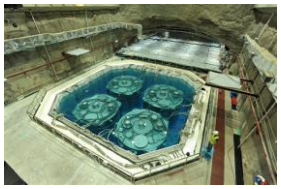
Status of IHEP distributed computing

Xiaomei ZHANG Xuantong ZHANG
Institute of High Energy Physics

10th DIRAC User Workshop
IN2P3 CC, Lyon
June 18, 2024

IHEP

- Institute of High Energy Physics (IHEP)
  - China's biggest laboratory in particle physics
  - Main campus located in Beijing, Dongguan campus in south China
  - Experiment sites spread widely in Dayaba and Jiangmen (Guangdong) neutrino physics, YangBaJing, Ali (Tibet) and DaoCheng (Sichuang) Astrophysics, Huangrou(Beijing) High Energy Photon Source and Dongguan(Guangdong) China Spallation Neutron Source
- Research area
  - Accelerator
  - Particle physics
  - Astroparticle physics
  - Multi-disciplinary research



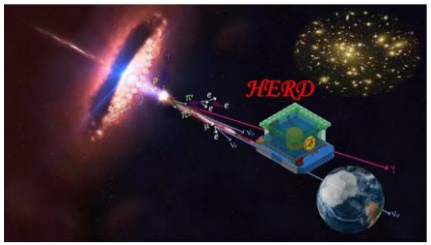
IHEP Experiments

- More than 9 experiments
  - BESIII, CEPC, Dayaba, JUNO, Yabajing, LHAASO, CSNS, HXMT, HERD, HEPS
- Four experiments are already using or plan to use distributed computing
  - JUNO (Neutrino experiment), BESIII (Collider experiment)
  - CEPC (Collider experiment), HERD (Astrophysics experiment)
- Also join many international experiments
  - LHC, BELLEII, PANDA, AMS, COMET……..

BESIII (Beijing Spectrometer III at BEPCII)

JUNO (Jiangmen Underground Neutrino Observatory)

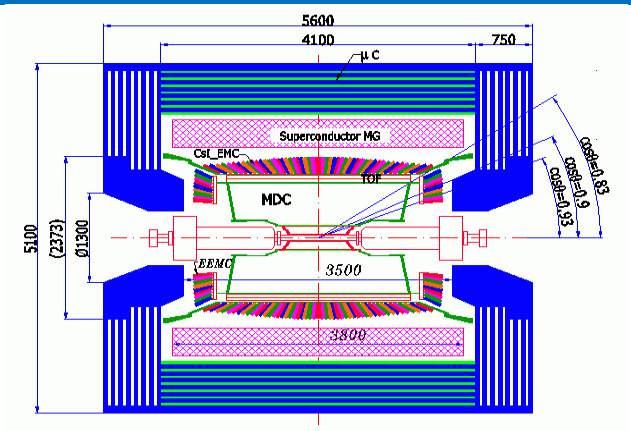
CEPC (Circular Electron Positron Collider)



HERD (High Energy cosmic Radiation Detection Facility)

DIRAC in IHEP experiments

- **BESIII** is the first IHEP experiment to use DIRAC
  - Started to use DIRAC since 2012
- In 2015, **CEPC** started to use DIRAC to organize resource for its detector study
- **JUNO** is the first IHEP experiment to consider distributed computing as main way of resource organization
  - DIRAC-based platform has been put into production since 2021
  - Will start data-taking in 2024



BESIII at the BEPCII accelerator, which is a e+ e- collider, for the studies of hadron physics and τ-charm physics

JUNO - neutrino experiment, located at Jiangmen, in Southern China

DIRAC for future experiments

- CEPC (Circular Electron Positron Collider)
  - Next Generation Accelerator in IHEP
  - CEPC collider is planed to build with the 100KM ring
  - Release CDR in Nov. 2018
  - Work on TDR now, plan to release it next year
- HERD (High Energy cosmic-Radiation Detection)
  - Space particle astrophysics experiment
  - Plan to start in 2027, already get funding
  - 2PB/5year(flight data), ~ 30PB in 5 years (include all data)
  - Will Run for more than 10 years
- Both of them plans to use both DIRAC and Rucio

Multi-VO DIRAC for IHEP communities

- Since more experiments hope to use distributed computing, IHEP plan to build a common distributed computing platform to save operation and development efforts
  - DIRAC as a service, also other common services: FTS, IAM, monitoring……
  - IHEP computing center will take care of “service” part
  - Experiments focus on “application” part with common solutions
- **Multi-VO DIRAC in IHEP** was built since 2015
  - Basic founctions are working: WMS, DMS……
  - But not fully VO independent in some parts: monitoring (job/pilot/transformation/….), management of VO……
- DiracX is expected to be more flexible and convenient on multi-experiment management?

