IHEP Site Report

On Behalf of IHEP-CC

Jingyan Shi

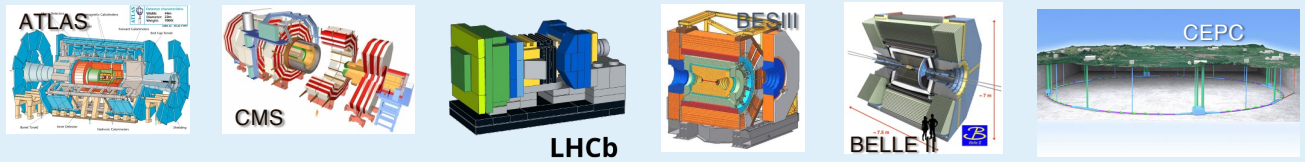
shijy@ihep.ac.cn

Outline

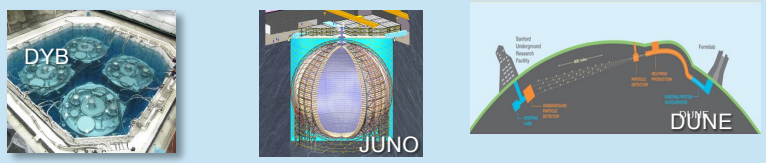
Brief Introduction to IHEP

- Institute of High Energy Physics
- The largest fundamental research center in China
  - Experimental Particle Physics
  - Theoretical Particle Physics
  - Astrophysics and cosmic-rays
  - Accelerator Technology and applications
  - Synchrotron radiation and applications
  - Nuclear analysis technique
  - Computing and Network application

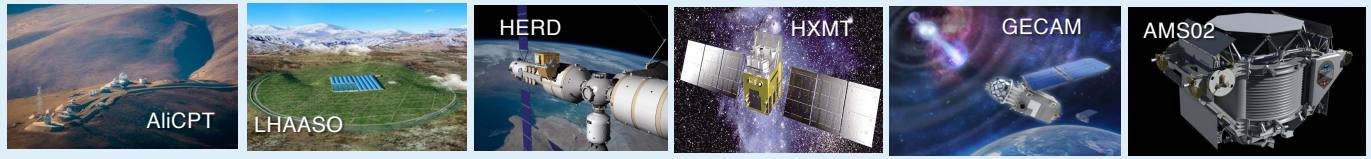
HEP Related Projects



Accelerator based particle physics



Neutrino physics



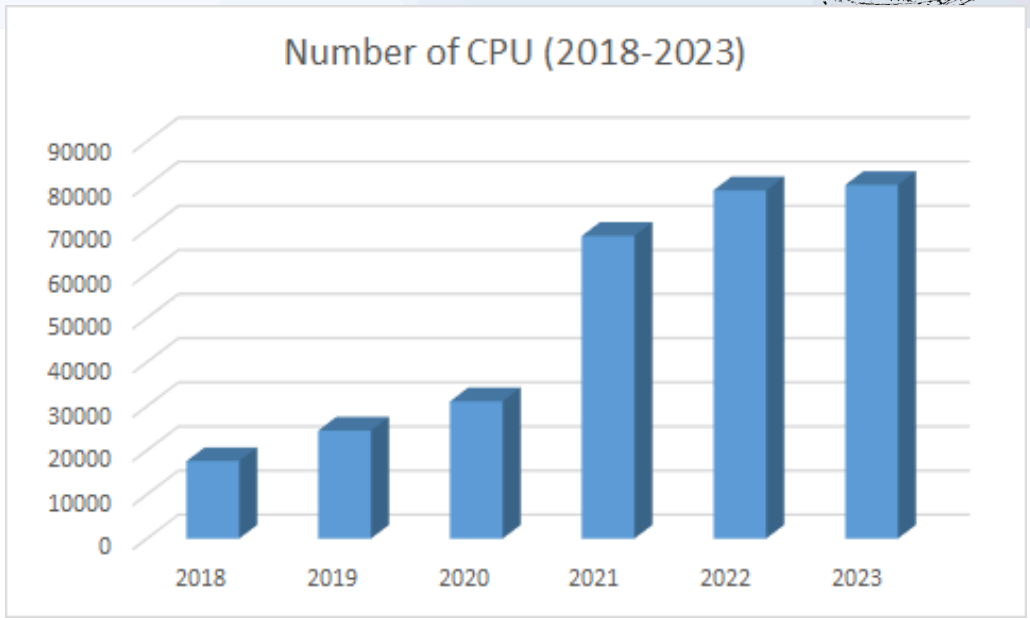
Cosmic ray and astrophysics experiments

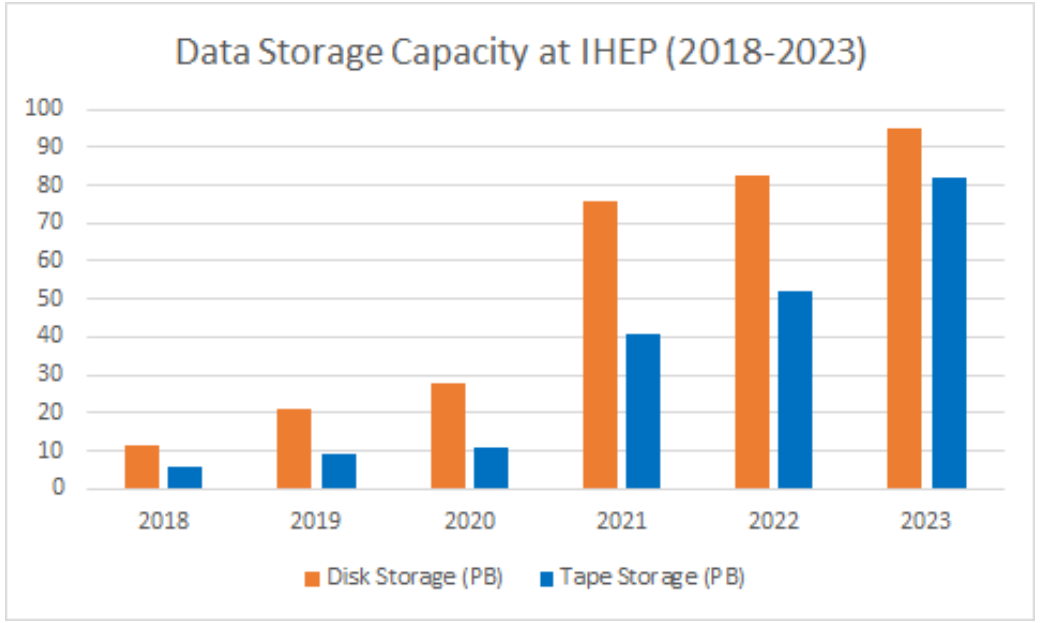
Neutron Source and Synchrotron Radiation Facilities

