraging WLCG technology
- Off-site data centers will receive ~6PB/year to store/process: CSCS, PIC, Frascati, DESY Zeuthen
- CTAO follows an open-data approach to support multi-wavelength/messenger astronomy

Emma Tolley showed the computational and data challenges for the upcoming large-scale radio telescope **Square Kilometer Array (SKA)**: processing the acquired radio signals from space will require huge computational resources

- Massive data processing needs: hours of observation require tens of thousands of core-hours
- I/O bottlenecks: new data formats (MSv4, Zarr) and GPU acceleration are being explored
- AI-based compression & efficient computing are critical to reduce CO_2 footprint
- Swiss SKA Regional Center: Switzerland will host a computing hub for global SKA research

Mid-long term evolution of facilities I

Topical session (HEPIX + OTF)

- **"What Will Sites and Site Services Look Like in Five or Ten Years?"**
- **Scope** of this session:
  - Deepen the strong connection between WLCG and the facilities
  - Explore the main issues about facilities that should be addressed in the WLCG Technical Roadmap (work ongoing, ETA for a skeleton a few months)

- Many **R&D efforts** are focused on understanding **analysis use cases**
  - Most sites supports several communities (also not-LHC)
  - Users support is key pillar

Mid-long term evolution of facilities II
Topical session (HEPIX + OTF)

- We heard from **Italy**, **Germany**, **Spain**, and **US ATLAS** about their plans for distributing or concentrating compute and storage and integrating HPC resources. The **UK** does not have HPC capacity at this time but would find concrete requirements from the experiments useful to help argue for funding
  - Different strategies for the various countries
  - Tensions in terms of HW planning: "world" seems to go toward GPUs, but it seems that LHC Experiments, in the next a few years, do not really need GPUs at scale
  - Full exploitation of HPCs (e.g. InfiniBand) is also very difficult with experiments workloads
  - if GPUs are not (massively) needed, is important to explain why
- Job submission with **K8S API** (direct, no layers like condor) is successful on sites in US and Canada

Site Reports I

12 site reports (7 from EU, 2 from US and 3 from Asia)
- CSCS, CERN, BNL, RAL, NDGF, DESY, KEK, IHEP, US ATLAS SouthWest Tier-2
- 3 (almost) new sites: Lisbon (NCG-INGRID-PT), Lanzhou and CTAO Swiss DC

Usual softwares (everywhere)
- Compute: HTCondor (mostly) or Slurm
- Storage: dCache, Lustre, CePH and EOS
- Virtualization: OpenStack
- Orchestration: kubernetes & GitLab (DevOps almost everywhere now)

New architectures still growing
- more GPU (even used for Quantum Computing)
- ARM
- sometimes integrated in production (GPU in PanDA queues, for ATLAS)

Site Reports II - News and trends

- Migration to (RH)**EL9** on-going (CentOS7 with ES still around)
- **Puppet** future not quite clear (some investigating OpenVOX)
  - "We're not forking Puppet; Perforce is forking Puppet"
- **Security challenges**, especially with IoT devices
- **New rooms/buildings** being or to be built to cope with the experiments needs
- More **AI**
  - add models training service, add support to HuggingFace
  - (strategic) discussions about GenIA

Computing and Batch Services I

- Reports from the **benchmarking WG**
  - Domenico ([link](#))
    - new releases of HEPScore and HEP benchmark-suite
    - Plugin system in the benchmark-suite (load, Watt)
  - Natalia ([link](#))
    - Continuous benchmarking of site to monitor the core power on WLCG and spot error in declared power (under or over)
    - Identifying badly configured sites
- **ARC 7**: New major release with fewer components and better performance; release date: 31.03.25
- Introduction into **MTCA systems** (Micro Telekom Computing Architecture) and into the StarterKit from powerBridge Computing (Sponsor)

Computing and Batch Services II

HEPiX

- 10th **European HTCondor [workshop](#)** -> 11th coming soon in 16-19 September 2025 (Prague)
- **[Auditor](#)** to manage huge amount of accounting data
- How to manage **batch queues** efficiently
  - Integrating GPU in the traditional CPU queue
  - …and [Quantum Processing Unit](#) (Sponsor - E4))
  - Adding caps on utilization and users
  - [Efficiency of Job Processing in Many Core Grid and HPC Environments](#)

