

HEPiX Fall 2025 Workshop



Report of Contributions

Contribution ID: 1

Type: **not specified**

Scientific SOC for STFC and RAL

Wednesday, 5 November 2025 13:30 (30 minutes)

In the research and education environment, cybersecurity threats are significant and growing. We must collaborate as a community protect our environment.

Effective protection requires the use of detailed, timely and accurate threat intelligence alongside fine-grained monitoring.

We illustrate the current capabilities of the SOC. Covering how we collect, enrich, analyse and use security relevant data.

Speaker release

Yes

Author: ATHERTON, Liam

Presenter: ATHERTON, Liam

Session Classification: Network & Security

Track Classification: Networking & Security

Contribution ID: 2

Type: **not specified**

CERN Site Report

Monday, 3 November 2025 13:30 (20 minutes)

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

Speaker release

Yes

Author: SINDRILARU, Elvin Alin (CERN)

Presenter: SINDRILARU, Elvin Alin (CERN)

Session Classification: Site Report

Track Classification: Site Reports

Contribution ID: 3

Type: **not specified**

Automating Resource Lifecycles and Service Subscriptions with Dynamic Authorization

Wednesday, 5 November 2025 11:00 (30 minutes)

Computing projects rely on a wide range of resources and services: databases to store data, containers to run simulations, websites to publish results, and the infrastructure needed to test and deploy software. High Energy Physics is no exception, and access to these resources must be regulated—for instance, only members of a given experiment may use a specific database cluster, or only staff in HR may access certain applications.

Without proper lifecycle management, resources and permissions are difficult to track, privileges accumulate over time, and ownership becomes unclear. This is both inefficient and a security risk: critical resources may remain active but unmaintained, or accessible to users who no longer require them.

At CERN, we are addressing these challenges by automating the management of resource ownership and service subscriptions. In 2025, one of the IT department's priorities has been the migration of services to a new system based on Authentication and Authorization technologies. This framework enables permissions to be defined through dynamic queries and eligibility criteria to be enforced consistently across services.

We will present an overview of the system architecture and technologies involved, the progress made during the migration, and aspects that may be of wider interest to other sites.

Speaker release

Yes

Author: TEDESCO, Paolo (CERN)**Co-authors:** LOSSENT, Alexandre (CERN); Mr CABAS ALVAREZ, Diego; ORMANCEY, Emmanuel (CERN); SHORT, Hannah (CERN); VAN VARK, Thomas; BRILLAULT, Vincent (CERN)**Presenter:** TEDESCO, Paolo (CERN)**Session Classification:** Software and Services for Operation**Track Classification:** Software and Services for Operation

Contribution ID: 4

Type: **not specified**

An update on the Web landscape at CERN

Wednesday, 5 November 2025 09:30 (30 minutes)

CERN, the birthplace of the World Wide Web, continues to evolve its web infrastructure to follow technology evolution and meet users' needs. This talk will provide an overview of the web landscape at CERN, from web governance to web hosting services, including recent and ongoing changes. We will look at the change of the Content Management System, and the transition from Drupal to a new WordPress hosting service. We will also discuss the migrations from legacy platforms (WebAFS, WebDFS and on-prem SharePoint 2013) to more modern alternatives (WebEOS and SharePoint Online), and trends such as the growing popularity of GitLab Pages. Along the way, we'll share lessons learned from these migrations, and mention possible future developments of web services at CERN.

Speaker release

Yes

Author: LOPIENSKI, Sebastian (CERN)**Presenter:** LOPIENSKI, Sebastian (CERN)**Session Classification:** Software and Services for Operation**Track Classification:** Software and Services for Operation

Contribution ID: 5

Type: **not specified**

Evolving Data Stores Services for Big Data at CERN

Wednesday, 5 November 2025 10:00 (30 minutes)

The Data Stores section of the CERN IT Database and Analytics group provides foundational services to store, process, and analyze scientific and operational data across the CERN community. These services currently rely on technologies such as Hadoop, OpenSearch, and NetApp. As data volumes and usage patterns continue to increase, the team is actively evolving the service portfolio to meet new requirements in scalability, flexibility, and performance.

For the Hadoop ecosystem, work is underway to integrate Apache Ozone as an alternative to HDFS. Ozone introduces object storage semantics, enabling improved scalability and integration with cloud-native workflows. In parallel, an evaluation of ClickHouse is being conducted as a potential new service to address online analytical processing (OLAP) use cases, with emphasis on high-throughput queries and near real-time analytics at scale.

The OpenSearch service is undergoing a strategic reassessment. While currently deployed on bare metal via Ironic, investigations are exploring the feasibility of running selected OpenSearch clusters on Kubernetes to enhance elasticity, resource efficiency, and lifecycle management.

This presentation will provide an overview of these initiatives, the motivations driving the service evolution, the technical challenges encountered, and the expected benefits for the CERN community. The roadmap for production integration will be discussed, together with knowledge sharing and lessons learned that may be of interest to peer institutes undertaking similar developments.