DIRAC set-up in IHEP

- The current DIRAC production is v8r0p39
  - OS: CentOS7
  - One master server in physical machine
    - Silver 4116(2.10 GHz) with 48 cores, 128GB memory
    - Hold all the services
    - WebAppDIRAC with Nginx
  - Two slave servers (in IHEP and in JINR) for load balance
    - IHEP: Configuration server, File Catalogue (FC), TS Agents
    - JINR: Configuration server
- FC redundancy service in slave server is not working stably
    - Connection to the FC Service in slave server often time out
    - Need debugging the problem before reusing it

The DIRAC functionalities heavily used

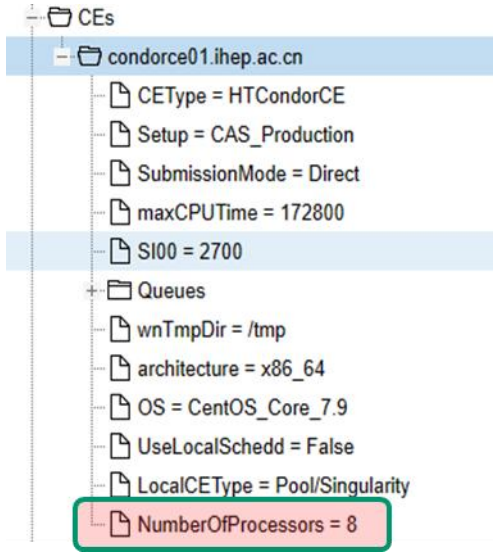
- **Workload management system**
  - Integrate HTCondorCE, ARCCE, SLURM, Cloud……
- **Data management system**
  - **DIRAC File Catalog (DFC)**
    - For both metadata and replica catalog
  - **TS+RMS+FTS3**
    - Bulk data replications
- **Transformation system**
  - MC simulation, reconstruction and Bulk transfer
- **Accounting system**

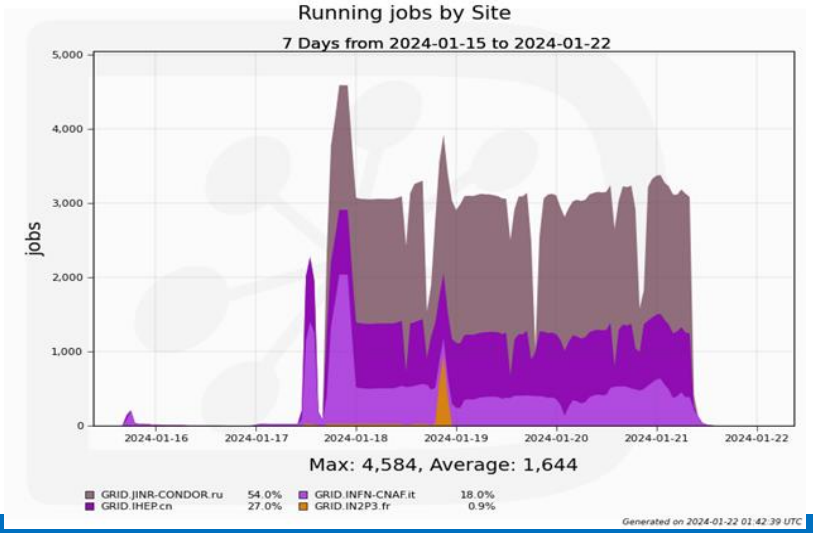
The DIRAC features in use

- Multi-VO
  - Support several experiments in one instance
- SingularityCE
  - Provide a unique sw environment
- PoolCE and tag feature
  - Integrate GPU resource and support multi-core jobs
- Monitoring and accounting
  - System, service, jobs, data transfers…..