CONCLUSIONS AND OUTLOOK

➡ **scaling corecount up from 8core poses challenges to processing efficiencies in WLCG**
  - ✳ ALICE, CMS and LHCb schedule multiple payloads per each pilot, the scheduling efficiency is being actively studied by the experiments
  - ✳ ATLAS runs one payload per job, needs scalability at the software level

➡ **the investigation on the ATLAS most used workloads shows that:**
  - ✳ **corecount scaling of single payloads negatively impacts CPUeff and event throughput**
  - ✳ there is not a common reason for the inefficiencies
  - ✳ tuning the job size can help, but marginally: 16-core seems so far the sweet spot
  - ✳ **whole-node (128+) only feasible for Full Sim, other workloads won't run to completion**
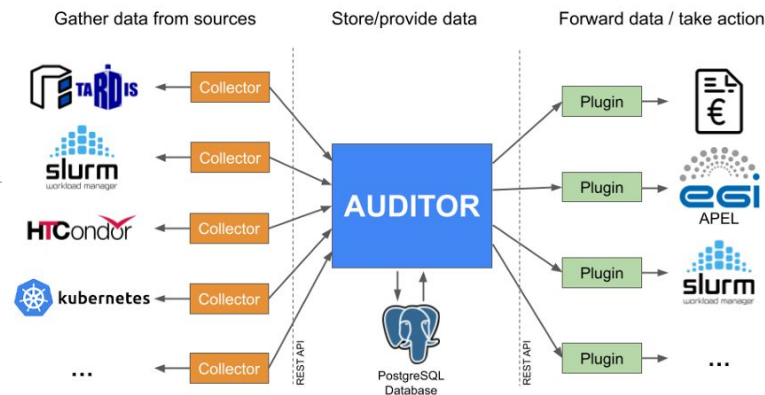
Cloud technology, Virtualization & Orchestration, Operating Systems I

- **Tier-2 on Kubernetes University of Victoria**
  - Running ATLAS (and also SKA) jobs on Kubernetes without a traditional batch system or compute element. Helm installation of T2Site. Leveraging WLCG with cloud-native components like Kuantifier, Frontier-squid, and EOS on CephFS, achieving high performance. Migration to ArgoCD for deployment management. Optimizations using Cilium and eBPF. Simplified upgrades for better maintainability

- **Kubernetes and Cloud Native at the SKA Regional Centres**
  - The presentation discusses the development of a federated SKA Regional Centre (SRC) Network to store, process, and manage SKA data. It outlines the roadmap, deployment strategies, and cloud-native approaches using Kubernetes. Utilizing tools like ArgoCD, JupyterHub, and Prometheus. Future focus on federated execution, monitoring models, and enhanced user engagement.

- **Hyper-converged cloud infrastructure at CSCS**
  - Updates from the CSCS private cloud infrastructure maintaining ~50 k8s clusters using SUSE Virtualization (Harvester) and SUSE Rancher, extending Kubernetes on HPC workernodes, and leveraging Crossplane as a single control plane to manage the entire cloud infrastructure.

Cloud technology, Virtualization & Orchestration, Operating Systems II

- **Keeping the LHC colliding: Providing Extended Lifecycle support for EL7**
  - The LHC relies on thousands of Front-End Computers (FECs), many running CERN CentOS 7, now outdated. Upgrading is difficult due to hardware limits, so CERN is rebuilding ELS7 packages to extend support until 2026. Long-term, they plan to transition to Debian

- **Roadmap to LS3: CERN's Linux Strategy**
  - CERN's Linux strategy includes AlmaLinux Kitten for testing, EPEL10 for better RHEL alignment, and an ISA shift to x86_64-v3 for performance gains. Dropped support for older hardware drives CERN to explore Debian as a backup to reduce RHEL dependency