Speaker release

Yes

Author: ANDRADE, Pedro (CERN)

Presenter: ANDRADE, Pedro (CERN)

Session Classification: Software and Services for Operation

Track Classification: Software and Services for Operation

Contribution ID: 6

Type: **not specified**

CNAF site report

Monday, 3 November 2025 13:50 (20 minutes)

Site report from INFN-T1

Speaker release

Yes

Author: PASCOLINI, Alessandro (Universita e INFN, Bologna (IT))

Presenter: PASCOLINI, Alessandro (Universita e INFN, Bologna (IT))

Session Classification: Site Report

Track Classification: Site Reports

Contribution ID: 7

Type: **not specified**

CERN's Apple Playbook: Smart Strategies for macOS and iOS Management

Thursday, 6 November 2025 10:00 (30 minutes)

Managing Apple devices at scale in a large organization presents unique challenges and opportunities. Our environment includes thousands of macOS and iOS devices used by a diverse workforce with varied needs. To support this, we have developed a management strategy built around automation, security, and user experience.

This session will explore how we deploy and manage Macs and iOS devices across the enterprise using Mobile Device Management (MDM). We will cover topics such as automated device enrollment through Apple School Manager, enforcing security baselines without compromising usability, enabling self-service for software distribution and configuration, and providing applications and updates to ensure users have the tools required for productivity, collaboration, and specialized workflows.

We will share how we balance centralized IT control with user autonomy, and the lessons we've learned in maintaining a secure, consistent, and user-friendly Apple ecosystem at scale. Attendees will leave with practical insights into Apple device management, approaches for integrating services, and methods for sustaining long-term operational efficiency.

Speaker release

Yes

Author: MUSZKOWSKI, Maciek (CERN)

Presenter: MUSZKOWSKI, Maciek (CERN)

Session Classification: Cloud Technologies, Virtualization & Orchestration, Operating Systems

Track Classification: Cloud Technologies, Virtualization & Orchestration, Operating Systems

Contribution ID: 8

Type: **not specified**

HEPiX-IPv6 status report

Wednesday, 5 November 2025 15:00 (30 minutes)

The HEPiX-IPv6 working group will present the status report. The current still ongoing GGus ticket campaign of the dual-stack protocol deployment at the wlcg tier-1/2 Worker node farms will be presented.

Even while the major data exchange of LHCOPN is over IPv6, some of the still remaining IPv4 flows will be shown and examples how to mitigate them.

The partial remove at end of September 2025 of IPv4 from LHCOPN as well as the pro and cons of the removal will be presented and the timing of the complete removal of IPv6 from LHCOPN will be discussed.

The next steps of going the roadmap towards IPv6 only will be highlighted

Speaker release

Yes

Author: HOEFT, Bruno Heinrich

Presenter: HOEFT, Bruno Heinrich

Session Classification: Network & Security

Track Classification: Networking & Security

Contribution ID: 9

Type: **not specified**

ICEPP Site Report

Monday, 3 November 2025 14:10 (20 minutes)

The International Center for Elementary Particle Physics (ICEPP) operates a WLCG Tier-2 site that provides essential computing resources for the ATLAS experiment.

This talk will present the current operational status of the site, covering hardware specifications, global network connectivity, recent operational developments, and ongoing R&D activities.

Speaker release

Yes

Author: SAITO, Masahiko (University of Tokyo (JP))

Presenter: SAITO, Masahiko (University of Tokyo (JP))

Session Classification: Site Report

Track Classification: Site Reports

Contribution ID: 10

Type: **not specified**

Recent developments of the tape infrastructure at CNAF

Tuesday, 4 November 2025 15:30 (30 minutes)

INFN-CNAF is the National Center of INFN (National Institute for Nuclear Physics) dedicated to research and development in the field of information technologies applied to subnuclear, nuclear and astro particle physics. CNAF hosts the largest INFN data center, which also includes a WLCG Tier1 site.

For more than 15 years, the Grid Enabled Mass Storage System (GEMSS), an in-house solution, has been adopted for managing tape data at CNAF. GEMSS middleware is based on a custom integration of the IBM Storage Scale file system with a tape backend powered by IBM Storage Protect, which provides Hierarchical Storage Management (HSM), also combined with the Grid Storage Resource Manager (StoRM) via the StoRM WebDAV and StoRM Tape REST API services.

After presenting a general overview of the whole tape infrastructure, and some interesting recent developments, we will focus on how the workflow has evolved to support any number of libraries and to efficiently use the NVMe partitions of the tape buffer.

Finally, we will describe how these changes proved particularly useful during the migration to the new INFN-CNAF Data Center at Bologna Tecnopolo.

Speaker release

Yes

Author: RENDINA, Andrea

Co-authors: ALKHANSA, Ahmad (INFN - CNAF); CAVALLI, Alessandro; PROSPERINI, Andrea (INFN - CNAF); LUPU, Daniel (INFN-LNL); MORGANTI, Lucia; SAPUNENKO, Vladimir (INFN - CNAF)

Presenter: RENDINA, Andrea

Session Classification: Storage & data management

Track Classification: Storage & data management

Contribution ID: 11

Type: **not specified**

KIT Site Report

Monday, 3 November 2025 11:40 (20 minutes)

We'll give an overview of current activities around the GridKa Tier-1 including updates on the compute, disk and tape resources and the internal network setup. In addition there are news on the HPC systems and the Large Scale Data Facility at KIT.

