高能同步辐射光源
High Energy Photon Source | HEPS

高能所計算中心
IHEP Computing Center

Design and implementation of HEPS scientific computing system for various interactive data analysis scenarios

IHEP-CC Qingbao Hu

2024.3

Outline

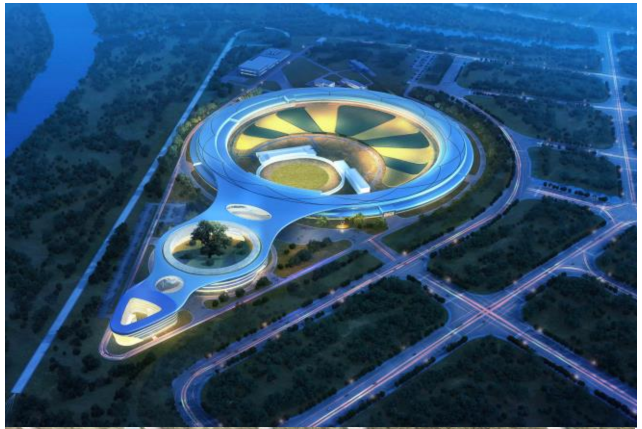
- HEPS introduction
- Scientific data analysis challenges
- System design
- System architecture
- Summary

Outline

- HEPS introduction
- Scientific data analysis challenges
- System design
- System architecture
- Summary

HEPS Introduction

- HEPS, located at Beijing HuaiRou, is an advanced public platform for multidisciplinary innovation research and high-tech development, as well as understanding many scientific questions in the fields of physics, chemistry, biology, etc.
- At HEPS, researchers will be able to observe the complex samples with more sensitive, finer, faster experimental tools, under condition close to the actual working environment. Therefore, researchers will be able to obtain the multidimensional, real-time, in-situ characterization of samples structure, as well as the dynamic evolution processes. HEPS can help researchers understand matter more accurately at the level of molecules, atoms, electrons and spin and in the dimensions of space, time and energy.



Data volume of HEPS

- Under the traditional model, the light source platform only provides data storage services. After users complete experiments on the online site, they download the experimental data to their personal computers for subsequent analysis, rarely using the computing resources of the light source platform.

- In the future, the data scale and throughput generated by advanced light sources will be quite huge, reaching the EB level. This brings unprecedented storage and computing challenges. Traditional data storage and analysis models will seriously hinder the discovery of research results based on HEPS for scientific experiments.

14 beamlines is built in HEPS, first phase 24PB raw experimental data per month

Table 1. Data volume of HEPS beamlines

Beamlines	Burst output (TB/day)	Average output (TB/day)
B1 Engineering Materials Beamline	600	200
B2 Hard X-ray Multi-analytical Nanoprobe (HXMAN) Beamline	500	200
B3 Structural Dynamics Beamline	8	3
B4 Hard X-ray Coherent Scattering Beamline	10	3
B5 Hard X-ray High Energy Resolution Spectroscopy Beamline	10	1
B6 High Pressure Beamline	2	1
B7 Hard X-Ray Imaging Beamline	1000	250
B8 X-ray Absorption Spectroscopy Beamline	80	10
B9 Low-Dimension Structure Probe (LODISP) Beamline	20	5
BA Biological Macromolecule Microfocus Beamline	35	10
BB pink SAXS	400	50
BC High Res. Nanoscale Electronic Structure Spectroscopy Beamline	1	0.2
BD Tender X-ray beamline	10	1
BE Transmission X-ray Microscope Beamline	25	11.2
BF Test beamline	1000	60
Total average:		805

Outline

- HEPS introduction
- Scientific data analysis challenges
- System design
- System architecture
- Summary

Necessity

- With such a huge amount of data, how to reduce the movement of large-capacity experimental data, provide the same data analysis environment as the user's personal computer, and implement the data generation and data analysis processes online, which is a new idea for modern light sources platform to provide services to users.