Multi-core jobs in JUNO

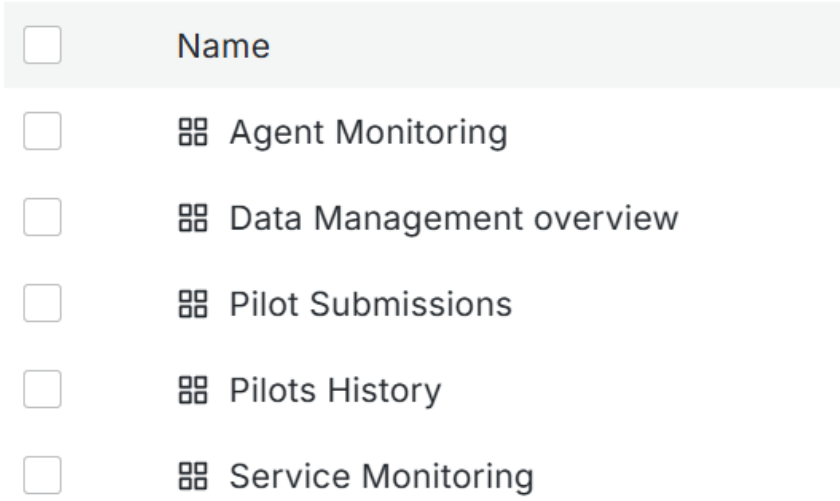
- Prepare to use multi-core for simulation and reconstruction in production
  - Use pool mode and tag mechanism
  - Data challenge with multi-core reconstruction on Jan 17
    - Total ~100K jobs, 8 core/each job (JINR: ~1900 cores, IHEP: ~900 cores, CNAF: ~400 cores)
- Some findings
  - Job accounting not showing number of jobs, but number of cores
  - Not easy to know jobs are running in multi-core or single-core

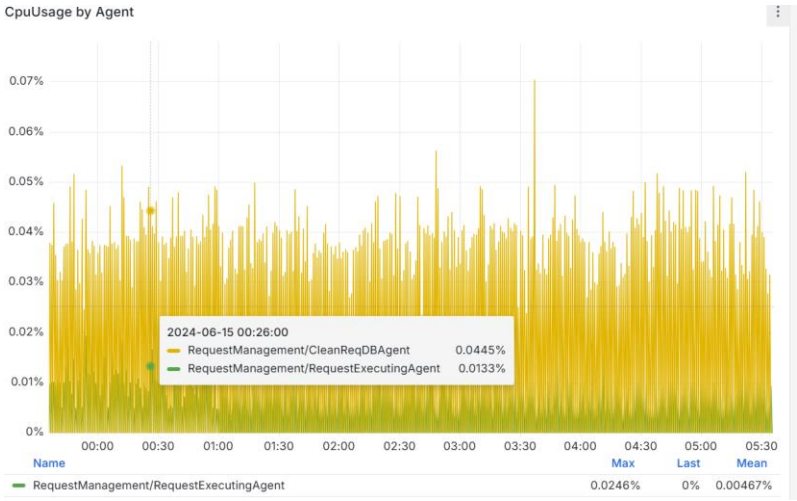




Monitoring

- ElasticSearch/Kibana/Grafana Infrastructure already set up
  - Haven’t moved to opensearch yet
- Monitoring data has been sent from DIRAC to ES
- Create plots out of data in ES to meet JUNO DCI requirements
- Design shifter dashboard with these plots



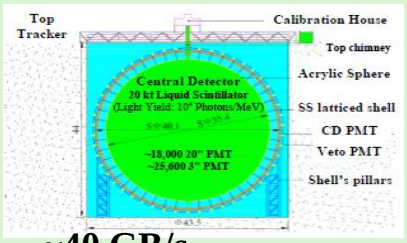


IHEPDIRAC extension

- Production and analysis activities
  - **Production system** for MC production and reconstruction
  - **Frontend** for user analysis
- Data operation
  - **Raw data distribution system** from IHEP to other data centers
  - **Operation tools**: bulk registration, bulk cleaning, consistency check….
- Monitoring
  - **SAM tests services and agents**, Extension to RSS

JUNO raw data processing

JUNO onsite



~40 GB/s

Trigger + OEC

60 MB/s

RAW (t/q, waveform)