Speaker release

Yes

Author: SCHNEPF, Matthias Jochen

Co-author: PETZOLD, Andreas (KIT - Karlsruhe Institute of Technology (DE))

Presenter: SCHNEPF, Matthias Jochen

Session Classification: Site Report

Track Classification: Site Reports

Contribution ID: 12

Type: **not specified**

Harnessing restrictive HPC environments with *fapptainer* for complex data processing workloads

Monday, 3 November 2025 15:50 (30 minutes)

The effective processing of experimental and simulated HEP data at scale on restrictive High Performance Computing (HPC) systems poses a known challenge to the physics communities. A significant aspect of such challenge is of technical nature. Many of the barriers have been successfully overcome or mitigated thanks to several integration projects over the last decade. Containerisation techniques have conspicuously helped in such achievements, yet several difficulties exist on a system-by-system basis, so the access threshold remains high. In this work, we address the impediment of running nested containers on systems where security policies prevent it. A novel tool is presented, *fapptainer*, that implements the “un-nesting” of the containers on the fly, running them sideways instead, in compliance with local security policies and without any modification to the workflow of the jobs. It runs unprivileged, thus not requiring system modifications by the local sysadmins. Other crucial aspects are addressed in the integration scheme, including remote access to the local job scheduler and software provisioning, leveraging open source tools like the Advanced Resource Connector (ARC) middleware, CernVM-FS (CVMFS), SSH Filesystem (SSHFS). The implementation, designed to be HPC centre agnostic, has allowed running ATLAS and CMS experiment computational workloads on the LUMI pre-exascale Supercomputer and on the MAHTI and PUHTI HPC systems at CSC in Finland.

Speaker release

Yes

Author: SLIUSAR, Ievgen (University of Oslo (NO))**Co-authors:** SCIACCA, Gianfranco (Universitaet Bern (CH)); LINDÉN, Tomas (Helsinki Institute of Physics (FI))**Presenter:** SCIACCA, Gianfranco (Universitaet Bern (CH))**Session Classification:** Computing & Batch Services**Track Classification:** Computing & Batch Services

Contribution ID: 13

Type: **not specified**

Building Inference-as-a-Service for LLMs on CSCS ALPS

Thursday, 6 November 2025 09:00 (30 minutes)

As part of the SwissAI initiative, CSCS resources were used to train APERTUS, an open-source multilingual language model whose entire development process including its architecture, model weights, training data, and recipes is openly accessible and fully documented.

In this presentation, I will describe the new services that CSCS is developing to deploy this model for inference using Kubernetes and NVIDIA GH200 nodes on our ALPS infrastructure.

Speaker release

Yes

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

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

Session Classification: Cloud Technologies, Virtualization & Orchestration, Operating Systems

Track Classification: Cloud Technologies, Virtualization & Orchestration, Operating Systems

Contribution ID: 14

Type: **not specified**

HEPiX Technology Watch Working Group Report

Tuesday, 4 November 2025 11:00 (1 hour)

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 report, we provide an overview of the hardware technology landscape and some recent developments, highlighting the impact on the HEP computing community.

Speaker release

Yes

Authors: Mr CHIERICI, Andrea (Universita e INFN, Bologna (IT)); Dr SCIABÀ, Andrea (CERN); YEN, Eric (Academia Sinica (TW))

Presenter: YEN, Eric (Academia Sinica (TW))

Session Classification: Environmental sustainability, business continuity, and Facility improvement

Track Classification: Environmental sustainability, business continuity, and Facility improvement

Contribution ID: 15

Type: **not specified**

ASGC site-report

Monday, 3 November 2025 11:20 (20 minutes)

ASGC site-report

Speaker release

Yes

Authors: YEN, Eric (Academia Sinica (TW)); 楊, 翔宇

Presenter: 楊, 翔宇

Session Classification: Site Report

Track Classification: Site Reports

Contribution ID: 16

Type: **not specified**

KEK Site Report

Monday, 3 November 2025 15:00 (20 minutes)

The KEK Central Computer System (KEKCC) is KEK's largest-scale computer system, providing essential services such as Grid and Cloud computing for the High Energy and Nuclear Physics community.

Following the procurement policy for large-scale computer systems, the KEKCC is replaced entirely every four years. The new system entered production in September 2024 and has successfully completed its first year of operation under the four-year contract (decommissioning is scheduled for August 2028).

In this talk, we will report on the operational status and key technical developments during the first year of the new KEKCC.

Key topics will include:

1. **Stable Operation and High Utilisation:** Maintaining stable service with high CPU utilisation (approximately 90%) across various experimental groups.
2. **OS Migration and Network Performance:** The ongoing operating system migration to RHEL9 (completed for components like CVMFS and StoRM data transfer nodes) and the achievement of 80+ Gbps throughput between KEK and Belle II Raw Data Centres for efficient data sharing.
3. **Migrating X.509/VOMS to JWT/IAM:** The launch of the new KEK Grid CA 2024 and the ongoing deployment of an Identity Provider (IdP) as part of the plan to migrate authentication and authorisation from the legacy X.509/VOMS to a modern JWT/IAM.

We will provide details on these efforts to enhance the computing and networking infrastructure supporting KEK's world-leading accelerator experiments.