Facilities of IHEP

Computing and Data Storage

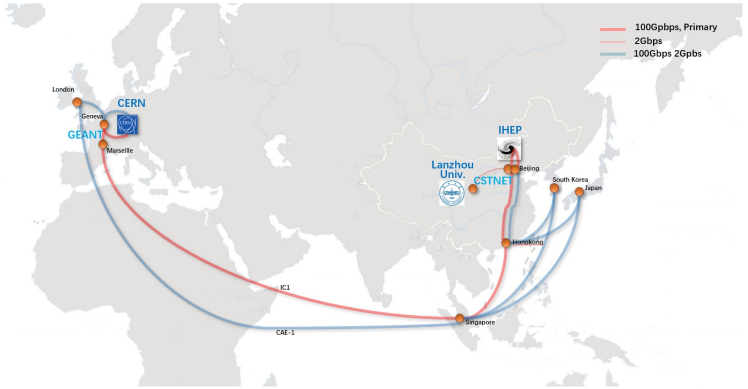
- Distributed Locations
  - Data center: Beijing, Dongguan
  - Exp. Onsite: Daocheng, Jiangmen, Tibet
- Quantity of resources grew exponentially
  - ~90k CPU cores
  - ~100PB disk storage
  - ~80PB Tape storage
- HTC and HPC for experiments
  - 28 experiments / applications

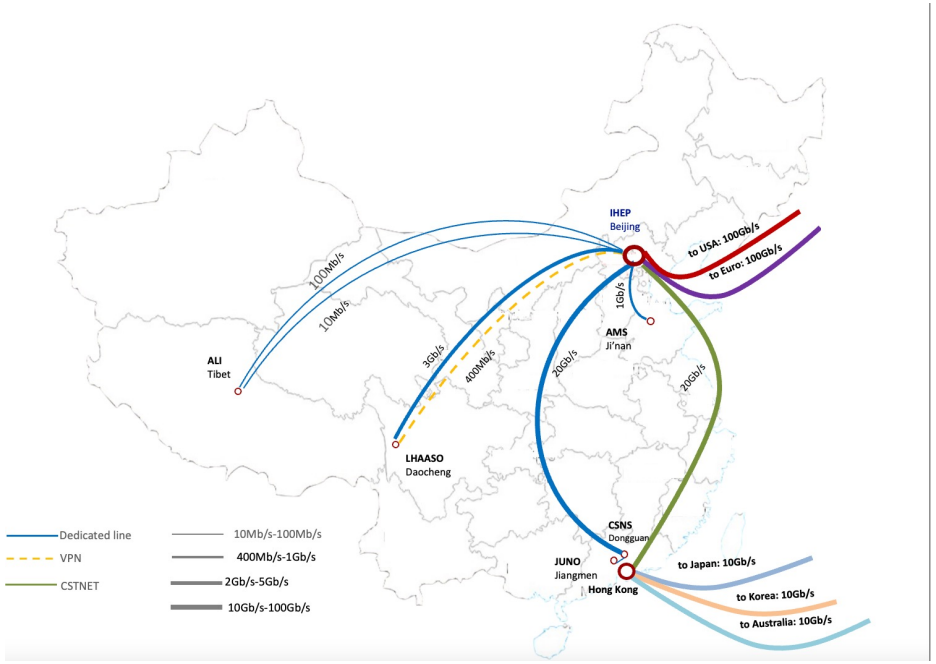




International Network

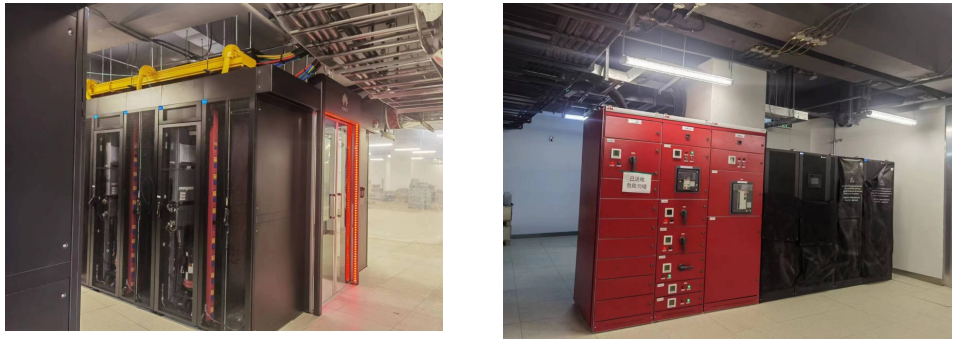
- International network link upgraded to 100Gbps in 2023
  - With the help from CSTNET, GEANT and CERN
  - The data transfer test showed the peak performance between IHEP and Europe reach to 50 Gbps
- Dedicated links between IHEP and domestic remote sites
  - HEPS-IHEP: 100 Gbps
  - CSNS-IHEP: 20 Gbps
  - LHAASO-IHEP: 2 Gbps(3 Gbps since Mar. 2024)
  - JUNO
  - Lanzhou Univ-IHEP: 2 Gbps
  - .....





IT Services for HEPS

- High Energy Photo Source（HEPS)
  - First high-energy synchrotron radiation light source in China
  - Located 100km north of IHEP and will be commissional in 2025
- 520m^2 New machine room for HEPS is ready
  - 2500kVA+2500kVA mains power supply and 800kVA UPS
  - 47 racks in Phase I: 20 for storage, 21 for computing and 6 for network
  - Storage and computing servers have been installed
- Software of data management and analysis have been deployed
  - Integration and performance tuning is undergoing



HEPS machine room

HEPS data service portal

Service Provision of CC-IHEP

- Computing Statistics last 6 months
  - HTCondor Cluster: 34k CPU cores
    - Job slot utilization: 87.74%
  - Slurm Cluster: 8864 CPU cores and 254 GPU Cards
    - Utilization of CPU: 79.23%
    - Utilization of GPU: 73%
  - Grid Computing

Exp.	Atlas	CMS	LHCb	BELLEII	JUNO
CPU hours	2,929,587	699,155.33	9,265,955.19	72,270.0	95,484
Efficiency	94.01%	58.18%	79.64%	96.71%	89.35%

Service Provision of CC-IHEP

- **Lustre: 31.01 PB with 65.81% usage ratio**

Exp.	BESIII	LHAASO	JUNO/DYW	CEPC	ASTRONOMY	Other
Capacity/Usage	14.3 PB / 9.72 PB	610 TB / 480 TB	3.1 PB / 2.4 PB	3.7 PB / 2.5 PB	6.8 PB / 3.75 PB	2.5 PB / 1.56 PB
Usage Ratio	67.9%	78.6%	77.4%	67.5%	55.1%	62.4%

- **EOS: 60.58 PB with 82.08% usage ratio**

Exp.	LHAASO	JUNO/DYW	ASTRONOMY	Other
Capacity/Usage	51.79 PB / 43 PB	7.96 PB / 5.76 PB	806 TB / 750 TB	240 TB / 220 TB
Usage Ratio	83%	72.3%	93%	91.7%

- **Grid: 9.64 PB with 27.3% usage ratio**

Exp.	Atlas	CMS	LHCb	BELLEII	JUNO
Capacity/Usage	398.43 TB / 317.68 TB	673.51 TB / 574.32 TB	6.69 PB / 0.25 PB	279.88 TB / 49.28 TB	1.6 PB / 1.44 PB
Usage Ratio	79.7%	85.27%	3.7%	17.6%	90%

Linux Operating System Upgrade

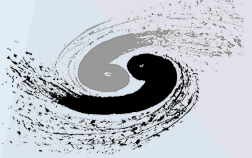
- CentOS 7 will be officially unsupported after 30th June 2024
- Evaluation on Alma Linux 9.3
  - System software deployment – almost done
  - Physics software evaluation – undergoing
- Upgrade plan of IHEP computing platform
  - Alma Linux 9 will be in production from Aug, 2024
    - Cent OS 7 will be supported in container