- **Exploring SUSE Open-Source Technology for Your Datacenter** (Sponsor)
  - Showcase of fully open-source projects maintained by SUSE, including openSUSE Linux distributions, Harvester (a cloud-native hypervisor solution), Rancher (a multi-Kubernetes manager), NeuVector for zero-trust security, and Longhorn as distributed storage for Kubernetes.
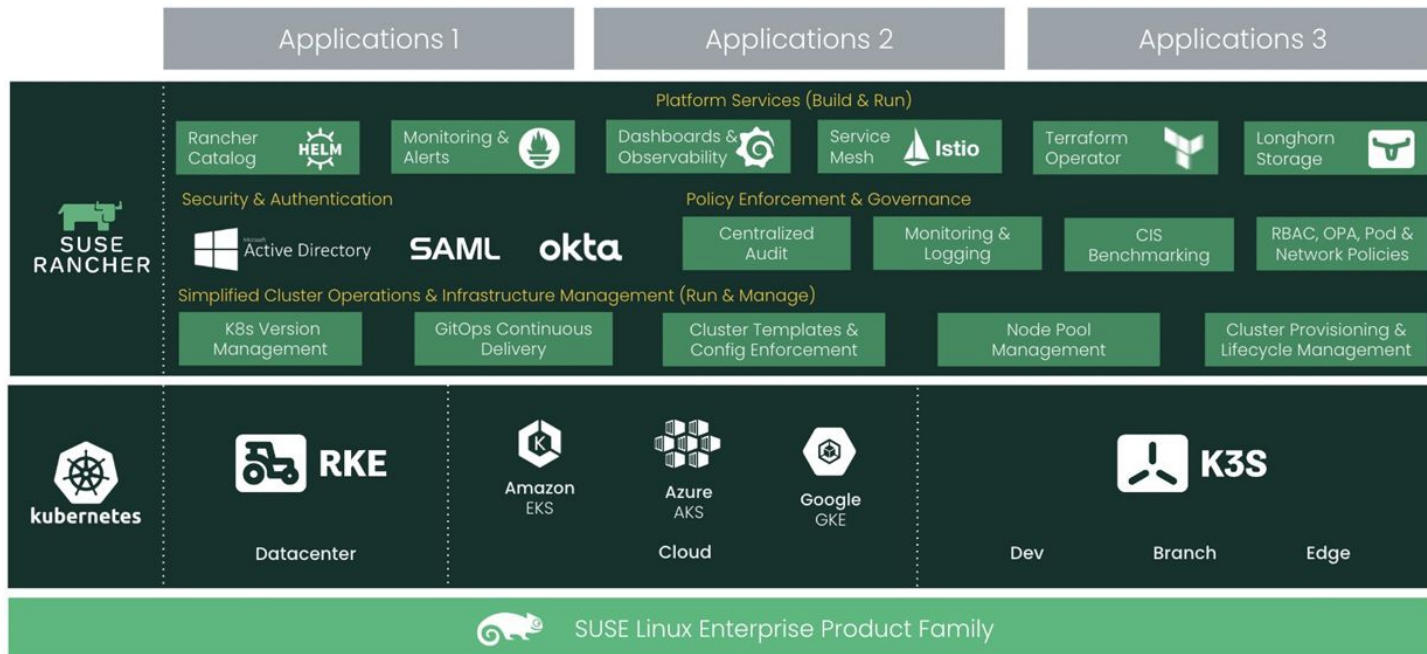
# Environmental sustainability, business continuity, and Facility improvement I

- **Updates on CPUs, GPUs and AI accelerators**: Michele gave a summary from the HEPiX Techwatch group of new CPU's etc including an AMD beast of a CPU which draws 500W. It uses 12 memory channels which need lots of space on the motherboard, so no way to have 4 nodes in 2u like old days. Next gen Xeon also 500W, Nvidia NVL144 600KW rack!
  - *See, in particular, slides about GPU data types and numerical precision*
- Eric Yen presented the **Trends of Energy Efficiency for computing and data centre**
- Tristan gave a nice summary of the challenges of providing simple to use **A/V** that just works at Nikhef and pitfalls to avoid
- **CSCS Sustainability: Utilizing Lake Water for Cooling and Reusing Waste Heat** Tiziano gave a description of how the CSCS DC use water taken from deep in the lake to cool multiple clusters using only one pump cycle, and then how the waste heat is used by a nearby school campus. PUE ~1.108

# Environmental sustainability, business continuity, and Facility improvement I

- **Natural job drainage and power reduction studies in PIC Tier-1 using HTCondor**: Pepe told us about simulation using AI to model time to drain PIC cluster with varied job mix, and how it affects the power utilisation
- **Transforming the Disaster Recovery of the Cloud Service**: Varsha provided a description of the automation of migration of services from one DC in CERN to the other to provide Disaster Recovery
- **Update on Energy Efficiency: AmpereOne and Turin**: Dave gave an update of the work Glasgow has been doing modelling how the power consumption varies with CPU clock frequency, which can be used to reduce power consumption during times when the electricity supply has a high carbon content

Characteristics of Natural Drainage Cycles
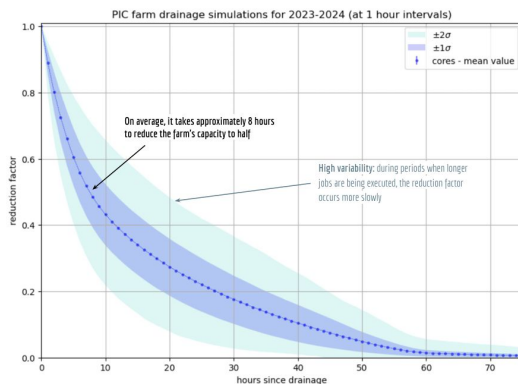
ATLAS makes significant use of the farm, with shorter jobs as compared to **CMS** and **LHCb**, making them the first to be drained

AI techniques for reduction predictions
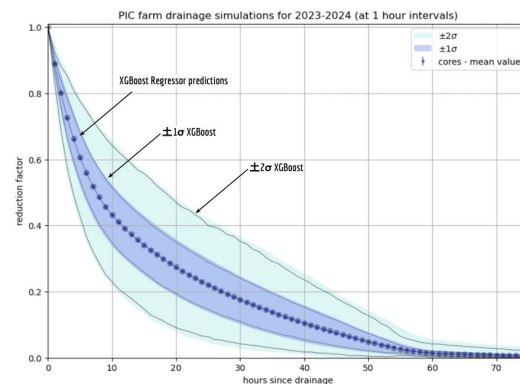
Network & Security I

The first network & security session Wednesday had 4 very nice presentations covering network infrastructure design, security monitoring, incident response workflows, and proactive threat management within the high-energy physics community

CERN Prévessin Datacentre Network Design **Presenter**: Vincent Ducret (CERN IT/CS-NE)
  Covered the design, implementation and plans for the CERN PDC network
Monitoring and Incident Response at CSCS **Presenter**: Fabio Zambrino (CSCS)
  Described activities and tools used at CSCS to monitor, alert and respond to security incidents for a diverse infrastructure with over 4,000 nodes
Endpoint Detection and Response (EDR) at CERN **Presenter**: Alexandros Petridis
  Described the CERN EDR solution to enhance endpoint security with continuous monitoring, live forensics, and proactive threat hunting capabilities
Computer Security Update **Presenter**: Jose Carlos Luna Duran (CERN)
  Provided the next installment of the CERN security update focusing on how to prepare for inevitable breaches by developing forensic toolkits, implementing robust monitoring systems, and maintaining comprehensive backups

Network & Security II

The second (and final) session had 5 clear and detailed presentations reporting on working group activities, network monitoring, CERN SSO evolution and the design and deployment of the HEPS network

