ISBN: 978-3-95450-250-9 ISSN: 2673-5520 doi:10.18429/JACoW-MEDSI2023-WEOBM03

THE DESIGN AND PROGRESS OF THE NETWORK AND COMPUTING SYSTEM FOR HEPS

H. Hu, F. Qi, X. Wang, Y. Cheng, Q. Hu, H. Zhang, Y. Hu, S. Zeng
Institute of High Energy of Physics, Beijing, China

Abstract

China's High Energy Photon Source (HEPS) is the first national high-energy synchrotron radiation light source. The 14 beamlines for the phase I of HEPS will produces about 300 PB/year raw data, it represents significant challenges in data storage, data access, data analysis and data exchange. HEPS Computing and Communication System (HEPSCC), is an essential work group responsible for the IT R&D and services for the facility, including IT infrastructure, network, computing, analysis software, data preservation and management, public services etc. This paper mainly introduces the design and progress of HEPSCC's work in addressing the data challenges faced by HEPS from various aspects, including machine room, network, storage, computing, and scientific software of data management and data analysis.

MANUSCRIPTS

HEPS is the first national high-energy synchrotron radiation light source and soon one of the world's brightest fourth-generation synchrotron radiation facilities [1], has been constructed from 2019 in Beijing's Huairou District, and will be completed in 2025. The 14 beamlines for the phase I of HEPS will produces about 300 PB/year raw data (see Table 1). Efficiently storing, analyzing, and sharing this huge amount of data presents a significant challenge for HEPS.

Table 1: Estimated Data Volume of HEPS at Phase I

Beamlines	Burst output(Byte/day)	Average output(Byte/day)
B1 Engineering Materials Beamline	600TB	200TB
B2 Hard X-ray Multi-analytical Nanoprobe (HXMAN) Beamline	500TB	200TB
B3 Structural Dynamics Beamline (SDB)	8TB	3TB
B4 Hard X-ray Coherent Scattering Beamline	10TB	3TB
B5 Hard X-ray High Energy Resolution Spectroscopy Beamline	10TB	1TB
B6 High Pressure Beamline	2TB	1TB
B7 Hard X-Ray Imaging Beamline	1000TB	250TB
B8 X-ray Absorption Spectroscopy Beamline	80TB	10TB
B9 Low-Dimension Structure Probe (LODISP) Beamline	20TB	5TB
BA Biological Macromolecule Microfocus Beamline	35TB	10TB
BB pink SAXS	400TB	50TB
BC High Res. Nanoscale Electronic Structure Spectroscopy Beamline	1TB	0.2TB
BD Tender X-ray beamline	10TB	1TB
BE Transmission X-ray Microscope Beamline	25TB	11.2TB
BF Test beamline	1000TB	60TB
Total average:		805.4TB/day, 24.16PB/month

HEPS Computing and Communication System (HEPSCC), also called HEPS Computing Center, is an essential work group responsible for the IT R&D and services for the facility, including IT infrastructure, network, computing, analysis software, data preservation and management, public services etc. Aimed at addressing the significant challenge of large data volume, HEPSCC has designed and established a network and computing system, making great progress over the past two years.

As the most fundamental part of the IT infrastructure, a deliciated and high-standard machine room, with about 900 m^2 floor space for more than 120 high-density racks in total, has been prepared for production since this August. The power system has two transformers for dual power supply and has a total capacity of 2,500KVA, with the UPS providing 800KVA of power capacity and offering a half-hour backup during emergencies. Row-Air conditioning with natural cooling is used for the refrigeration of the machine room, which can greatly reduce the energy consumption.

For the data center network, we designed it as a spine-leaf architecture which makes it very easy to scale out. The backbone bandwidth of the data center network can support speeds up to 4*400 Gb/s, which can fully meet the demands of high-speed data exchange. Meanwhile, we also support RoCE [2] (RDMA over Converged Ethernet) to provide a lossless and high-performance network environment for scientific workload in HEPS data center. Previous test evaluations showed that RoCE can reach the same performance as InfiniBand (IB) in both point-to-point and collective tests.