Alma Linux 9.3 Evaluation

System Software	Status
HTCondor	√
Slurm	√
Lustre	√
EOS	√
NFS	√
CVMFS	√
AFS	√
Kerberos	Undergoing
Container	√

Timeline for Alma Linux 9 at IHEP

Linux Operating System Upgrade

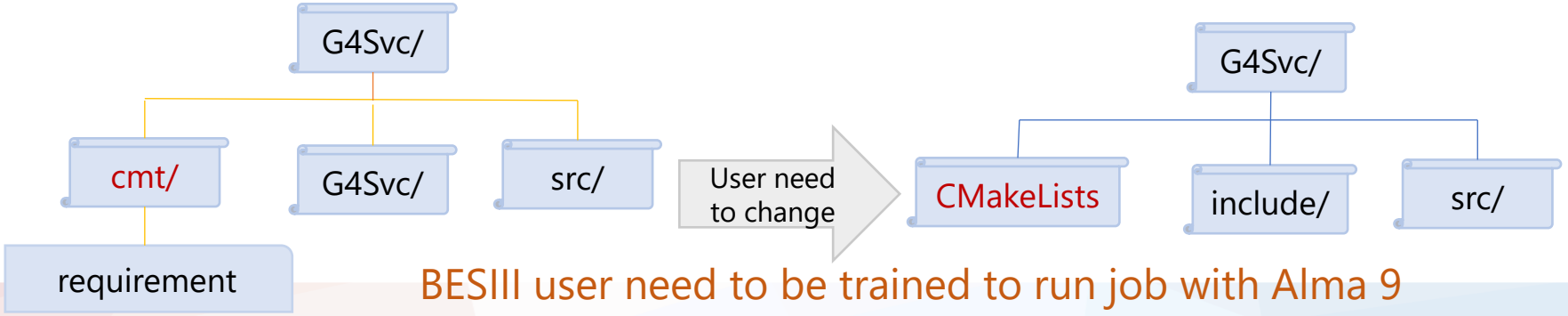


- Basic software compilation
  - Compilers, basic libraries, program languages and HPC software are compiled
    - OpenMPI, MPICH, Python, gcc, …
    - fftw, lapack, scalapack, gromacs, cp2k, …
    - CUDA, cuDNN …
  - Modulefiles are created to load software environment
  - New license need to be bought/renewed for some commercial software
    - MATLAB, Mathematica, Comsol, gdfidl
  - Tips
    - AlmaLinux 9 has nouveau driver enabled by default
      - NVIDIA GPU cards could not be identified as a result
    - Some kernel modules are not installed by default
      - e.g. knem

Exp. Software for Operating System Upgrade

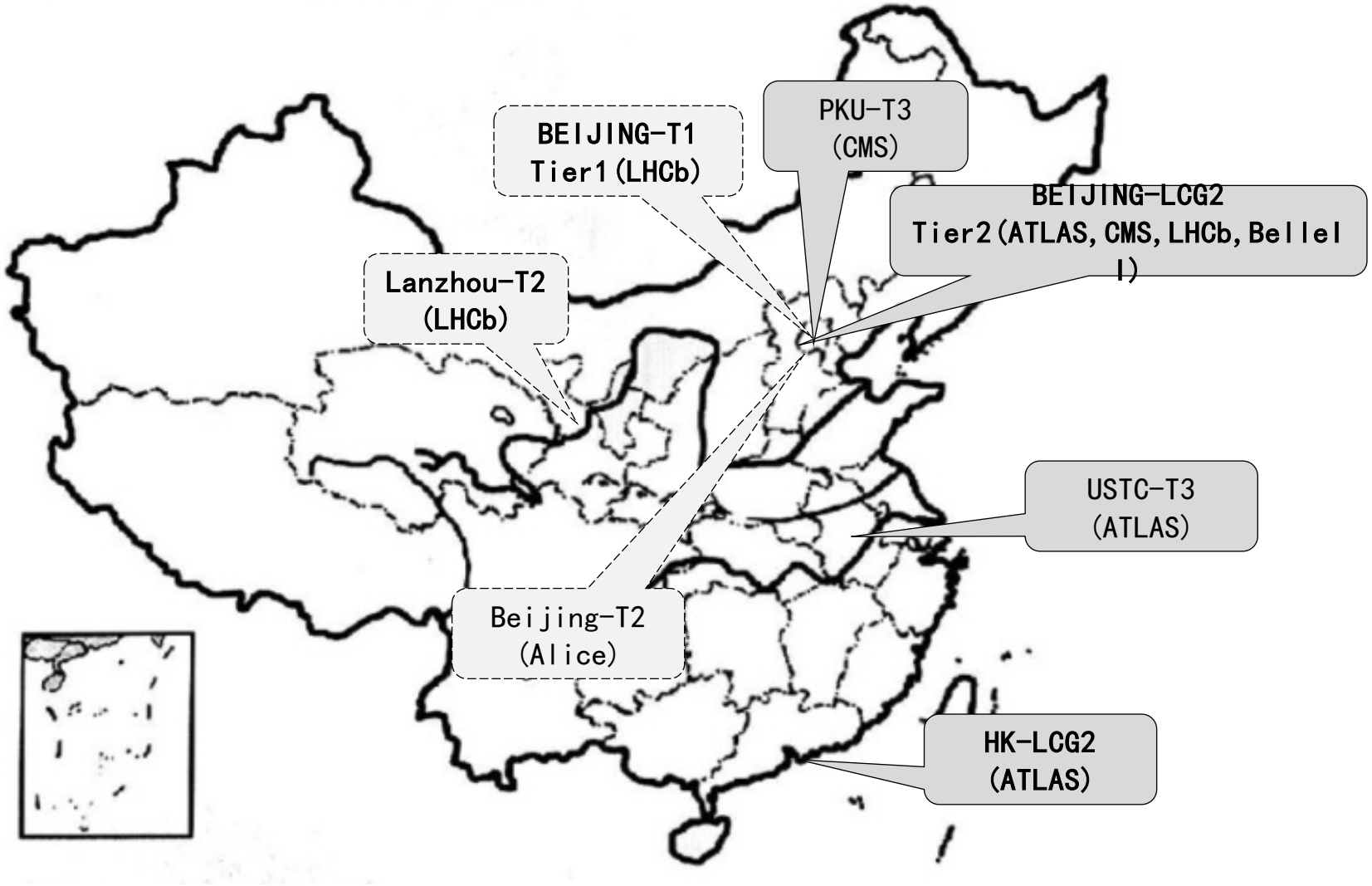
- Most software of the Exp. led by IHEP compiled and be ready
  - LHAASO, JUNO, etc.
- Compilation of BESIII software BOSS is undergoing
  - BOSS has been the BESIII software since 2009
    - Cernlib has been used by BOSS till now
      - not available in LCG
      - Necessary by BESIII generators (KKMC, BesEvtGen)
    - need to be compiled within Alma 9
  - Important changes to upgrade to Alma Linux 9
    - CMT → Cmake
      - More simple structure for BOSS but a lot of work need to be done for the complitation
    - A big version upgraded for GCC and Gaudi → Big Changes
  - Current progress
    - LCG, Gaudi, Geant, BersGDMS were compiled and be ready
    - Run HelloWorld successfully in BOSS