Update from the HEPiX IPv6 Working Group **Presenter**: Bruno Hoeft
    Summarized the status and plans of the working group toward IPv6-only
Activities Update Research Networking Technical Working Group **Presenter**: S McKee
    Described ongoing work in network visibility, optimization and SDN
WLCG Network Monitoring Infrastructure and perfSONAR Evolution **Presenter**: S McKee
    Presented on network monitoring status and upcoming changes
Single Sign-On Evolution at CERN **Presenter**: Paul Van Uytvinck
    Described the history and current status of SSO work at CERN
Network design and implementation status of HEPS **Presenter**: Shan Zeng
    Presented on the HEPS light source campus network design and status
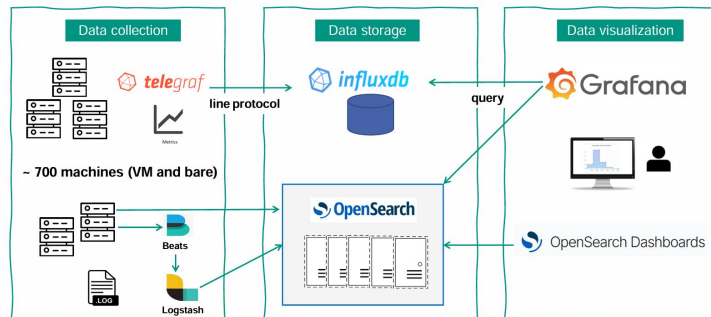
Track Summary: 9 presentations, a broad range of Network+Security topics

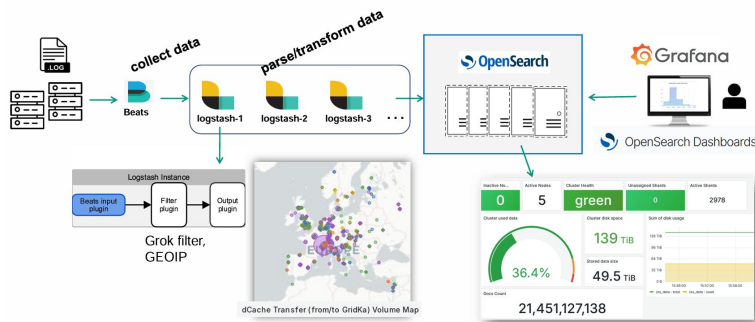# Software & Services for Operation I

7 contributions in this track

- **Infrastructure Monitoring for GridKa and beyond** – Evelina Buttitta
  Current monitoring architecture is in flux, as replacement technologies are under consideration. (e.g. Influxdb → VictoriaMetrics, OpenTelemetry?)
- **Windows device management at CERN: A new era** – Siavaș Firoozbakht
  Implementing Intune and MCM at CERN for over 10k Windows devices according to use-cases 'buckets'
- **JUNO Distributed Computing Infrastructure and Services Monitoring System** – Xuantong Zhang
  Monitoring system constructed from scratch, in a workflow-based probing architecture. including HTCondor tasks
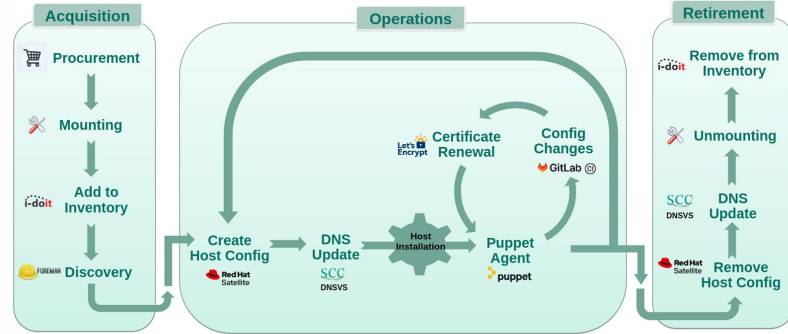
Software & Services for Operation II

- **Integrated Configuration Management at Karlsruhe Institute of Technology (KIT)** – Nico Schlitter

  Automation for consistent and reproducible host deployment. Removing the admin from low-level tasks

- **From Batch to Interactive: The "INK" for High Energy Physics Data Analysis at IHEP** – Dr Jingyan Shi

  "Interactive aNalysis worKbench" Enabling interactive, web-based analysis using batch system resources

- **Grafana dashboards as code Feat. Jsonnet & Grafonnet** – Ewoud Ketele

  Comprehensive job monitoring across all of WLCG; bringing 5600 lines of JSON for a dashboard down to 76 lines of code

- **MarmotGraph @ CSCS** – Oliver Schmid

  "Management Applications for Rich Metadata ObjecTs" Originally for the Human Brain Project, this graph database has over 45k entities to query for all sorts of operational matters
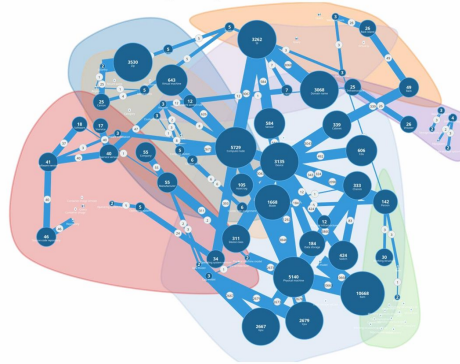
Host Life Cycle

Scientific Computing Center
Karlsruhe Institute of Technology (KIT)

Loose Architecture of INK
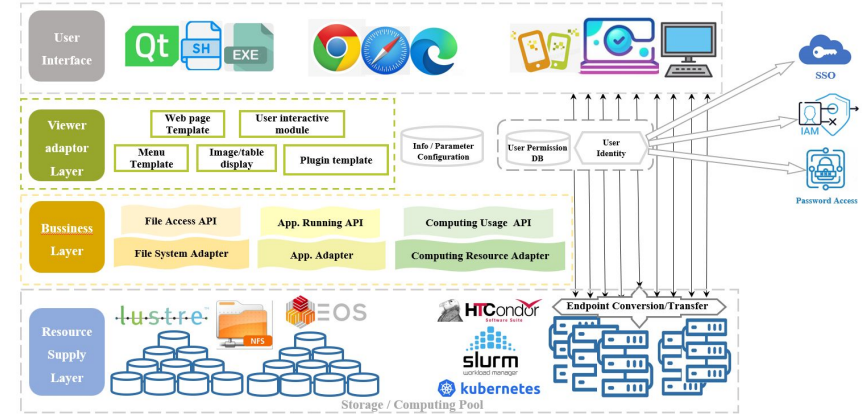
The current graph

CSCS
45'824 entities*
(506'117 triples, 66 types)

EBRAINS
>4 Mio entities*
(>29 Mio triples, ~200 types)

*State on Apr 3rd, 2025

ETH zürich

24

Storage & data management I

- **EOS, CERN's open-source storage system**, presented by Elvin, has evolved to meet growing performance and user demands, ensuring readiness for the upcoming HL-LHC era
- Roberto showed the **CERN's custom backup tool "cback"**, that ensures scalable, secure, and efficient backups for EOS and CephFS, addressing disaster recovery needs through innovative strategies and design choices applicable to broader backup solutions
- Octavian talked about refurbishing the **Meyrin Data Centre**: Storage and Operations, scalability, and reliability in the coming years
- Cedric presented his **Distributed Storage Odyssey from CentOS7 to ALMA**
- Marina described how to setup **Label-based Virtual Directories In dCache**
- Mohamad presented **NVMe-HDD Solution-Level Usage Models, Features and Advantages** (Sponsor)

Storage & data management II

- **How CERN's New Datacenter Enhances Cloud Infrastructure and Data Resilience with Ceph:** Roberto demonstrated how CERN orchestrates Ceph clusters across its two locations to support high-availability services
- **CERN Update on Tape Technology:** Vladimir gave a highly detailed and engaging talk on the evolution of tape systems over the years
- **Evolution of Continuous Integration for the CERN Tape Archive (CTA):** Niels shared the significant progress that has been made in improving CI for CTA
- **The CERN Tape Archive Beyond CERN:** Niels highlighted the key concepts and features of CTA, including its integration with EOS and dCache
- **Storage Technology Outlook:** Ed (Sponsor) discussed how HDD technology is not keeping pace with data growth and emphasized the strong future of tape storage
- **Online Seamless HDD Self-Healing Options & Capabilities:** Curtis (Seagate) outlined various strategies for self-healing HDDs, providing an in-depth analysis of the pros and cons of each approach

Miscellaneous I

- Fabien Wernli presented the **CC-IN2P3's computer history museum** in detail. He highlights the importance of preserving IT infrastructure for education and science, and showed the challenges of running such museum
- With green computing in focus and the high-luminosity LHC approaching, Daniël Geerts explored **optimizing ATLAS analysis code** to improve performance, enhance resource efficiency, and reduce processing time
- Gino Marchetti shared five years of experience managing **CC-IN2P3 user documentation**, covering technology choices, content update procedures, and the automated software stack that supports 2,700 computing users and 10,000 GitLab users
- Thomas Jyothish provided an overview of **RAL's use of the XRootD Cluster Management System for ECHO storage**, including configuration, custom scripts, and insights from different setups

Miscellaneous II

- Ricardo Goncalves presented the **CERN's Microsoft SQL Server infrastructure**, detailing server management, database solutions, high availability, backup strategies, and data integrity measures for critical applications
- The evolution of **PIC's Big Data Analysis Facility** was shown by Francesc, detailing its Hadoop-based architecture, key services like CosmoHub and SciPIC, integration with PIC's computing ecosystem, and future plans for interoperability, identity federation, and multi-messenger astronomy support
- Andrey Shevel showed that **Machine Learning**, particularly Retrieval Augmented Generation (RAG) with a large language model (LLM), aids developers and administrators in **managing complex computing systems**
- Ruben Gaspar introduced the **Transcription and Translation as a Service** (TTaaS) project at CERN, which leverages Automated Speech Recognition (ASR) and translation technologies to enhance accessibility, diversity, and searchability of CERN's audiovisual content
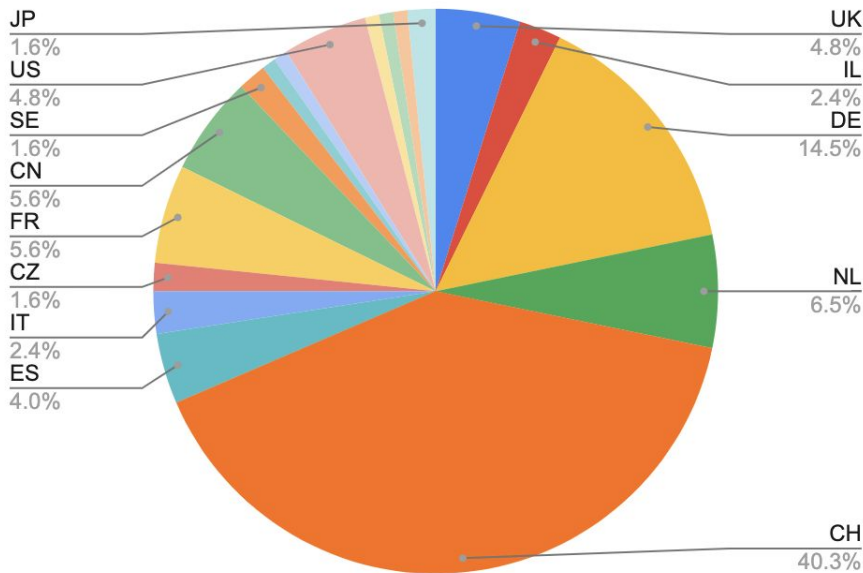
Extras

Participation

Count of Country

~40 institutions
18 countries

Most of the participants European
7 participants from North America
14 participants from Asia
Many 1st time attendees
+16% wrt last EU HEPIX ws

HEPIX Board meeting

- Lengthy review of the current meeting
- **Board was pleased** with the way the meeting was running, the success of the topical session with OTF, and the content in general
  - Track names were adjusted (we might review tracks for future meetings)
  - Early announcement and registration process + clear deadlines helpful
  - Code of conduct added to HEPiX workshop indico page
- Report on progress for the Fall meeting, to be held **03 – 07 November, 2025 at Lanzhou, China**
  - Arrangements are in progress, with announcement forthcoming shortly and registration to open by June 2025 - **save the date!**
- Presentation of **LIP proposal to host HEPiX in Lisbon** in Spring 2026
  - The HEPiX Spring 2010 Workshop was disrupted by the Icelandic volcano eruption… The board agreed for Lisbon next year!
- Reports from all three **working groups**
- Fresh eyes are welcome to improve workshops organization…**please contact us!**