DAQ Event Rate: 1kHz

Data Volume: *60 MB/s*

Each file size: *5 GB/file*

N (Events) per file: *83,333* (=5GB / 60MB/s * 1kHz)

IHEP

RAW data Volume: *5.2 TB/day, 2 PB/year*
Number of files: *~1k/day, 400k/year*
RTRAW data Volume: ROOT format of RAW, *2/3 RAW*

2 Gbps network (dedicated)

RAW (t/q, waveform)

Tape

RTRAW

Sampling + Reconstruction

Reconstruction

ESD (KUP)

IAD (KUP)

Data Quality Monitoring

Offline Data Monitoring

Delay: ~hour

Delay: ~half day

JUNO Data Monitoring Web

JUNO Data centers

Requirements: one year data needs to be reconstructed within 2 months.

100 Gbps network (LHCONE)

RAW (t/q, waveform)

Tape

RTRAW

Reconstruction (new calibration)

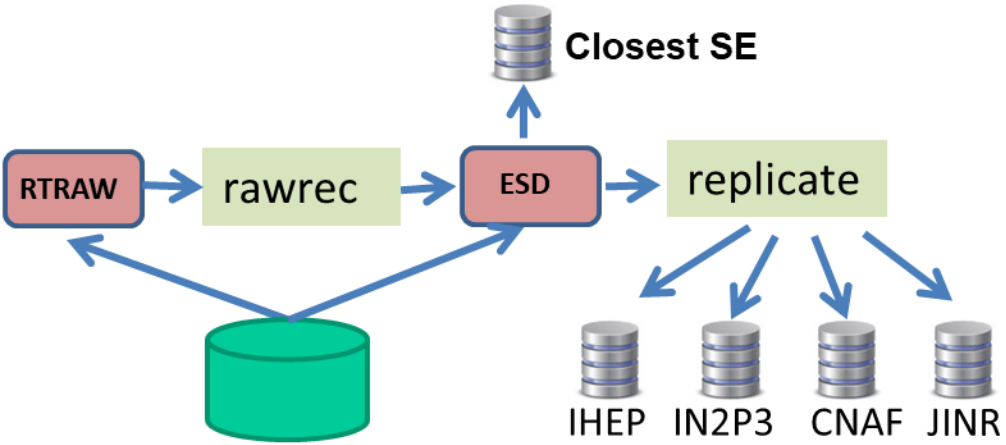
IAD (PP)

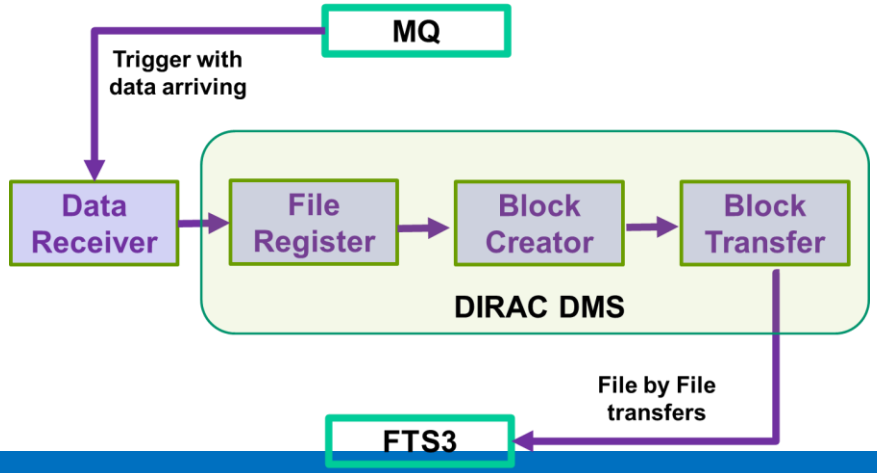
ESD (PP)