	BOSS 7.1.1	Upgraded Version
OS	CentOS 7	Alma Linux 9.3
GCC	4.9.3	GCC 13
CMake	cmt	3.26.2
Python	2.7.10	3.9.12
Gaudi	v27r1	v36r14
LCG	LCG_84	LCG_104
ROOT	6.20.02	6.28.04
Geant4	10.7	10.7



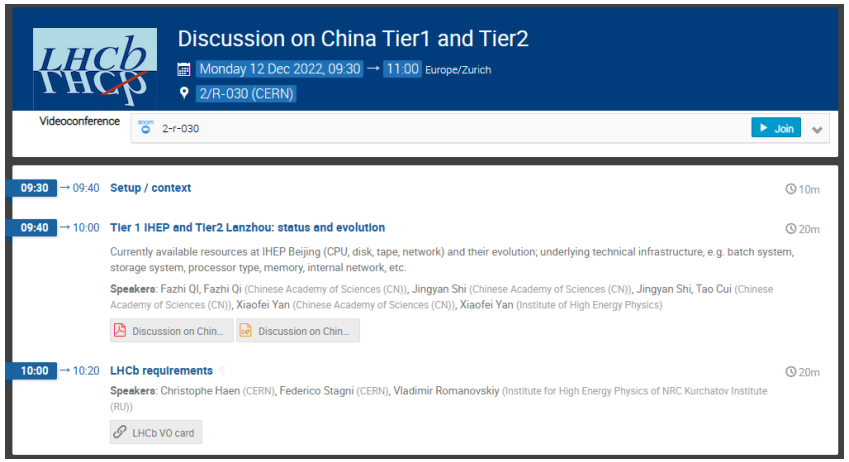
BESIII user need to be trained to run job with Alma 9

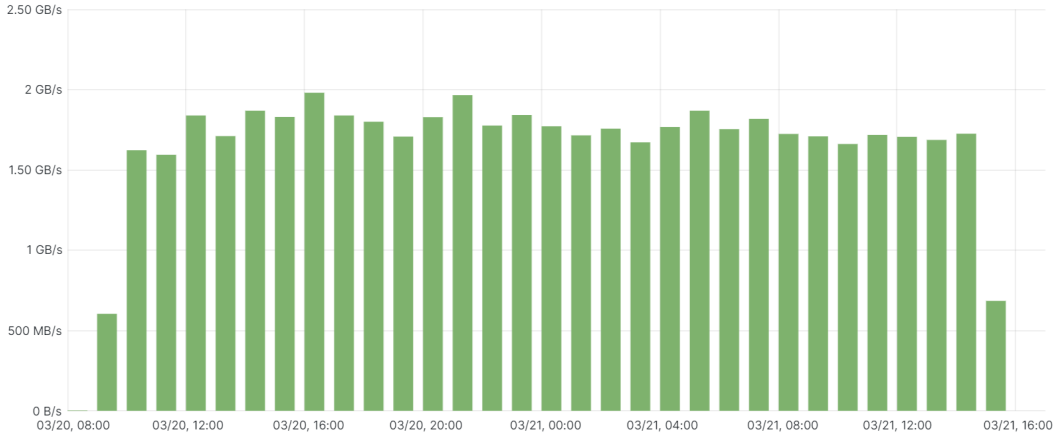
Overview of WLCG Sites in China



Construction and Resource of LHCb Beijing Tier-1 Site

- Construction
  - Oct. 2023: Chinese LHCb collaboration and CC-IHEP decided to construct Tier-1 Site for LHCb
  - Dec. 2023: Discussed and received the approval from WLCG
  - Feb. 2024: Construction completed
- Resource provided for LHCb Beijing Tier-1
  - Computing:
    - 40 worker nodes (Intel & AMD) with 3216 CPU cores: 67,000 HepScore
  - Disk storage
    - 4 sets of storage arrays provide 3.2 PB
  - Tape storage:
    - 4 drivers (IBM) and 170 tapes with 3 PB
  - Network equipment and management server:
    - 6 switches, 1 router, 2 band cards and 10 servers
- First data challenge has been done in Mar. 2024
  - 189 TB data was transferred into IHEP Site in ~2 days
  - Average transfer speed is about 1.55GB/s (Max is 1.98)
  - Transfer efficiency was close to 100%
- Will be in production in Jun. 2024