Speaker release

Yes

Author: IWAI, Go (KEK)

Co-authors: Dr MURAKAMI, Koichi; SUZUKI, Soh; NAKAMURA, Tomoaki; KISHIMOTO, Tomoe

Presenter: IWAI, Go (KEK)

Session Classification: Site Report

Track Classification: Site Reports

Contribution ID: 17

Type: **not specified**

Development and Practice of AI Agents and Framework for Scientific Data Processing

Tuesday, 4 November 2025 16:00 (30 minutes)

2025 is widely recognized as the Year of the AI Agent. Large language models have moved beyond conversational interfaces to become callable tools that boost productivity—evident in the rapid adoption of systems like Manus, Claude-Code, and Cursor. AI Agent technologies are also increasingly being applied in scientific research to assist in data analysis and literature exploration, as demonstrated by systems such as SciMaster, FutureHouse, Machine Chemist, and SciToolAgent.

The Computing Center of the Institute of High Energy Physics (IHEP) initiated research on scientific AI Agents in 2023 and developed Dr.Sai, an intelligent agent for BESIII (Beijing Spectrometer) physics analysis. Building upon this experience, we present OpenDrSai—a scientific AI agent framework designed to accelerate the development and deployment of AI agents for scientific data processing.

OpenDrSai integrates core capabilities including self-learning and reflection, real-time human-agent interaction, long task management, and multi-agent collaboration. The framework offers modular components for multimodal scientific data perception, knowledge and memory management, scientific tool orchestration, and complex workflow execution. It also features a flexible multi-agent architecture, a scalable backend system, an interactive human-machine interface, and standardized APIs. These features address key challenges in scientific AI development, such as integrating complex tools, managing long-running tasks, and handling domain-specific data and knowledge.

OpenDrSai is already deployed or planned for use in several large-scale scientific experiments, including the China Spallation Neutron Source, Beijing Synchrotron Radiation Facility, Large High Altitude Air Shower Observatory (LHAASO), JUNO Neutrino Experiment, and the Deep-Sea Neutrino Telescope. Some specialized agents—such as DataAgent, RongZai Agent, and BOSS8 Assistant—have been developed to support tasks including neutron diffraction and PDF refinement, as well as data processing for large-scale experimental facilities.

Speaker release

Yes

Authors: XIONG, DONGBO (Institute of High Energy Physics, CAS); ZHANG, Zhengde (中国科学院高能物理研究所); QIFAZHI, 齐法制

Presenter: XIONG, DONGBO (Institute of High Energy Physics, CAS)

Session Classification: Storage & data management

Track Classification: Storage & data management

Contribution ID: 18

Type: **not specified**

The development and upgrades of BESIII Offline Software System

Tuesday, 4 November 2025 09:30 (30 minutes)

The BESIII experiment has been operating since 2009 and received several upgrades, to study the τ -charm physics utilizing the BEPCII accelerator. The offline software system is the fundamental tool for physics analysis. It is developed to process the raw data from BESIII detector and Monte Carlo data from simulation, produce the reconstructed data which contains various physics information of the primary and secondary particles. This presentation focuses on the development and upgrades of offline software to meet the requirements from various aspects.

Speaker release

Yes

Author: Prof. DENG, Ziyang (Institute of High Energy Physics, Chinese Academy of Sciences)

Presenter: Prof. DENG, Ziyang (Institute of High Energy Physics, Chinese Academy of Sciences)

Session Classification: Computing & Batch Services

Track Classification: Software and Services for Operation

Contribution ID: 19

Type: **not specified**

Building the data system for the application of light-source facilities and AI4S research

Tuesday, 4 November 2025 15:00 (30 minutes)

During the operation and scientific research of large-scale scientific facilities such as synchrotron light sources (BSRF, HEPS) and spallation neutron sources (CSNS), the massive, complex, and heterogeneous data are generated continuously. Enabling systematic management and efficient use of facility data; experimental data (diffraction, scattering, imaging, spectroscopy, etc.); simulation data; and textual literature data has become a pivotal challenge that support both scientific research and industrial applications. As artificial intelligence and data-driven research paradigms advance, it is imperative to build an end-to-end data system that spans the full data lifecycle and serves AI4S (AI for Science). This report introduces an integrated “Policies & Standards - Software & Tools - Services & Providing” framework across the collect-store-process-analyze-use lifecycle. The aim is to improve data quality, accessibility, and reusability, ultimately yielding high-quality scientific databases and data products. The system is grounded in unified policies and standards, supported by both general-purpose and domain-specific software and tools, and delivered through diversified service channels. It enables standardized management and efficient circulation of multi-type, multimodal, and cross-disciplinary data, meeting the combined needs of facility operations, scientific innovation, and industrial application. Data policies and standards include: file-format specifications for different data types; metadata standards that cover facilities, experiments, simulation, and literature, especially scientific metadata standards of experiments; traceable data-processing pipelines and methods; and institutionalized designs for long-term data preservation and usage. We are currently collaborating with multiple domestic institutions to apply for a national standard titled “Metadata for Photon and Radiation Experimental Data”, in order to promote coordinated development and enhanced interoperability across large facilities in China. The software-and-tools layer is built and integrated around principles of standardization and reproducibility: the mamba data-acquisition software; the Domas data-management software framework; the Daisy data-processing software framework; and platforms for simulation and literature-driven knowledge. These tools are tightly coupled to the data standards to ensure normative and consistent data production, management, and processing. In parallel, “data agents” are introduced to automate and augment key stages-data cleaning, fusion and alignment, analysis and processing, and knowledge extraction-thus providing a high-quality data foundation for AI4S model training and inference. On the services and providing side, the system conducts cross-modal data fusion and alignment tailored to different experimental methodologies and simulation data; constructs domain knowledge graphs from textual literature to enable knowledge organization and semantic search; and offers multiple channels for data access and use, including the Domas service portal, a data-download client, RESTful APIs, and llm-based intelligent Q&A interfaces. Building on systematic data cleaning, fusion, and processing, we curate high-quality datasets for diverse disciplinary research scenarios and support unified access, download, and use via the data-fusion platform and AI platform-forming productized data outputs that directly support research and applications. Through the integrated coordination of “Policies & Standards - Software & Tools - Services & Provisioning,” this system provides end-to-end support from data production to application for the multimodal, complex data generated by large scientific facilities. It markedly enhances data management and utilization, delivering sustainable data system for the AI4S research paradigm and for facility-driven fundamental scientific innovation.

Speaker release

Yes

Authors: HU, Hao (Institute of High Energy of Physics); HUP, 胡鹏; QIFAZHI, 齐法制

Presenter: HUP, 胡鹏

Session Classification: Storage & data management

Track Classification: Storage & data management

Contribution ID: 20

Type: **not specified**

HEPiX Benchmarking Working Group: Status Update

Monday, 3 November 2025 15:20 (30 minutes)

A large distributed computing environment with different groups of interest (funding agencies, resource providers, users), such as the WLCG, requires a score to define the amount of needed and provided computing resources.

The HEPiX Benchmarking Working Group has developed and provided benchmarks for high-energy physics (HEP) for years.

In addition to the classic CPU benchmark, we are currently developing a GPU benchmark for the HEP. With the HEP Benchmark suite, additional metrics such as, load and power consumption can be recorded during the benchmark. These metrics can help find the best hardware or optimize existing hardware in various directions, such as energy efficiency.

We give an overview of the working group, the features of the provided tools, and the status of the development of the GPU benchmark.

Speaker release

Yes

Author: SCHNEPF, Matthias Jochen

Presenter: SCHNEPF, Matthias Jochen

Session Classification: Computing & Batch Services

Track Classification: Computing & Batch Services

Contribution ID: 23

Type: **not specified**

Computer Security Update

Wednesday, 5 November 2025 15:30 (30 minutes)

This presentation aims to give an update on the global security landscape from the past months. The global political situation has introduced a novel challenge for security teams everywhere. 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.

Speaker release

Yes

Author: LUNA DURAN, Jose Carlos (CERN)**Presenter:** LOPIENSKI, Sebastian (CERN)**Session Classification:** Network & Security**Track Classification:** Networking & Security

Contribution ID: 24

Type: **not specified**

Investigating Routing Anomalies and Performance Degradation in WLCG Networks (Case Studies)

Wednesday, 5 November 2025 14:00 (30 minutes)

This presentation reports on the analysis of network incidents in WLCG infrastructures utilizing data from perfSONAR. We examine a few case studies of reported incidents in the past three years, which involved routing path changes and performance impacts across site pairs, highlighting IP-level anomalies and their correlation with throughput and latency metrics. Through graph-based anomaly detection and topology-performance correlation, we identify patterns of routing instability, such as path inflation and asymmetry through statistical baselining, and their temporal alignment with performance degradation. Findings reveal that routing changes often coincide with performance issues. This work demonstrates the value of integrated data analysis for proactive network management in WLCG collaborations.

Speaker release

Yes

Authors: BABIK, Marian (CERN); VASILEVA, Petya (University of Michigan (US)); MCKEE@UMICH.EDU, Shawn (University of Michigan)

Presenter: VASILEVA, Petya (University of Michigan (US))

Session Classification: Network & Security

Track Classification: Networking & Security

Contribution ID: 25

Type: **not specified**

Accounting with AUDITOR: Experiences and News

Tuesday, 4 November 2025 10:00 (30 minutes)

AUDITOR is a flexible framework designed to address the challenges of managing large volumes of accounting data across diverse computing environments, including institute clusters, Grid sites, and Kubernetes clusters. Due to its modular design, it is possible to gather job information from various batch systems. AUDITOR effectively stores this data in a database for easy access and processing. It also enables publishing the collected accounting data to external services such as APEL or the use of Python notebooks for in-house analysis.

Besides the accounting of CPU performance in HEPSPC06 and HEPSCore23, it is possible to account for other scores, such as GPU performance or power usage.

Electricity prices or dynamic CPU core power are critical metrics for sustainable computing. However, these metrics can change over the runtime of a job. Therefore, we want to expand AUDITOR for an accurate accounting of dynamic values.

This presentation will show AUDITOR's ecosystem structure, existing applications, and highlight its potential in advancing sustainable computing solutions.

Speaker release

Yes

Author: SCHNEPF, Matthias Jochen

Co-authors: SAMMEL, Dirk (University of Freiburg (DE)); VON CUBE, Florian (KIT - Karlsruhe Institute of Technology (DE)); GIFFELS, Manuel (KIT - Karlsruhe Institute of Technology (DE)); Dr KÜHN, Max (Karlsruhe Institute of Technology); BOEHLER, Michael (University of Freiburg (DE)); VIJAYAKUMAR, Raghuvar (University of Freiburg (DE)); KLEINEMUEHL, Raphael

Presenter: SCHNEPF, Matthias Jochen

Session Classification: Computing & Batch Services

Track Classification: Computing & Batch Services

Contribution ID: 27

Type: **not specified**

Design of Torch Computing Platform and The Platform Application in HEPS

Thursday, 6 November 2025 09:30 (30 minutes)

China's High-Energy Photon Source (HEPS)—the country's first national high-energy synchrotron radiation light source—is currently in the design and construction phase. The HEPS Computing Center serves as the core provider of high-performance computing (HPC) resources, data resources, and technical services for HEPS scientific experiments. The overarching mission of the HEPS scientific computing platform is to accelerate scientific discovery in light source experiment-related research through advanced HPC and data analysis capabilities. To address the diverse data analysis needs across light source disciplines, a dedicated scientific computing platform has been developed to deliver multi-modal computing services, including desktop analysis, interactive analysis, and batch analysis. This platform enables scientists to access the computing environment via a web-based interface anytime and anywhere, facilitating rapid analysis of experimental data. This paper presents the design of a scientific computing platform—the Torch Computing Platform—tailored to meet HEPS's diverse analysis requirements. First, it elaborates on the varied analysis demands of HEPS research scenarios. Second, it identifies and analyzes the key challenges faced by the HEPS scientific computing system. Third, from a user-centric perspective, it details the architecture and service workflow of the Torch Computing Platform, while providing in-depth explanations of several critical technical implementations. Finally, it demonstrates the practical application effects and performance of the platform, validating its effectiveness in supporting HEPS scientific research.

Speaker release

Yes

Author: HU, Qingbao (IHEP)**Presenter:** HU, Qingbao (IHEP)**Session Classification:** Cloud Technologies, Virtualization & Orchestration, Operating Systems**Track Classification:** Cloud Technologies, Virtualization & Orchestration, Operating Systems

Contribution ID: 28

Type: **not specified**

New feature implementation of Interactive aNalysis worKbench (INK)

Wednesday, 5 November 2025 11:30 (30 minutes)

Ink (Interactive Analysis Workbench) is a self-developed software tool by the Computing Center of IHPE. It provides seamless access to cluster computing and storage resources through flexible API interfaces. Since its initial release and deployment on IHEP computing platform in March this year, Ink has been well-received by its users.

This report will focus on the introduction of several key recently implemented features in Ink, detailing their technical solutions and implementation outcomes. The highlighted features include: the file sharing mechanism, enhanced support for multiple authentication methods, and the “User Job Resource Consumption Tracking” system.

Speaker release

Yes

Author: SHI, Jingyan (IHEP)

Presenter: SHI, Jingyan (IHEP)

Session Classification: Software and Services for Operation

Track Classification: Software and Services for Operation

Contribution ID: **30**

Type: **not specified**

IHEP Site Report

Monday, 3 November 2025 11:00 (20 minutes)

IHEP Site Report.

Speaker release

Yes

Author: Mr JIANG, Xiaowei (IHEP (中国科学院高能物理研究所))

Presenter: Mr JIANG, Xiaowei (IHEP (中国科学院高能物理研究所))

Session Classification: Site Report

Track Classification: Site Reports

Contribution ID: 31

Type: **not specified**

An AI multi-agent system for physics analysis

Friday, 7 November 2025 11:00 (30 minutes)

The data processing and analyzing is one of the main challenges at HEP experiments. To accelerate the physics analysis and drive new physics discovery, the rapidly developing Large Language Model (LLM) is the most promising approach, it have demonstrated astonishing capabilities in recognition and generation of text while most parts of physics analysis can be benefitted. In this talk we will discuss the construction of a dedicated intelligent agent, an AI assistant names Dr.Sai at BESIII based on LLM, the potential usage to boost hadron spectroscopy study, and the future plan towards a AI scientist.

Speaker release

Yes

Author: SHANG, Zijie (Lanzhou University)**Co-author:** LI, Ke**Presenter:** SHANG, Zijie (Lanzhou University)**Session Classification:** Miscellaneous**Track Classification:** Miscellaneous

Contribution ID: 32

Type: **not specified**

Elastic Resource Expansion and Scheduling Optimization at IHEP

Monday, 3 November 2025 16:20 (30 minutes)

We built a coordinated scheduling layer across our local cluster and grid resources. Using glideins, idle grid capacity is elastically federated into our local HTCondor pool, while busy periods trigger proactive glidein removal to promptly return resources—achieving cross-domain elastic scale-out/in. On the local side, we introduced performance-aware labeling: nodes are tiered (fast/medium/slow) by CPU microarchitecture, and annotated in ClassAds with local disk I/O, network, and storage capabilities. Jobs can then express precise constraints to target suitable resources, reducing mismatch-induced failures, increasing utilization and throughput, and making queueing and runtime more predictable.