Physics Analysis

JUNO data challenge 1/2

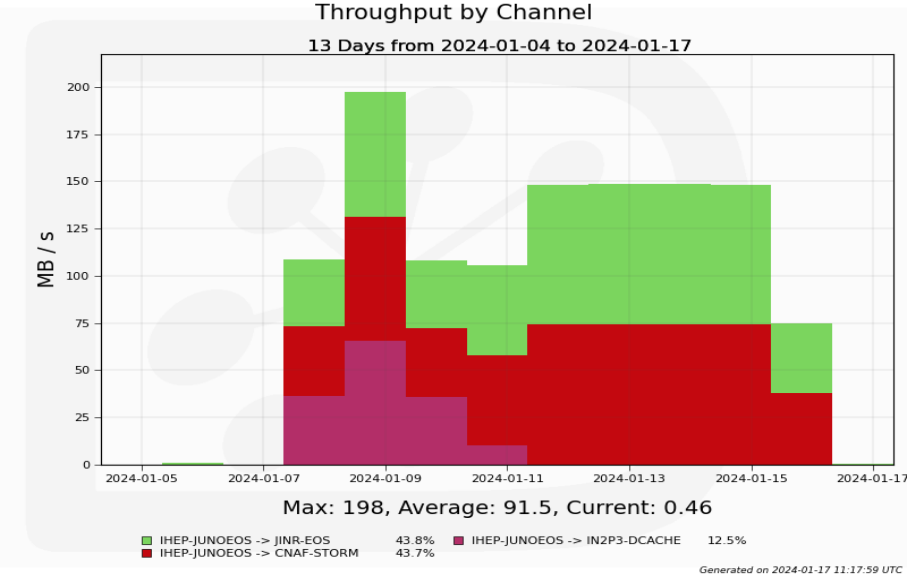
- From second half of 2023, JUNO started offline data challenges to prepare for the coming data-taking
- The first data challenge is to simulate and validate the complete raw data processing chain for the first time, JUNO DCI part covering
  - Receive RTRaw data from online chain
  - Register and replicate RTRaw from T0 to T1s
  - 2nd Reconstruction run in all data centers
- All the implementation are based on DIRAC DMS and transformation

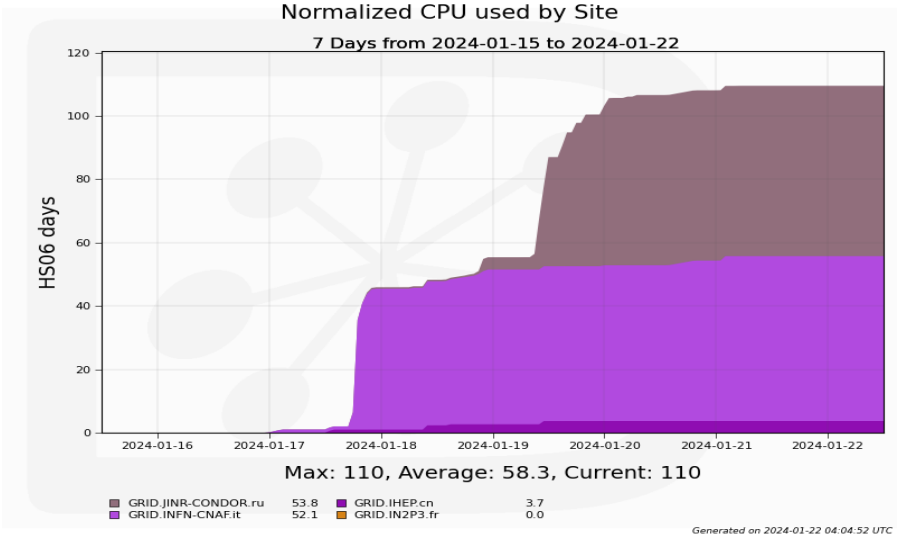




JUNO data challenge 2/2

- Tests show the system is working fine
  - 7 blocks, 103TB data were transferred to data centers in time with 100% success rate
  - ~100K jobs in four days
- More challenge tests are being prepared
  - Try to understand performance and bottleneck





Man power

- The colleague from IHEP computing center Xuantong Zhang joined us
- IHEPCC plans to have one more employee and student to join the DIRAC group
- Our postdoc position to work on DIRAC or DiracX keeps open
  - Not only for JUNO, also Herd and more experiments

Summary and plan

- JUNO starts to prepare for its data taking with data challenges
- DIRAC DMS provides an easy solution on automated raw data management system
- A common infrastructure and solution is in considerations to reduce operations efforts for more IHEP experiments
- Thank for help from DIRAC communities!