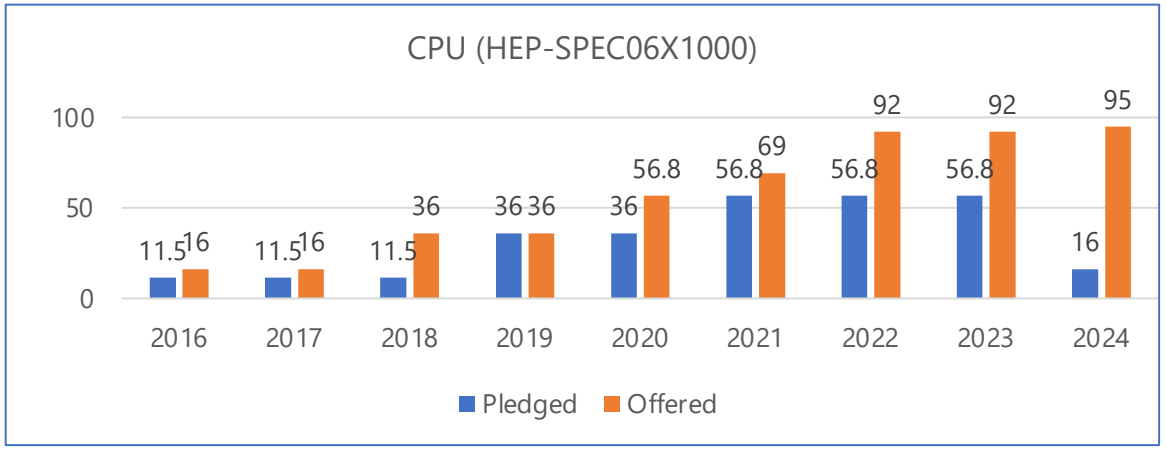


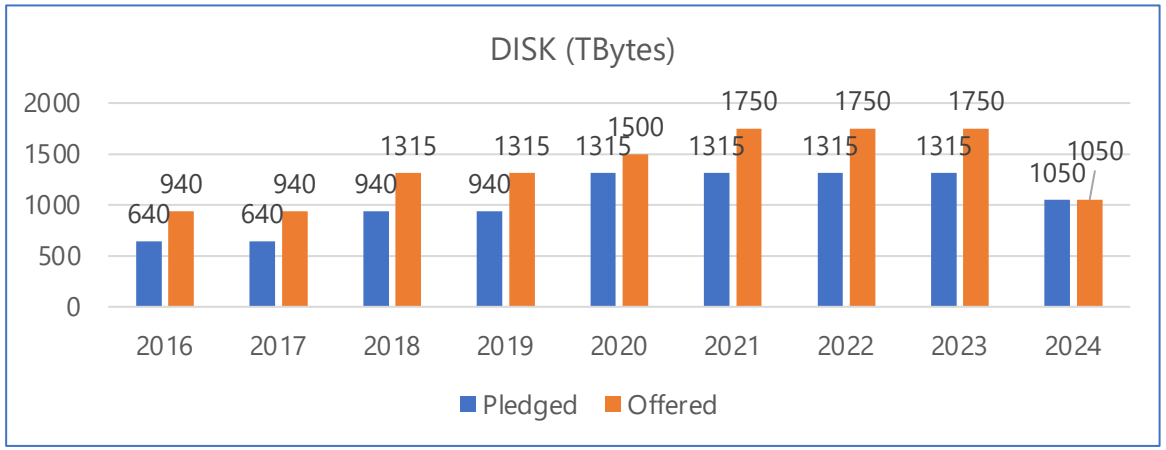


Network traffic of the first data challenge

Chinese Tier-2 Site Federation

- CPU: 4472 cores with 95,000 HepScore
  - AMD 9654: 1152 Cores
  - Intel Golden 6338: 1280 Cores
  - Intel Golden 6140: 1152 Cores
  - Intel E5-2680V3: 696 Cores
  - Intel X5650: 192 Cores
- CE & Batch: HTCondorCE & HTCondor
- VO: ATLAS, CMS, LHCb, BELLEII, JUNO, CEPC
- Storage: 1050TB
  - 4TB * 24 slots with Raid 6, 5 Array boxes
  - DELL MD3860 8TB * 60 slots
  - DELL ME4084 10TB * 42 slots
  - DELL ME4084 12TB * 84 slots
- Got new budget for upgrading ATLAS and CMS Tier-2 in 2024
  - CPU：60,000 HepScore
  - Disk storage：2.5PB

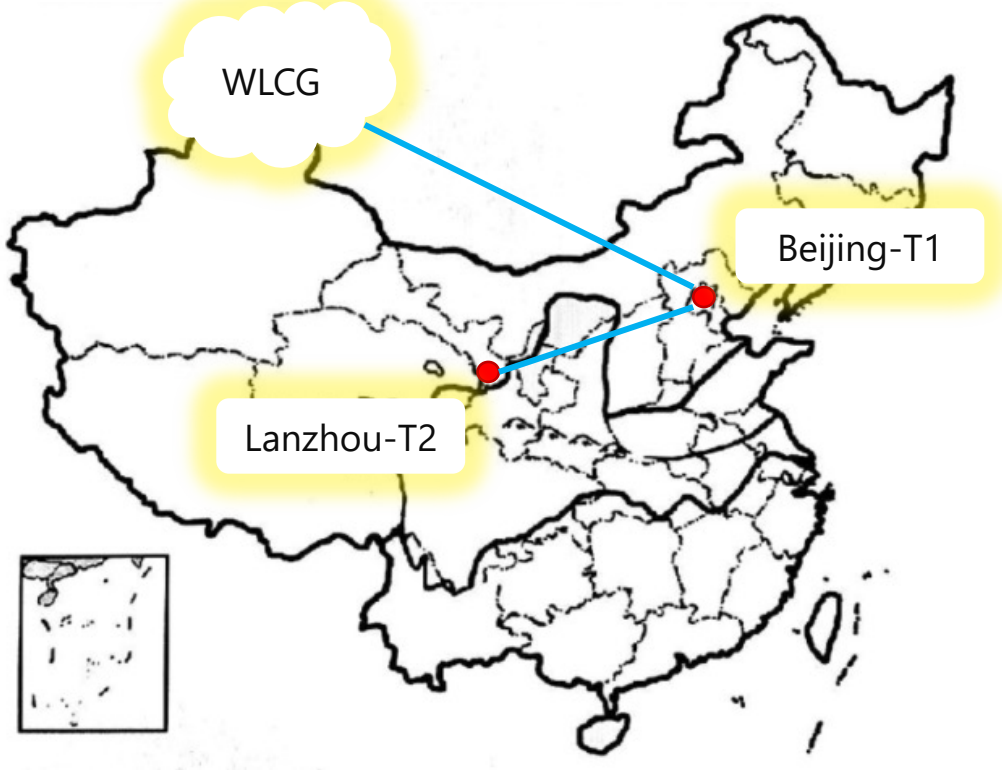




Computing and Storage Pledge of BEIJING LCG Tier- 2

Construction of New WLCG Tier-2 Sites

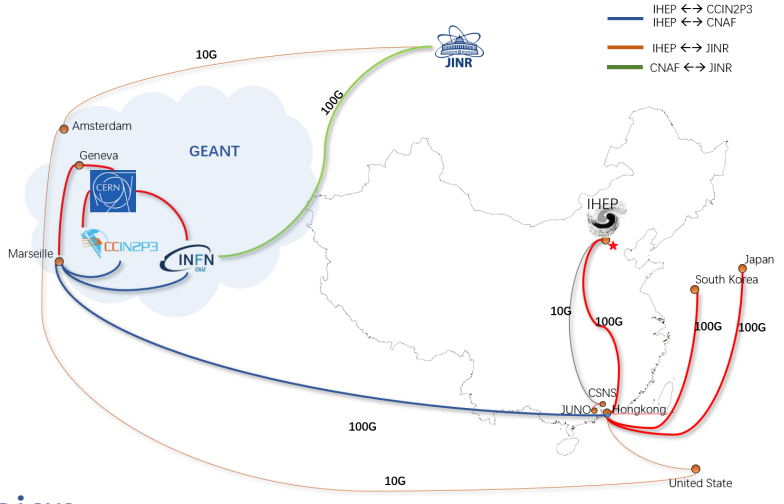
- LHCb Tier-2 site in Lanzhou
  - Construction started in Oct. 2023
    - ~3500 CPU cores with 77,000 HepScore
    - ~3 PB Disk Storage
    - Dedicated 2 Gbps link between IHEP and Lanzhou Univ.
  - Software deployment started in Apr. 2024
  - Jointly maintained by CC-IHEP and Lanzhou Univ.
    - Hardware maintenance: Lanzhou Univ.
    - Software deployment and maintenance: CC-IHEP
- Alice Tier-2 at IHEP
  - Chinse Alice collaboration would like to build Tier-2
    - 1152 CPU cores with 30,600 HepScore
    - 840TB disk storage
    - CC-IHEP to be responsible for the overall maintenance
  - Expected Production: Aim for in production in 2024



Lanzhou Univ. LHCb Tier-2 Site

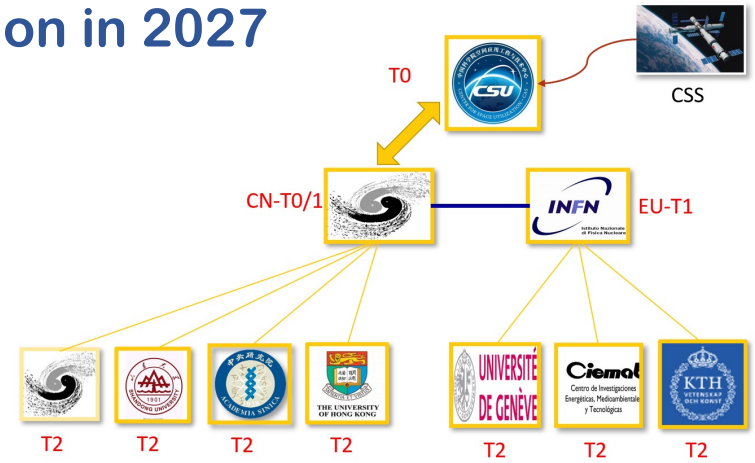
Grid Computing of JUNO & HERD at IHEP

- JUNO: Jiangmen Underground Neutrino Observatory, will start production in 2024
  - IHEP site runs as Tier-0 and Tier-1
    - Storage: 8 PB by EOS, Lustre on disk, 4 PB by EOS-CTA on tape.
    - Computing: 180 KHS06 by HTCondor on x86, Slurm on ARM and GPU.
  - Network: 10Gbps(From JUNO-onsite to IHEP), 100Gbs(From IHEP to GEANT).
  - Grid computing platform: DIRAC system with IHEP-extensions.