- It is of great significance to build a HEPS scientific data computing system, provide users with computing resources and analysis environment during and after the experiment, and help users discover one-stop scientific research activities.

Scientific data analysis challenges

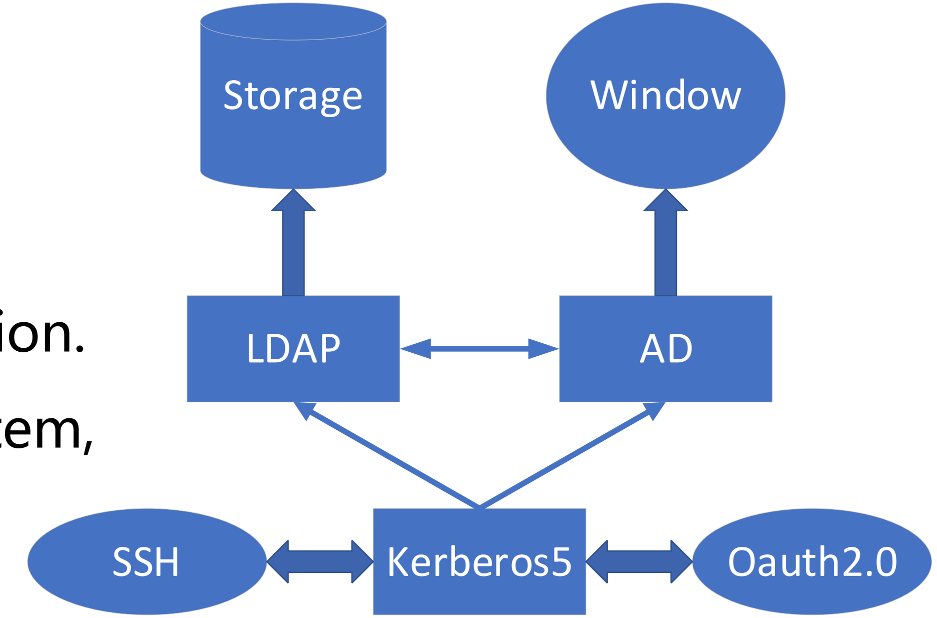
- Compared with the traditional high-energy physics computing platform, the construction of a computing platform for light source users has many challenges:
  - The number of users is huge.
  - User data access is strict.
  - Support multiple analysis scenarios
    - Multiple analysis algorithms.
    - Interactive data analysis scenarios need to be provided.
    - Multiple operating systems support. (Windows/Linux…)
    - Services are provided based on web browser and ssh terminal.
  - Efficient IO access performance.

Outline

- HEPS introduction
- Scientific data analysis challenges
- System design
- System architecture
- Summary