Speaker release

Yes

Author: GUO, Chaoqi (Institute of High Energy Physics of the Chinese Academy of Sciences)

Co-authors: SHI, Jingyan (IHEP); Mr JIANG, Xiaowei (IHEP (中国科学院高能物理研究所))

Presenter: GUO, Chaoqi (Institute of High Energy Physics of the Chinese Academy of Sciences)

Session Classification: Computing & Batch Services

Track Classification: Computing & Batch Services

Contribution ID: 33

Type: **not specified**

GNN application for BESIII/STCF tracking

Thursday, 6 November 2025 14:00 (30 minutes)

Track reconstruction is one of the most important and challenging tasks in the offline data processing of collider experiments. The BESIII and Super Tau-Charm Facility (STCF) is the current and a next-generation electron-positron collider running in the tau-charm energy region proposed in China, where conventional track reconstruction methods face challenges from the higher background environment especially for STCF.

In this contribution, we demonstrate a novel hybrid tracking algorithm based on Graph Neural Network (GNN) method and traditional methods for the BESIII/STCF drift chamber. In the GNN method, a hit pattern map representing the connectivity between drift cells is constructed considering the geometrical layout of the sense wires, based on which we design an optimal graph construction method, then an edge-classifying graph neural network is trained to distinguish the hit-on-track from noise hits. Finally, the result after the noise filtering is integrated into tracking software where a track-finding algorithm based on the DBScan (for BESIII) or Hough transform (for STCF) is performed and a track-fitting algorithm based on GENFIT is used to obtain the track parameters.

Preliminary results based on the MC sample, showing promising performance. Furthermore, the GNN based noise filtering algorithm can also be potentially applied to other collider experiments with similar drift chamber based trackers.

Speaker release

Yes

Authors: JIA, Xiaoqian (Shandong University); QIN, Xiaoshuai (Shandong University (CN)); LI, Teng (Shandong University); HU, Xiaoqian (Shandong University); ZHOU, Qidong (Shandong University); SONG, Shuangbing (Shandong University); HUANG, Xingtao (Shandong University); ZHANG, Xueyao (Shandong University); ZHANG, Yao (IHEP); YUAN, Ye (IHEP)

Presenter: QIN, Xiaoshuai (Shandong University (CN))

Session Classification: Science talks

Track Classification: Miscellaneous

Contribution ID: 34

Type: **not specified**

WLCG site monitoring with ps-dash: insights, alarms, and visualization for network performance

Wednesday, 5 November 2025 16:30 (30 minutes)

We will present the latest updates in **WLCG site network monitoring** through **ps-dash**, a web-based dashboard designed to visualize and analyze network performance data collected by perfSONAR. The tool presents alarms derived from alerting data measurements to help identify **site-specific network issues**. It reports on routing anomalies, packet loss, variations in bandwidth, and throughput that may indicate bottlenecks or congestion, etc., provides visibility on Tier-1 site connectivity and perfSONAR test coverage across hosts. Particular focus is given to the site report page, which consolidates per-site statistics and alarm history —offering an intuitive interface for observability and performance comparison. Together, these developments enhance the visibility, reliability, and diagnostic capacity of the WLCG network infrastructure.

Speaker release

Yes

Author: HOLOBORODKO, Yana (IRIS-HEP, Princeton University (US))

Co-authors: BABIK, Marian (CERN); VASILEVA, Petya (University of Michigan (US)); MC KEE, Shawn (University of Michigan (US))

Presenter: HOLOBORODKO, Yana (IRIS-HEP, Princeton University (US))

Session Classification: Network & Security

Track Classification: Networking & Security

Contribution ID: 35

Type: **not specified**

Empowering early-career researchers through AI-driven studies on computing infrastructure at PIC

Friday, 7 November 2025 09:30 (1 hour)

The growing complexity and scale of modern scientific computing infrastructures, such as the Port d'Informació Científica (PIC), a Tier-1 center within the Worldwide LHC Computing Grid (WLCG), require continuous optimization to maintain performance, reliability, and energy efficiency. Artificial Intelligence (AI) and Machine Learning (ML) techniques provide powerful means to tackle these challenges by enabling predictive and adaptive management of computational resources.

This contribution presents ongoing and prospective research directions focused on identifying and addressing operational challenges in large-scale computing environments through data-driven approaches. Illustrative examples include predicting the reduction in resource/cores utilization during compute-farm drainage periods, enhancing data cache management via intelligent eviction policies beyond traditional LRU mechanisms, identifying “hot” files to dynamically migrate them to higher-performance storage tiers (such as SSDs), and forecasting job execution times to improve scheduling and throughput.

Beyond the technical scope, this initiative aims to engage young researchers and students through short, well-defined projects that offer hands-on experience with real operational data and infrastructure from the PIC center. In addition, a set of future project ideas will be presented to inspire new collaborations and exploratory studies. By bridging educational efforts with practical challenges, this initiative seeks to foster innovation, attract emerging talent to the field of scientific computing, and contribute to the sustainable advancement of large-scale computing facilities within the WLCG ecosystem.

Speaker release

Yes

Author: FLIX MOLINA, Jose (CIEMAT - Centro de Investigaciones Energéticas Medioambientales y Tec. (ES))

Presenter: FLIX MOLINA, Jose (CIEMAT - Centro de Investigaciones Energéticas Medioambientales y Tec. (ES))

Session Classification: Miscellaneous

Track Classification: Miscellaneous

Contribution ID: 36

Type: **not specified**

CENI Briefing

Wednesday, 5 November 2025 16:00 (30 minutes)

China Environment for Network Innovations (CENI) is China's first national major science and technology infrastructure in the field of communications and information. It is an open, user-friendly, and sustainable large-scale general test facility, which can provide a simple, efficient, and low-cost test and verification environment for researching the innovative architecture of future networks.

This facility supports China in achieving major breakthroughs in key aspects of network science and cyberspace technology research, such as key equipment, network operating systems, routing control technologies, network virtualization technologies, secure and trusted mechanisms, and innovative service systems. Moreover, it can verify network scenarios suitable for Internet operation and services, and explore the technical routes and development paths of future network development.

Speaker release

Yes

Author: WU, Bingqing (Jiangsu Future Network Innovation Institute)

Presenter: WU, Bingqing (Jiangsu Future Network Innovation Institute)

Session Classification: Network & Security

Track Classification: Networking & Security

Contribution ID: 37

Type: **not specified**

Ceph Storage for the Tier1

Tuesday, 4 November 2025 16:30 (30 minutes)

This talk provides a quick overview of Echo, the largest disk storage cluster operated at STFC's Scientific Computing Department. Echo is the storage cluster supporting the UK's WLCG Tier 1.

Echo has been running continuously for over eight years and has scaled to 137PiB.

We will share key lessons learned from managing a cluster of this scale, including best practices, operational challenges, and practical tips for maintaining reliability and performance over time.

In particular, we will discuss the issue of past OSD maps building up in the monitor and OSD DBs. If too many are accumulated, these can degrade cluster performance. We will detail the strategies and techniques we use to monitor and avoid this scenario.

Finally, we will outline our future plans for Echo, highlighting upcoming improvements and software and hardware upgrades and how we intend to develop the service to meet growing data demands.

Speaker release

Yes

Author: Mr ABUAJAMIEH, Maksim (UKRI)

Presenter: Mr ABUAJAMIEH, Maksim (UKRI)

Session Classification: Storage & data management

Track Classification: Storage & data management

Contribution ID: 38

Type: **not specified**

Computing challenges of HEP experiments: the LHCb case

Tuesday, 4 November 2025 14:00 (30 minutes)

Achieving higher precision of particle physics observables from large part depends on ever-increasing amount of recorded collisions. Together with improved sensitivity and increasing granularity of detectors, which leads into online processing of O(TBit/s) data with output as O(GBit/s). Such a conditions are leading into strict computing requirements for current and future HEP experiments.

This talk will overview existing HEP computing challenges, both online and offline computing, mainly on example of LHCb experiment as the first LHC experiment with a full software trigger. Furthermore, challenges for HL-LHC era and LHCb Upgrade II will be discussed.

Speaker release

Yes

Author: SAUR, Miroslav (Lanzhou University (CN))**Presenter:** SAUR, Miroslav (Lanzhou University (CN))**Session Classification:** Science talks**Track Classification:** Miscellaneous

Contribution ID: 41

Type: **not specified**

JUNO Status

Thursday, 6 November 2025 13:30 (30 minutes)

Speaker release

Presenter: Dr JIA, Yi (IHEP)

Session Classification: Science talks

Track Classification: Miscellaneous

Contribution ID: 42

Type: **not specified**

Data analysis and Software at LHAASO

Tuesday, 4 November 2025 13:30 (30 minutes)

Presenter: Dr YIN, Liqiao (IHEP)

Session Classification: Science talks

Track Classification: Miscellaneous

Contribution ID: 43

Type: **not specified**

Welcome

Monday, 3 November 2025 09:30 (15 minutes)

Presenter: NAKAMURA, Tomoaki

Session Classification: Welcome

Track Classification: Miscellaneous

Contribution ID: 44

Type: **not specified**

Introduction to Lanzhou University

Monday, 3 November 2025 09:45 (30 minutes)

Speaker release

Presenter: Prof. LIU, Zhi-yi (Lanzhou University)

Session Classification: Welcome

Track Classification: Miscellaneous

Contribution ID: 45

Type: **not specified**

Logistics Information

Monday, 3 November 2025 10:15 (15 minutes)

Presenter: XIAO, Dong (Lanzhou University)

Session Classification: Welcome

Track Classification: Miscellaneous

Contribution ID: 46

Type: **not specified**

Wrap-up

Friday, 7 November 2025 11:30 (30 minutes)

Presenter: NAKAMURA, Tomoaki

Session Classification: Wrap-up

Track Classification: Miscellaneous

Contribution ID: 47

Type: **not specified**

Current status and future perspectives of projects involving Asian sites

Thursday, 6 November 2025 11:00 (30 minutes)

Presenter: NAKAMURA, Tomoaki

Session Classification: (Topical session) Current status and future perspectives of projects involving Asian sites

Contribution ID: 48

Type: **not specified**

discussions

Thursday, 6 November 2025 11:30 (30 minutes)

Session Classification: (Topical session) Current status and future perspectives of projects involving Asian sites