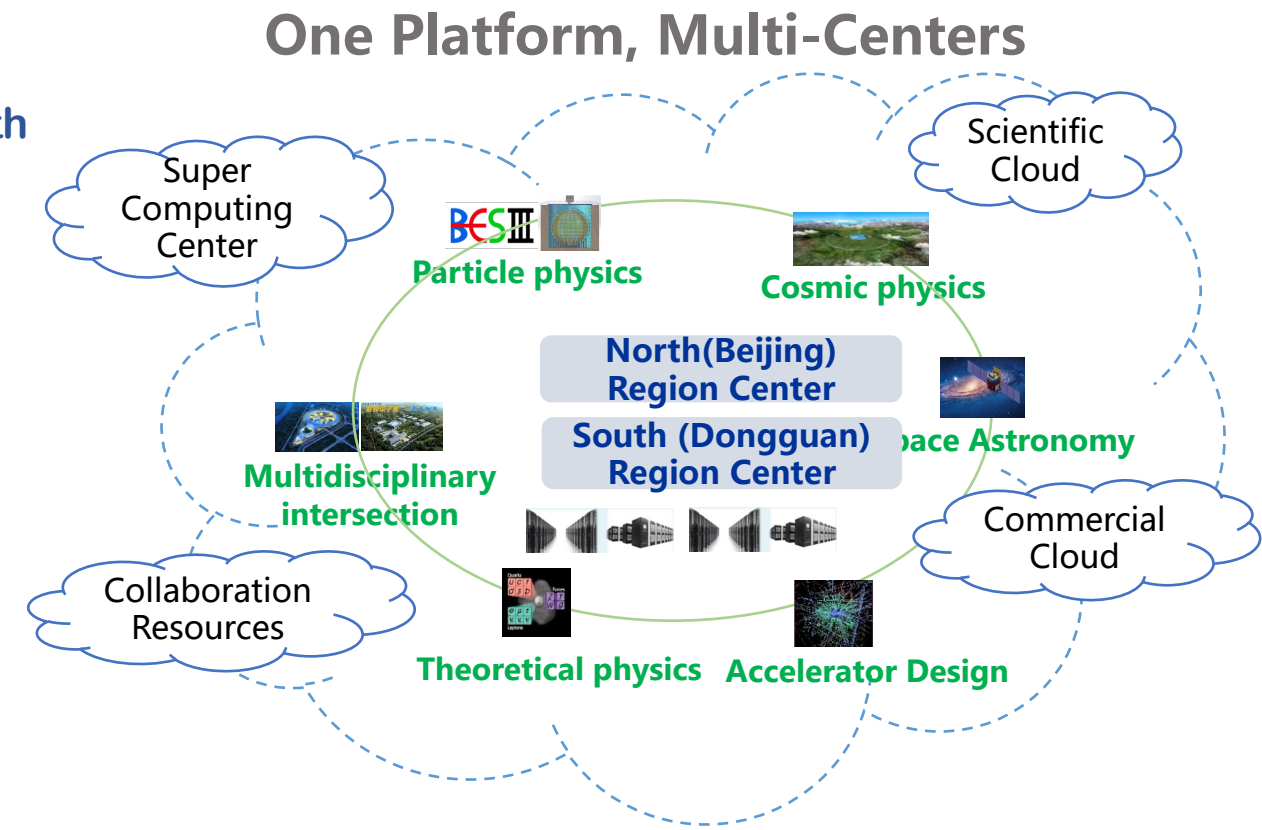
- JUNO Data Challenge 1:
  - 12th ~ 26th Feb 2024, corresponding to WLCG DC24.
  - Pressure transfer (500-1000 Mbps) with 4-8 times throughput than JUNO design.
    - IHEP -> CNAF/IN2P3 transfer worked well, almost no failure.
      IHEP->JINR is bad and always get stuck.

- HERD: High Energy Radiation Detection Facility, will start production in 2027
  - Storage requirements: 70 PB in 10 year.
  - Computing requirement: 16000 CPUs in 10 years.
  - Two Tier-1 sites will run at China and Europe
  - Grid computing platform
    - DIRAC + dHTC (HTC&HPC) for computing, Rucio for data management



One-Platform, Multi-Centers

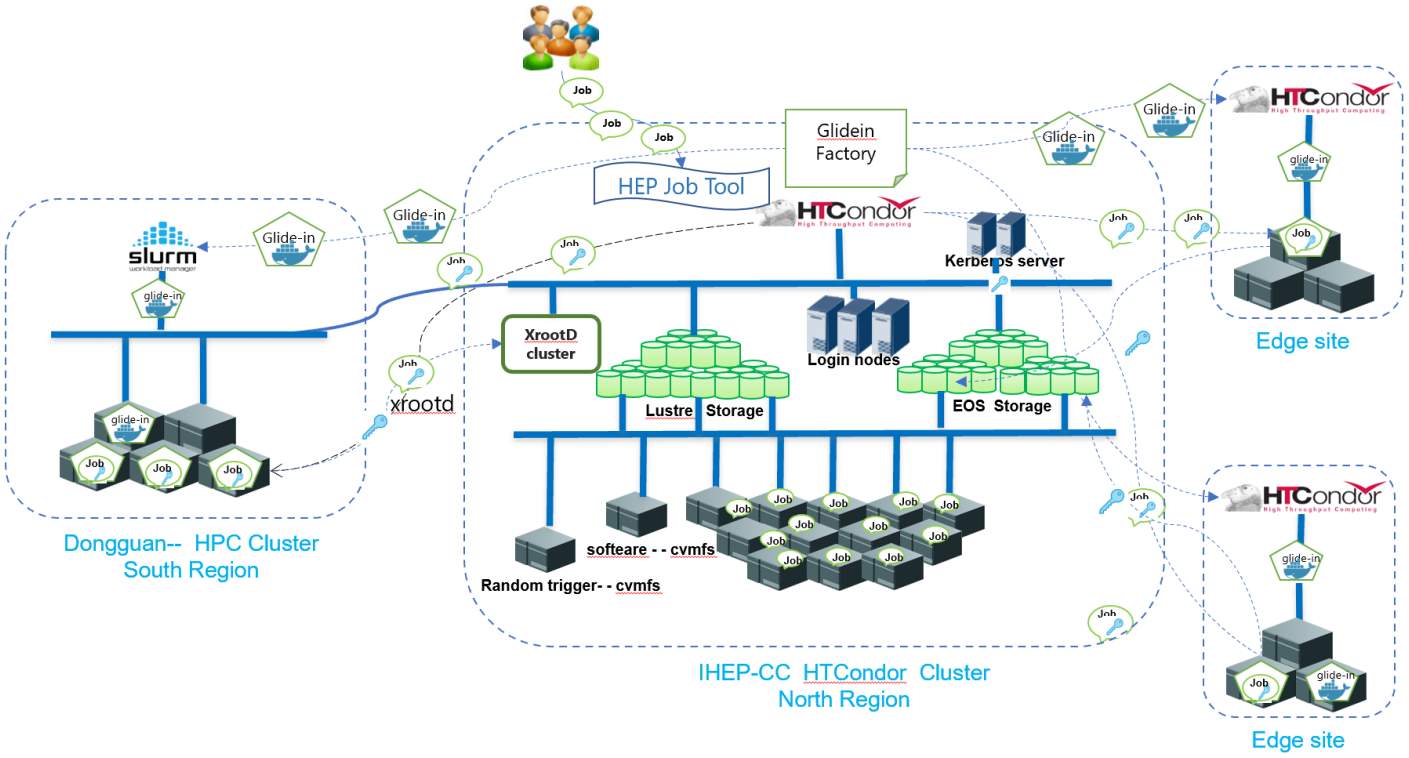
- Construct a cross-regional unified data processing platform for HEP in China
  - North and south region centers
  - Combining multiple remote sites and Compatible with heterogeneous hardware
  - IT services deployed to the HEP Exp. Facilities
  - Collaboration member IT resources
  - Commercial clouds
  - Super computing Center
- Two models
  - Grid model
  - Local cluster expanding
    - Expand HTC cluster to the remote resources



HTC Local Cluster Expanding

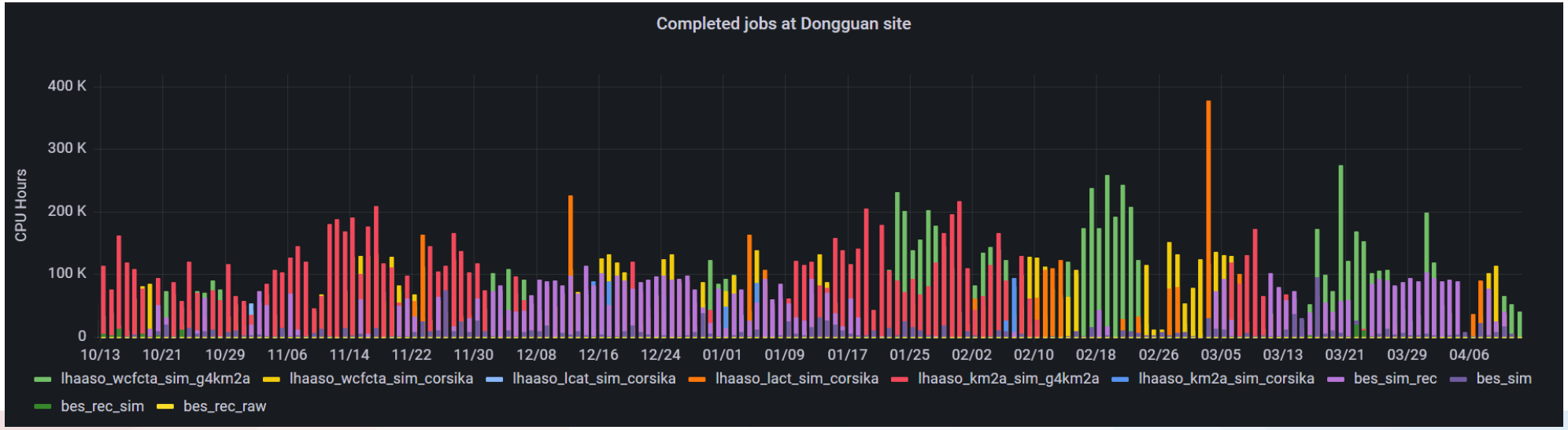
- Local cluster is the main way of data processing to the HEP Exp. Led by IHEP
  - Job slots remain highly utilized and with significant job queuing time
  - Users are accustomed to the “cluster way” than the “grid way”
    - Easier and more convenient
    - No data management tool and only local file system storage for job running

- Expand IHEP local HTCondor cluster to the remote resources
  - Classified the job and site
    - Simulation and reconstruction job are dispatched to the remote worker node to suit the limit network bandwidth
  - User authentication from remote ends based on Kerberos
  - Necessary files transfer to/from remote via xrootd



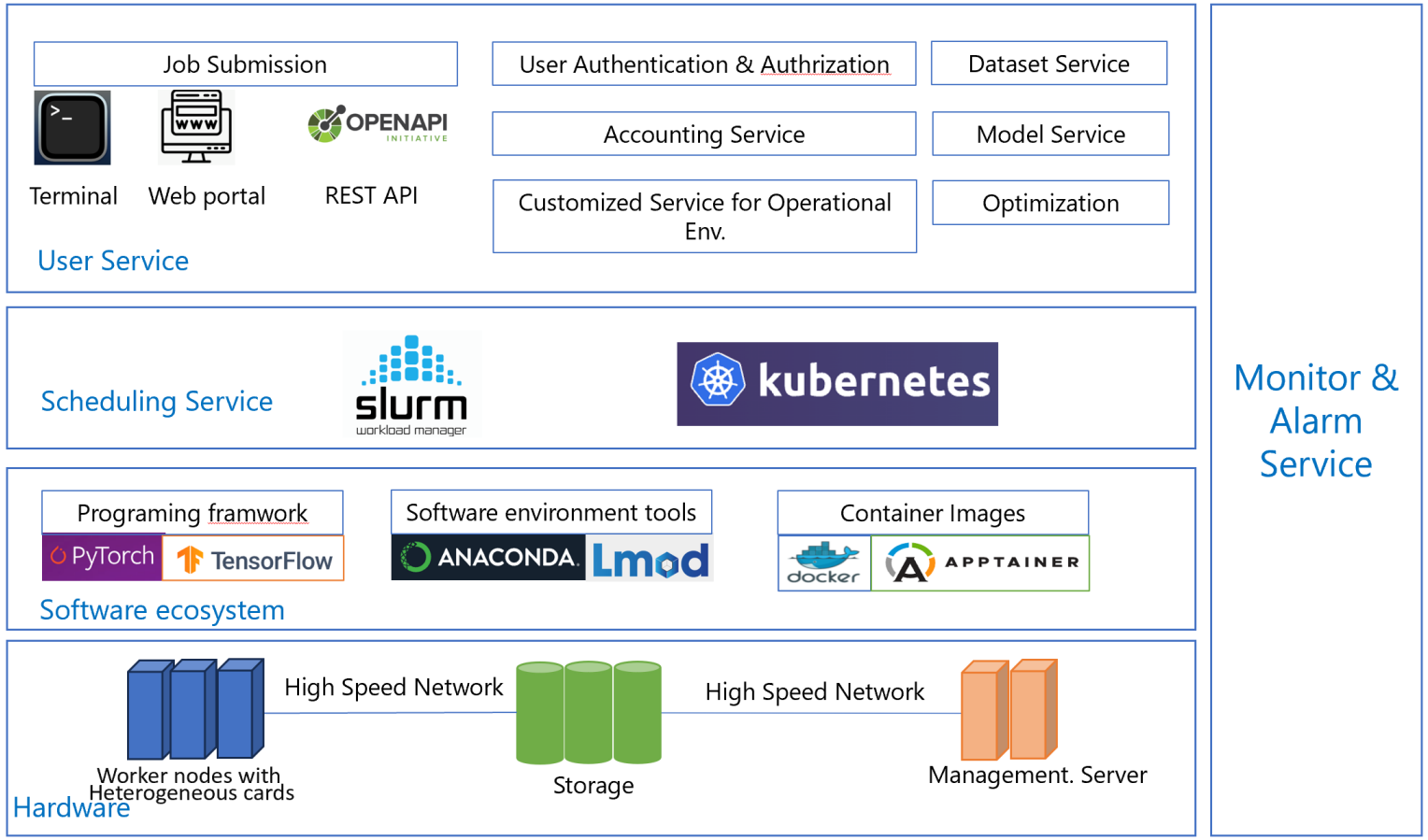
HTC Cluster Jobs dispatched to Dongguan

- Dongguan slurm cluster contribute part of the CPU cores to IHEP HTC jobs
  - 8,000 cpu core + 10,000 arm cores
- Simulation and reconstruction jobs of LHAASO and BESIII have been dispatched to Dongguan based on HTC local cluster extension
- Statistic of the jobs running at Dongguan last 6 months
  - LHAASO: 30,701,524 cpu hours, 44.5% of the whole LHAASO jobs
  - BESIII: 7,582,166 cpu hour, 9.2% of the whole BESIII jobs



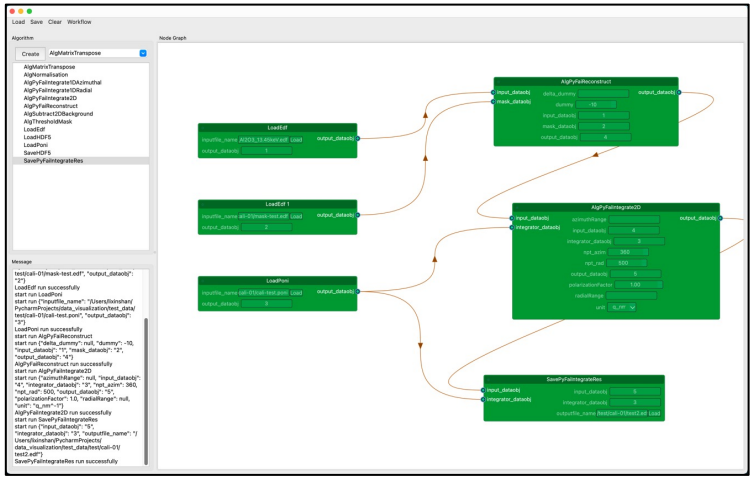
Plan for AI Platform at IHEP

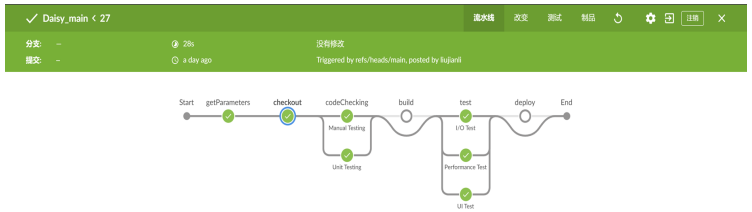
- Got new budget for AI Platform
- Research is undergoing

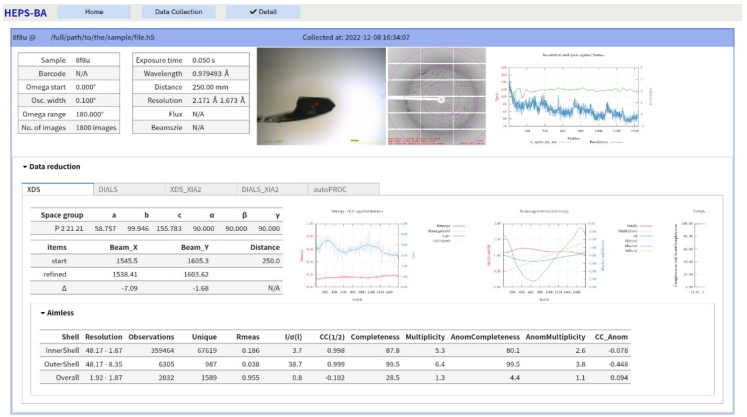


Data analysis framework

- Aim
  - Develop a basic and common high-performance scientific data analysis software framework to address the exabyte data challenge of advanced scientific facilities.
- Daisy Framework
  - Basic framework for integrating algorithms
  - Streaming and distributed data processing to handle high throughput
  - Workflow for flexible and general data processing tasks
  - CI/CD system to accelerate the software development lifecycle
- Scientific applications integration
  - Several HEPS and HXMT Applications integrated
    - HEPSPtycho, Daisy-BMX, Daisy-PDF，HEPSCT，Daisy-HXMT
  - AI-based method is under development
- More Platform supported
  - Several scientific software and algorithms ported from CUDA to the ROCm (AMD machine from domestic vendor)





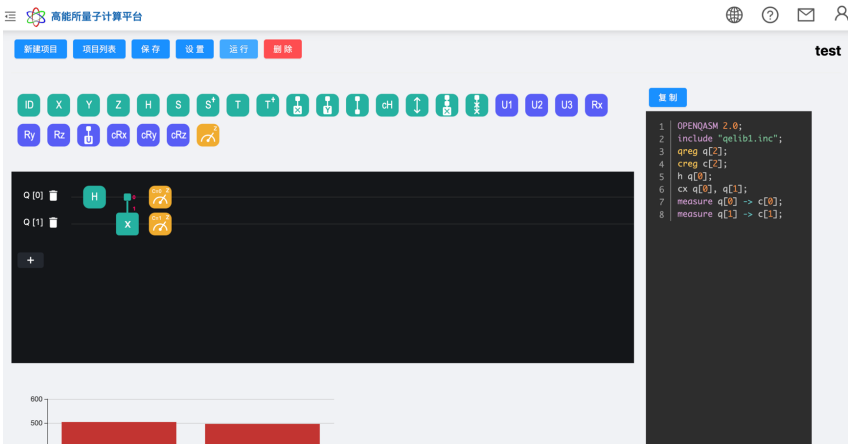


Already applied in synchrotron light source and space astronomy

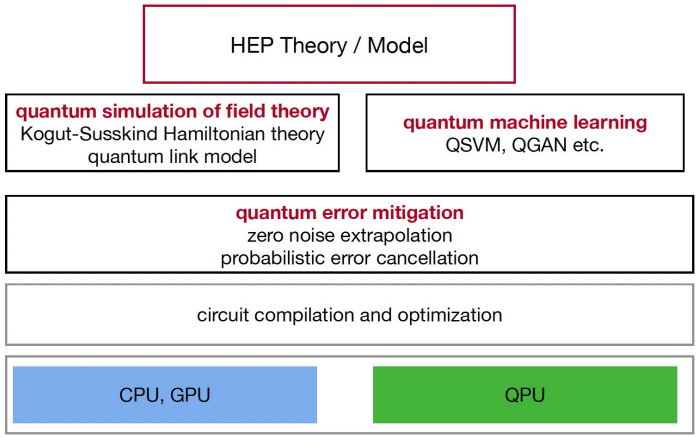
https://daisydoc.readthedocs.io/

Quantum Computing

- IHEP quantum computing platform provided
  - Interactive drag and drop module for learning basic quantum circuits
  - Jupyter based interactive developing environment with CPU and GPU simulator backend
  - High performance qiskit-aer GPU simulator with cuQuantum speed up based on the HPC clusters
- Research on quantum algorithms
  - Utilize quantum machine learning algorithms such as particle transformer
  - Explore the quantum-centric high performance computing in theoretical physics



Web page of IHEP quantum computing platform



Design of IHEP quantum computing platform

Summary

- The computing and storage resource have grown exponentially to satisfied the more exp. Requirement
- IHEP is responsible for the most work of the Chinese grid sites construction and maintenance
- “One platform, multi-centers” provides exp. more resource usage
- Research work at IHEP goes well

Thank You!

Question?