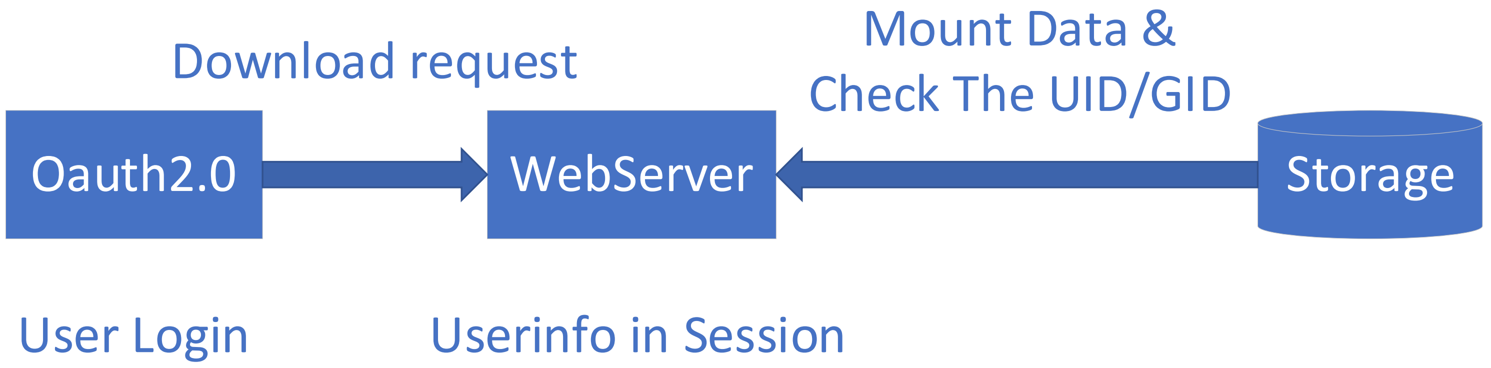
User account management

- HEPS user management is designed based on user access patterns. In the HEPS scientific computing system, users can access the computing environment through three modes:
  - 1. Web page in Oauth2.0 authentication mode.
  - 2. Terminal interface for ssh login.
  - 3. Windows login interface.
- Centralized account management.
  - Use Kerberos5 to store account passwords, and use LDAP and AD domains to store user attribute information.
  - Combined with the HEPS account management system, the consistency of account names in the three storage tools are ensured.



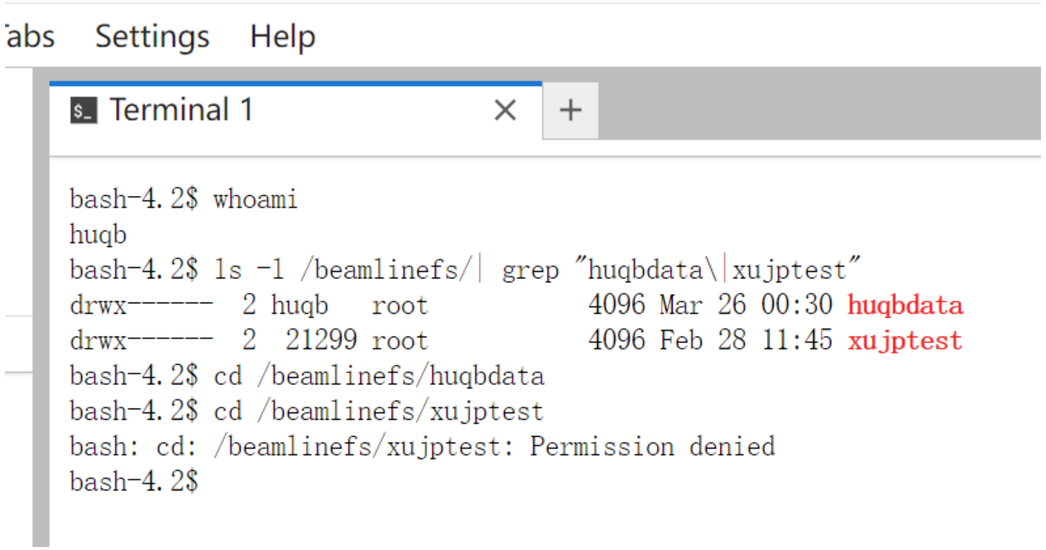
Data access management

- After the HEPS experimental data is generated from the DAQ module, it is written to the POSIX storage system and the ACL corresponding to the user list is set.
- User can access data with the following three modes:
  - Download data from web site.
    - Users can only see the authorized data list on the page.
    - check user info from web session, and check the object file’ s acl contain the current user.

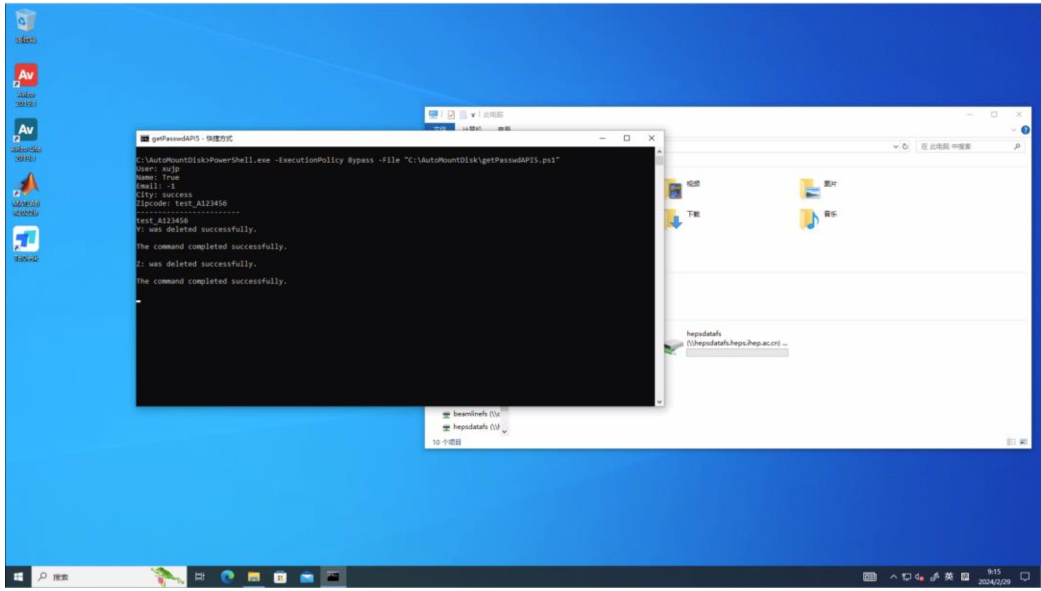


Data access management

- Use the jupyterLab tool to analyze the data.
  - Restrict the user identity of the user's operating system session, and the user directly accesses the data using ACL with POSIX semantics.

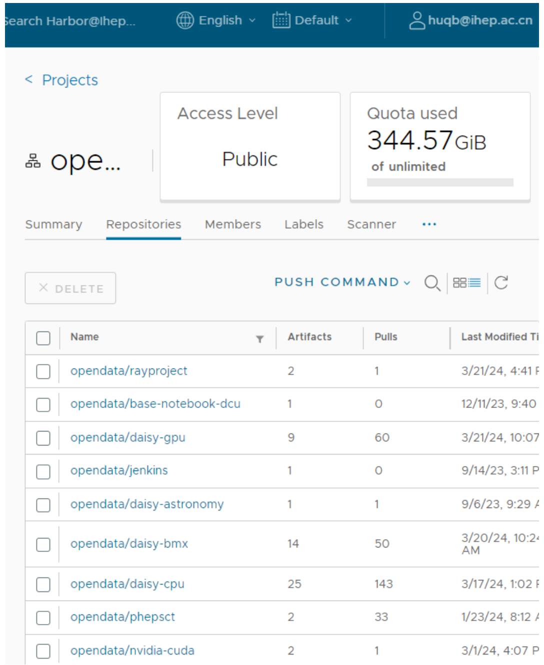


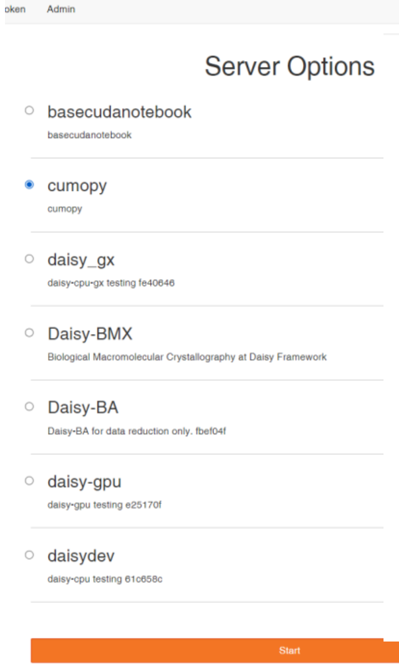
- Use virtual cloud desktops to analyze the data.
  - Map the Microsoft Windows security identifiers (SIDs) to the Linux user/group identifiers (UID/GID). SMB/CIFS

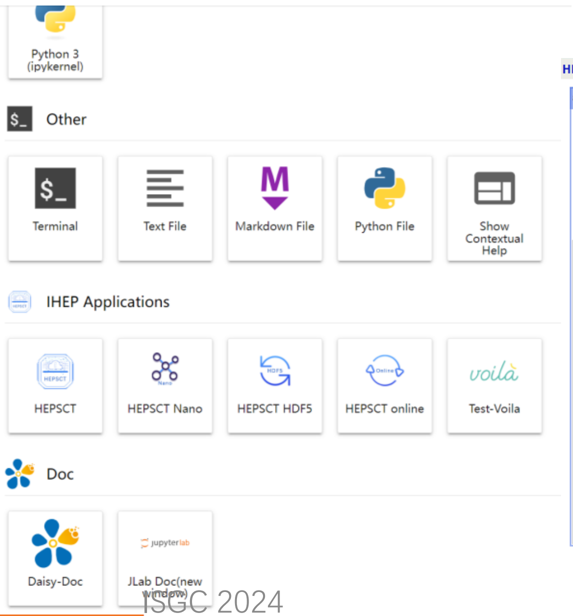


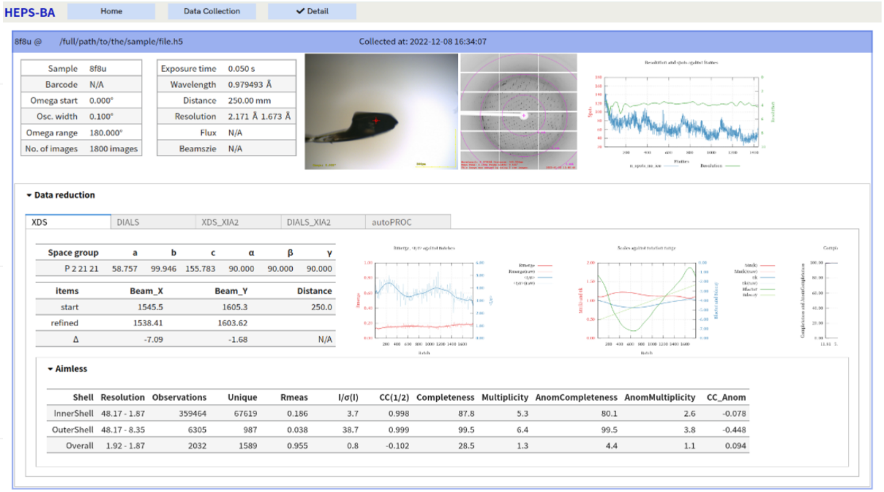
Multi-method analysis scenarios

- Container virtualization solution
  - Build multiple methodologies into jupyterLab base container images.
  - Uses jupyterhub to provide users with a selection list of different scenarios, and manages the corresponding container images.
  - The user selects the corresponding methodology to start the analysis scenario.



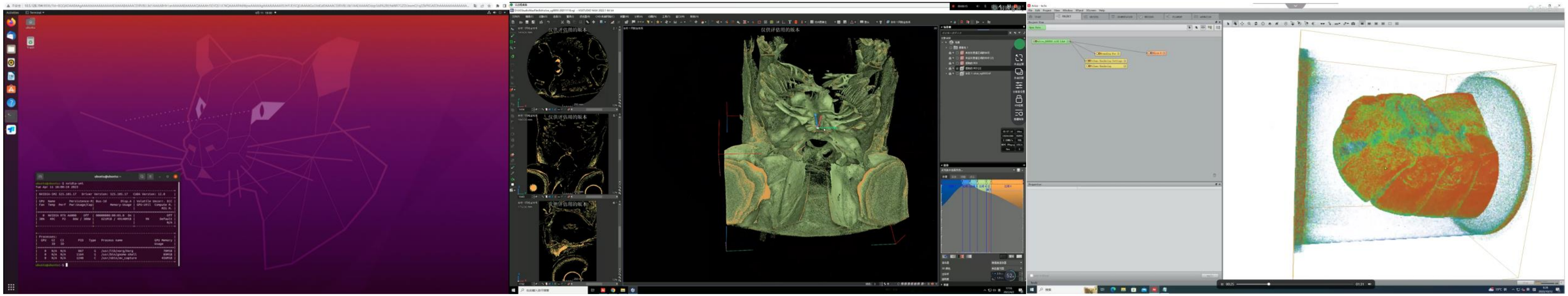






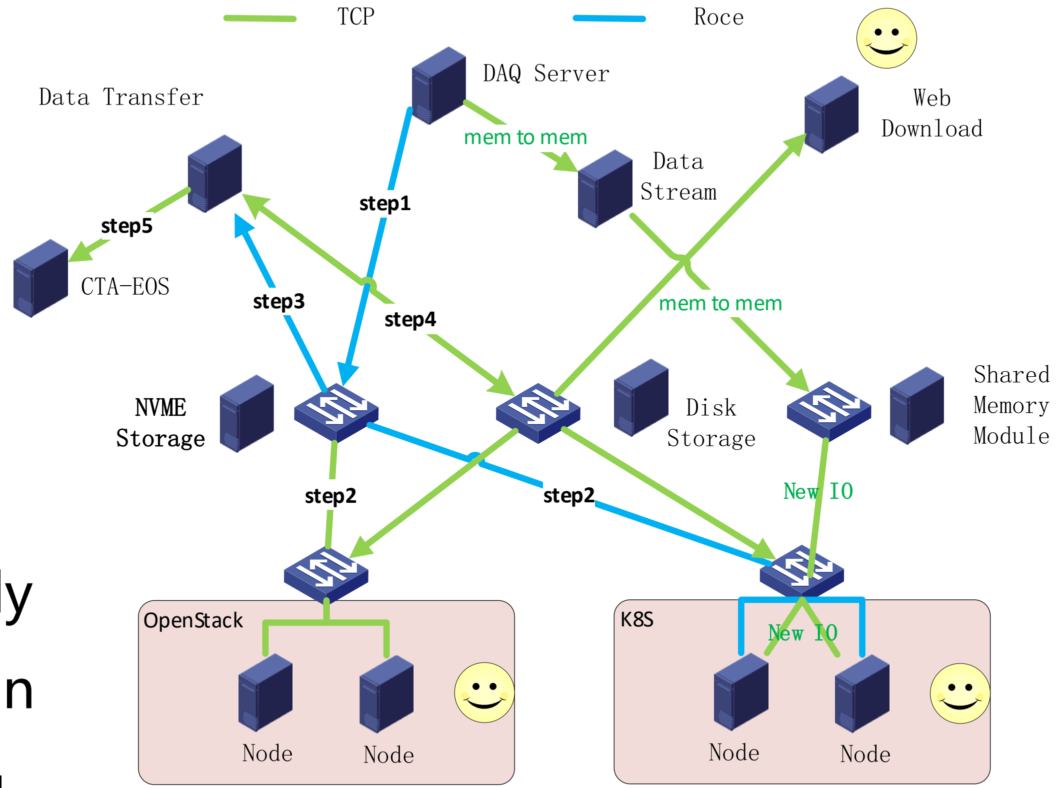
Multi-method analysis scenarios

- Virtual machine solution
  - Preinstall the software on the operating system and build VM image.
  - GPU hardware is directly connected to the virtual machine operating system.
  - Provide data analyze service to user through the virtual cloud desktop.
  - A streaming server is deployed and the image compression algorithm is optimized to transmit desktop images in low-bandwidth mode.



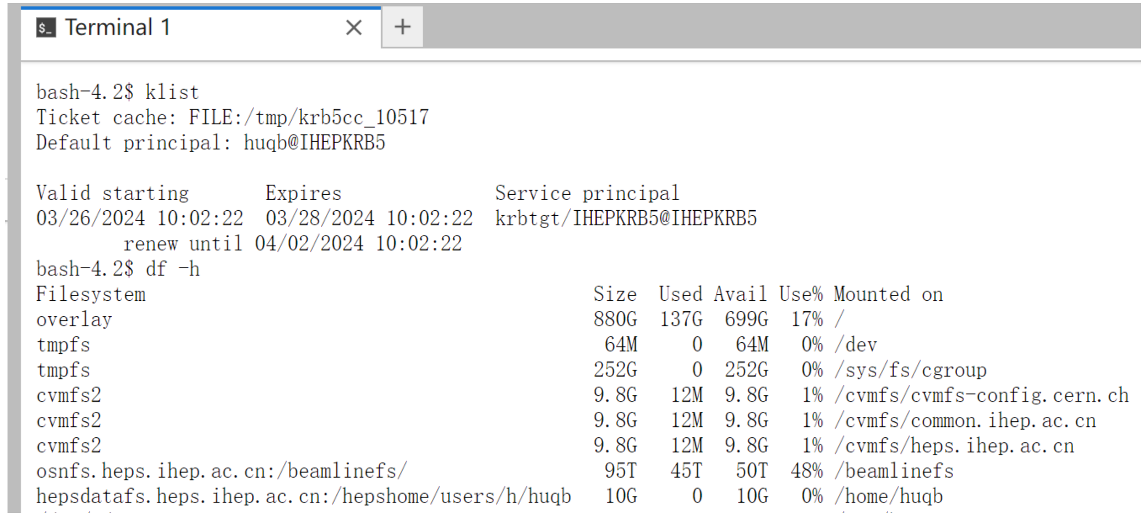
High-performance IO

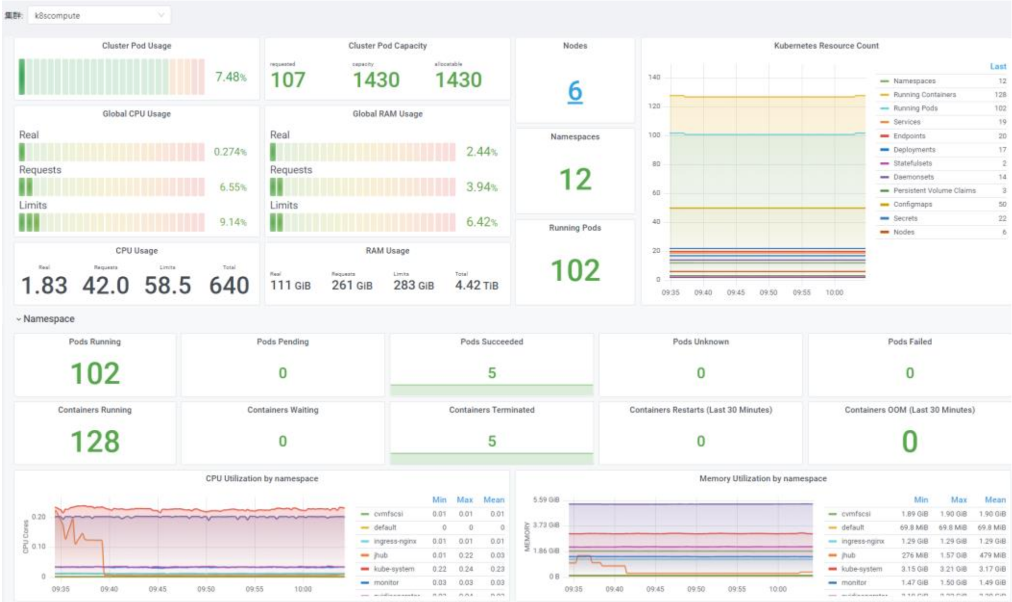
- In order to provide high-performance IO,
  - NVME disks to store data for one week.
  - STAT disks to store data for one month.
  - The Roce network connection is used between the client and the NVME storage server. The client uses RDMA technology to access stored data and provide optimal IO performance.
- After the data is transmitted from DAQ, it is directly stored in the memory message queue. The solution of using data flow for data processing is still being explored.



Other basic services support

- Cvmfs is used to deploy platform-dependent software.
  - CVMFS CSI plugin provides read-only mounting of CernVM-FS (CVMFS) repositories in HEPS analysis container.
- Krb5 TGT is used to identify users.
  - User TGT writes into the environment variables when starting the container environment.
- Prometheus tool is used to monitor cluster status.





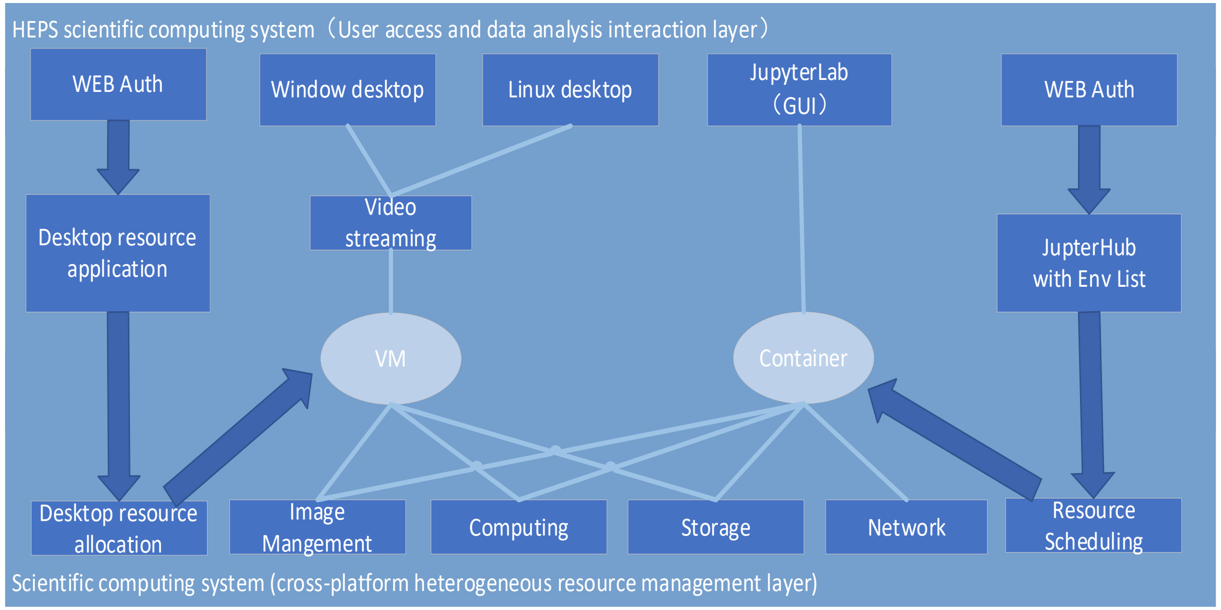
Outline

- HEPS introduction
- Scientific data analysis challenges
- System design
- System architecture
- Summary

Platform architecture

- There are two layers in the platform
  - User access and data analysis interaction layer
  - Cross-platform heterogeneous resource management layer

- Platform service process.
  - Users pass web authentication.
  - Select interactive data analysis scenarios.
  - Apply resources and enter the analysis interface to perform data analysis



Summary

- In response to the many challenges of the light source computing platform, container virtualization and virtual machine technology are used to achieve heterogeneous resource integration and resource elastic expansion. The data analysis environment can be flexibly deployed according to the diverse needs of users.
- The centralized account management model combined with the storage system enables data security under multiple data access modes.
- Based on virtual cloud desktop technology and jupyterLab scientific analysis tools, it provides users with a flexible interactive analysis method.
- Deployment of high-performance network and high-performance storage to provide high-performance IO for data analysis.
- In order to further improve IO performance, a stream-based data access model is under development.

Thanks!