In order to balance the cost-effectiveness of storage devices and realize the high reliability of data storage, a three-tier storage is designed for storing experimental data, including beamline storage, central storage, and tape. There is a storage policy for data preservation (see Fig. 1), the raw data and processed data are stored on the beamline storage for a maximum of 7 days, on the central storage for a maximum of 90 days, and only the raw data are archived to tape for long-term storage with two copies. Of course, this data storage policy could be adjusted according to the actual data volume and funding situation of HEPS. The beamline storage utilizes distributed all-flash SSD arrays to achieve high data input/output speeds, while offering a total storage capacity of 800 TB. The central storage leverages distributed high-density HDD arrays to get medium-high speed data IO, providing a total capacity of 30 PB. The tape storage is compliant with the LTO9 [3] standard, and provides 2 PB at the first stage although we have no budget for tape.

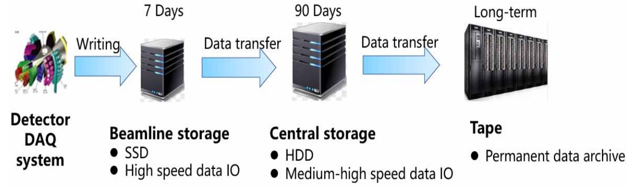

Figure 1: Storage policy for HEPS data.

To meet the requirements of data analysis scenarios for HEPS, a computing architecture has been designed and deployed in three types (see Fig. 2), including Openstack [4], Kubernetes, and Slurm. Openstack integrates the virtual cloud desktop protocol to provide users with remote

Content from this work may be used under the terms of the CC-BY-4.0 licence (© 2023). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI

ISBN: 978-3-95450-250-9 ISSN: 2673-5520 doi:10.18429/JACoW-MEDSI2023-WEOBM03

desktop access services, and supports users to use browsers to access windows/Linux desktop, running commercial visualization data analysis software. Kubernetes manages container clusters, and starts container images with multiple methodological software according to user analysis requirements [5]. Slurm is used to support HPC computing services and meet users' offline data analysis needs.

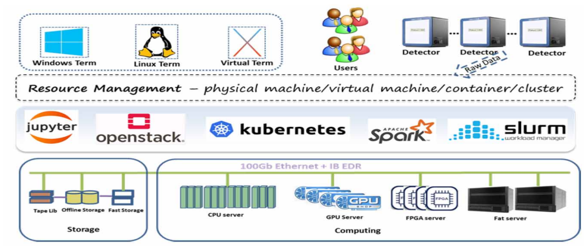

Figure 2: the computing architecture.

In addition to providing the IT infrastructure, HEPSCC also designed and developed the software for the data management and analysis, DOMAS and DAISY.

DOMAS (Data Organization, Management and Accessing Software stack), which is aimed for automating the organization, transfer, storage, distribution and sharing of the scientific data for HEPS experiments, provides the features and functions for metadata catalogue, metadata acquisition, data transfer, data web portal. The metadata catalogue module is responsible for storing metadata into database and providing RESTful APIs to access metadata. The data transfer module facilitates the automatic migration of data between different levels of storage media. The metadata acquisition and processing module is responsible for obtaining metadata from different stages and systems involved in data management and integrating them into the database. The data service module provides users with a web portal for data access, viewing, downloading, and offline analysis.

The core functional module of DOMAS has been completed, and we are progressively open-sourcing each module. Combined with the specific scientific data policy of HEPS [6], standardized data file formats, and data storage directory designs, we have completed the development of the scientific data management system for HEPS. Meanwhile, the system implemented the fully automated movement and management of data among three-tiered storage (beamline storage, central storage, and tape).

The Daisy (Data Analysis Integrated Software System) [7,8] has been designed for the data analysis and visualization of X-ray experiments. Daisy framework consists of a highly modular C++/Python architecture composed of four pillars (see Fig. 3): algorithm, workflow, workflow engine and data store. The algorithm defines the scientific domain model and support the integration of third-party library. The workflow is a sequence of algorithms and defines the scientific domain architecture. The workflow engine manages the runtime environment of workflow. The data store manages the data objects. The architecture supporting customized plug-in functions can easily access visualization tools and the Python-based scientific computing ecosystem.

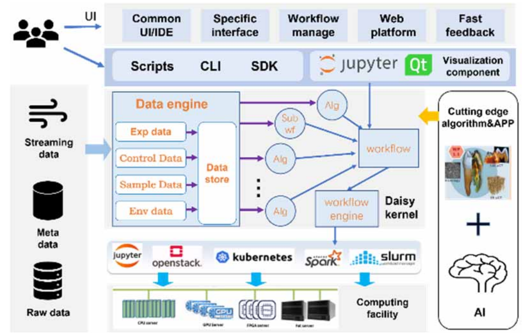

Figure 3: The architecture of DAISY.

Based on the Daisy framework, several specific scientific applications have been developed. For example, HEPSCT integrated an in-house developed 3D tomographic reconstruction package, cumopy, into the framework. It can conduct real-time computing to remove the background, control the white balance, carry motion correction, detect the position of rotation axis according to the parameters setting, then the reconstruction algorithms are executed on GPUs. Daisy-BMX is an AI-based biological macromolecular data-processing application. It implemented a data processing pipeline from diffraction to structure, including the real-time data processing, AI-based structure prediction, and model refinement. The Pair Distribution Function (PDF) implemented a pipeline from diffraction to PDF, include the pre-processing, azimuth integration, PDF transition and post-process. In addition, DAISY implemented and integrated several X-ray absorption spectrum analysis applications, include spectrum matching application and PCA/LCF spectrum component analysis application. Other data analysis algorithms/software will be continuously integrated to the framework in the future.

In 2021, A testbed was set up at beamline 3W1 of BSRF, which is a running synchrotron radiation facility and provides the technology R&D and test platforms for HEPS. The 3W1 beamline, which is dedicated to test high-throughput instruments for HEPS. It is an ideal candidate to set up the testbed where we can deploy the system and verify the functions and the whole process of data acquisition, data processing, data transfer, data storage and data access. The integration and the verification of the whole HEPSCC system at 3W1 beamline achieved great success. It strongly proved the rationality of the design scheme and the feasibility of the technologies. After the optimization and upgrade of the functionality, in July 2022, the HEPSCC system were deployed at 4W1B, which is a running beamline at BSRF, can provide full process service for beamline users.

Now HEPSCC entered the stage of equipment installation and system debugging. The machine room is prepared and provides the operation environment for another IT equipment. The campus network and data center network have been ready since this September, awaiting access from other devices. The software for data management and analysis are deploying and testing.

Content from this work may be used under the terms of the CC-BY-4.0 licence (© 2023). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI

ISBN: 978-3-95450-250-9 ISSN: 2673-5520 doi:10.18429/JACoW-MEDSI2023-WEOBM03

REFERENCES

[1] Y. Jiao *et al.*, "The HEPS project", *J. Synchrotron Radiat.*, vol. 25, pp. 1611-1618, Nov. 2018. doi:10.1107/S1600577518012110

[2] S. Zeng *et al.*, "Research and Evaluation of RoCE in IHEP Data Center", *EPJ Web Conf.*, vol. 251, p. 02018, 2021. doi:10.1051/epjconf/202125102018

[3] LTO9, https://www.lto.org/lto-9/

[4] A. Celesti *et al.*, "Improving desktop as a Service in OpenStack", in *Proc. 2016 IEEE Symposium on Computers and Communication (ISCC)*, Messina, Italy, Jun. 2016. doi:10.1109/ISCC.2016.7543755

[5] J. Stubbs *et al.*, "Integrating jupyter into research computing ecosystems: Challenges and successes in architecting jupyterhub for collaborative research computing ecosystems", *Practice and Experience in Advanced Research Computing*, pp. 91-98. Jul. 2020. doi:10.1145/3311790.3396648

[6] H. Hu, F. Qi, H. Zhang, H. Tian, and Q. Luo, "The design of a data management system at HEPS", *J. Synchrotron Radiat.*, vol. 28, pp. 169-175, Jan. 2021. doi:10.1107/S1600577520015167

[7] Y. Hu *et al.*, "Daisy: Data analysis integrated software system for X-ray experiments", *EPJ Web Conf.*, vol. 251, p. 04020, Aug. 2021. doi:10.1051/epjconf/202125104020

[8] Y. Hu and H. Tian, "Data Analysis Integrated Software System in IHEP, design and implementation", in *Proc. International Symposium on Grids & Clouds 2022 (ISGC2022),* Taipei, Taiwan*,* March 2022. doi:10.22323/1.415.0028

Content from this work may be used under the terms of the CC-BY-4.0 licence (© 2023). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI