CEPC Computing

Xiaomei Zhang

Feb 13, 2023

HK-IAS mini-workshop

HKUST, HongKong

Contents

- Status of CEPC computing
- WLCG evolutions and follow-ups
- R&D activities and outlook
- Summary

Data estimation and challenge

- **Very preliminary estimation** of data volume at data-taking
  - ~3PB/year in Higgs/W factory, modest comparing to LHC Run2
  - Reach EB scale in Z factory, comparable to HL-LHC, but a few years after
- It is helpful to join WLCG collaborations and benefit from WLCG experience
- **Distributed computing (grid computing)** is planned for CEPC
  - Organize resources for data processing activities in both R&D and data-taking
- **The CEPC distributed computing system has been built up**
- **The system was proved to work well for R&D detector simulation**

CEPC Distributed Computing

- **DIRAC** is chosen as CEPC distributed computing framework
  - Originally from LHCb, now used for many other experiments: BELLEII, ILC, CTA, SKA…..
  - **DIRAC Consortium** establishes a strong collaboration on the DIRAC project
  - **IHEP has joined DIRAC Consortium since 2016, and get much support from DIRAC team**
- **CVMFS** for software distribution
  - stratum0 operated @IHEP : /cvmfs/cepc.ihep.ac.cn/, stratum1 @IHEP and @RAL
- **VO CEPC** for resource authentication and authorization
  - VOMS hosted @IHEP : https://voms.ihep.ac.cn:8443/voms/cepc/
- **CEPC users can access resources everywhere with web or client**
  - Web sites: https://dirac.ihep.ac.cn
  - IHEPDIRAC Client in cvmfs: /cvmfs/dcomputing.ihep.ac.cn/dirac/IHEPDIRAC/

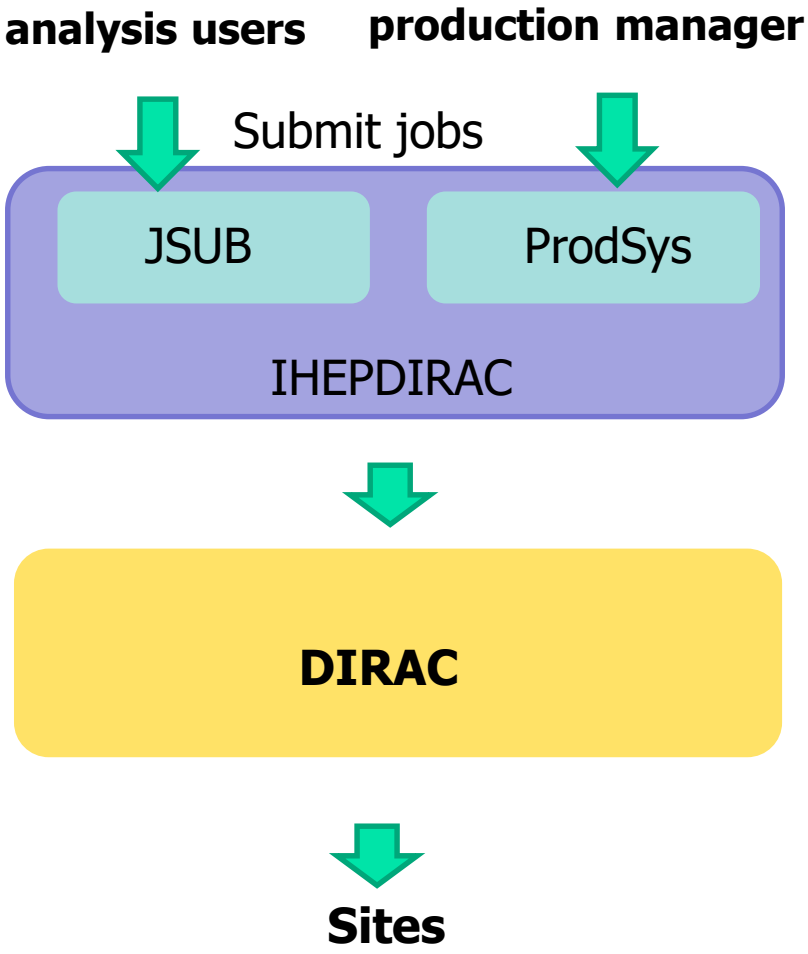
Current status

Resources and sites

- About 4600 cores in the system
  - IHEP has dedicated resources
    - CPU: 2000 cores, 640 cores shared with ILC in 2022
    - Storage: 2.5PB
  - Five joint sites from UK and other China universities
    - ~2600 CPU cores
    - Shared with other experiments
- Last year ARM resources have been tested and added to the system

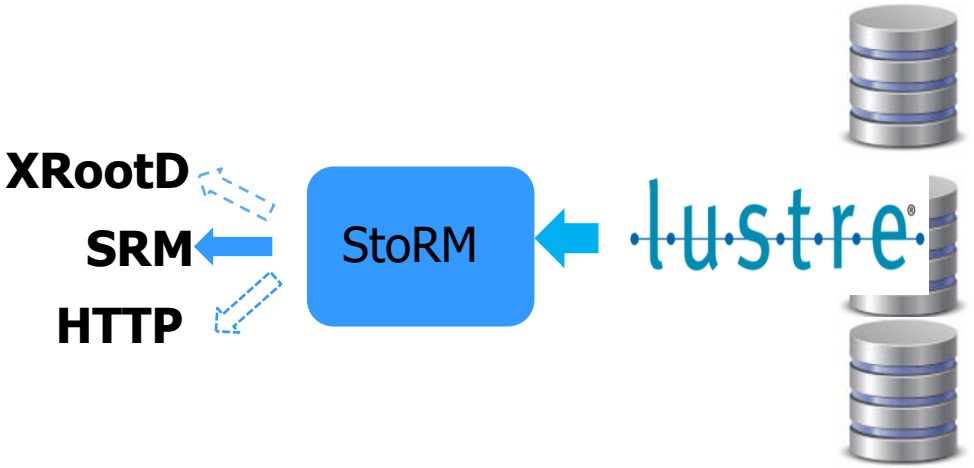
Workload management and Job submission

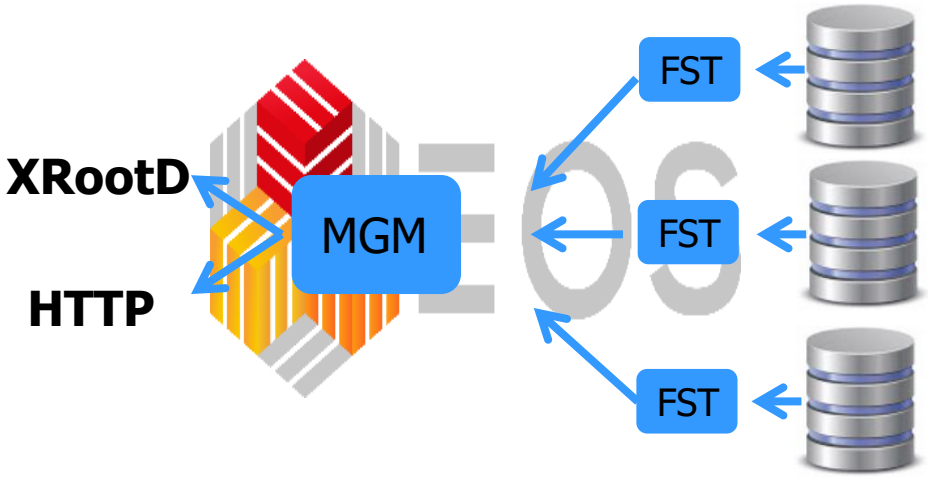
- **DIRAC**
  - Provide a middle layer between jobs and resources to hide complexity from users
  - Allow to support the connection to various resources: cloud, grid, cluster…..
  - v7.3.21, **migrated to python3 in 2022**
  - Plan to upgrade to v8.1 to get token support in 2023
- **IHEPDIRAC (v2.1 with python3)**
  - Extension to DIRAC, provide interface and tools for running CEPC applications
  - **Two job submission tools in prototype**
    - **ProdSys:** developed to submit and manage MC simulation tasks
    - **JSUB :** developed as a job submission frontend for analysis users
- **CEPC users can use IHEPDIRAC interface to submit jobs**



Central SE (Storage Element)

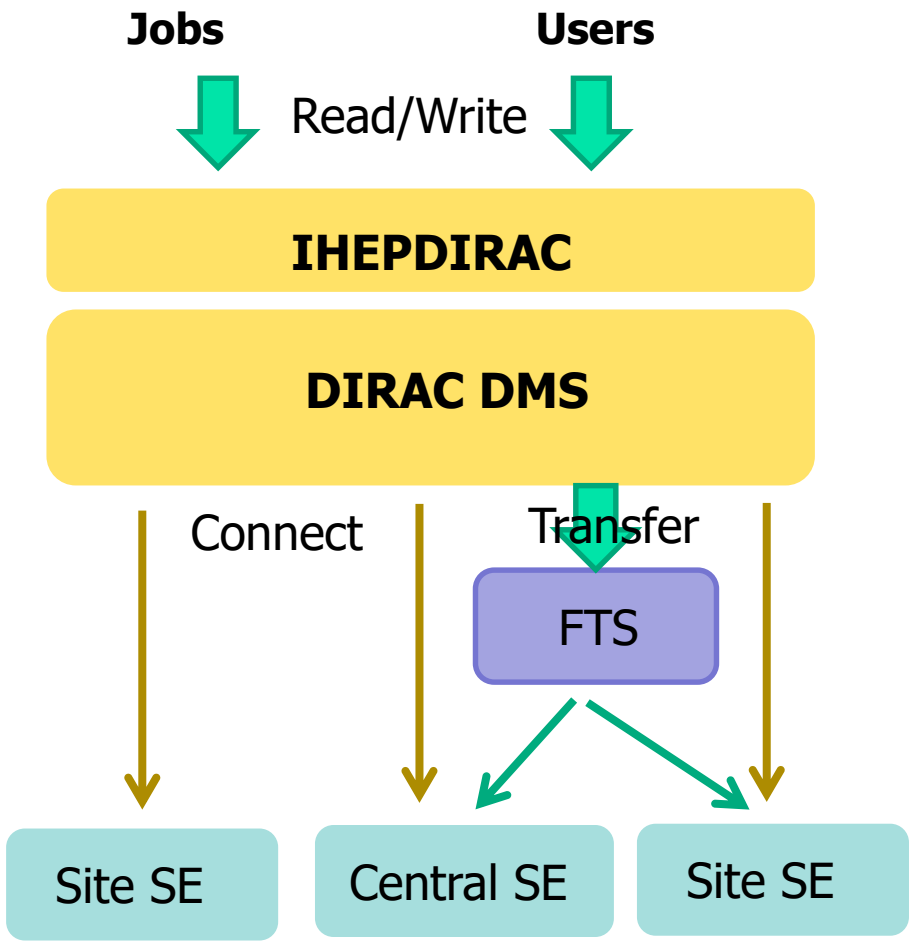
- StoRM is used as central SE
  - Lustre is backend file system
  - Hosted @IHEP, just upgraded to the latest version v1.11.21
  - Support both XRootD and HTTP TPC
  - Support macaroon token and scitoken
- In future, plan to migrate to EOS
  - EOS will become main storage in IHEPCC
  - EOS instance has been tested and proved to be able to provide HTTP TPC
  - Work on support to scitoken
- Have close contact with StoRM team in CNAF and EOS team at CERN





Data management

- **DIRAC DMS (Data Management System)**
  - Provide a global view of data in grid environment, including
    - File Catalogue: mapping of LFN to PFN
    - Meta Catalogue: dataset management
- **FTS (File Transfer System)**
  - Manage large number of file movements between different SEs
  - fts3 server installed at IHEP: https://fts3.ihep.ac.cn
- We also use the infrastructure for other IHEP experiments: JUNO, BES



Data access

- With DIRAC client, CEPC users are able to query, access and share CEPC data in grid environment:
  - ***-bash-4.2$ dirac-dms-filecatalog-cli***
    - *Starting FileCatalog client*
    - *FC:/> cd /cepc*
    - *FC:/cepc>**ls -l***

    drwxr-xr-x 0 yyang cepc_user 0 2021-12-19 13:30:28 cefs

    drwxrwxr-x 0 zhangxm cepc_user 0 2022-10-13 06:23:08 lustre-ro

    drwxrwxr-x 0 zhangxm cepc_user 0 2022-10-13 06:22:58 user

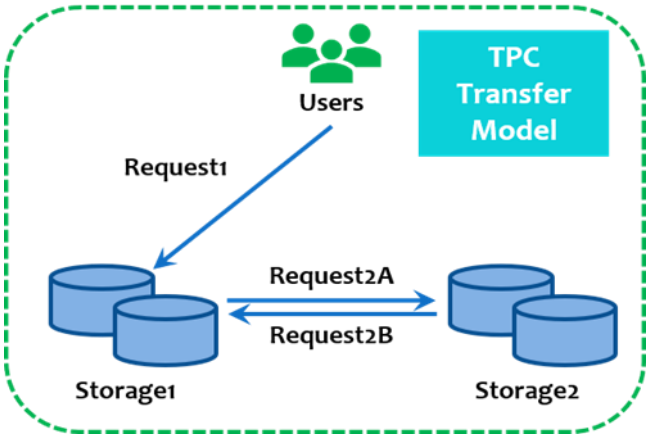
    - *FC:/> **get /cepc/user/z/zhuyf/test_001/sim/zz_sl0mu_down.e0.p0.00035_sim.slcio***

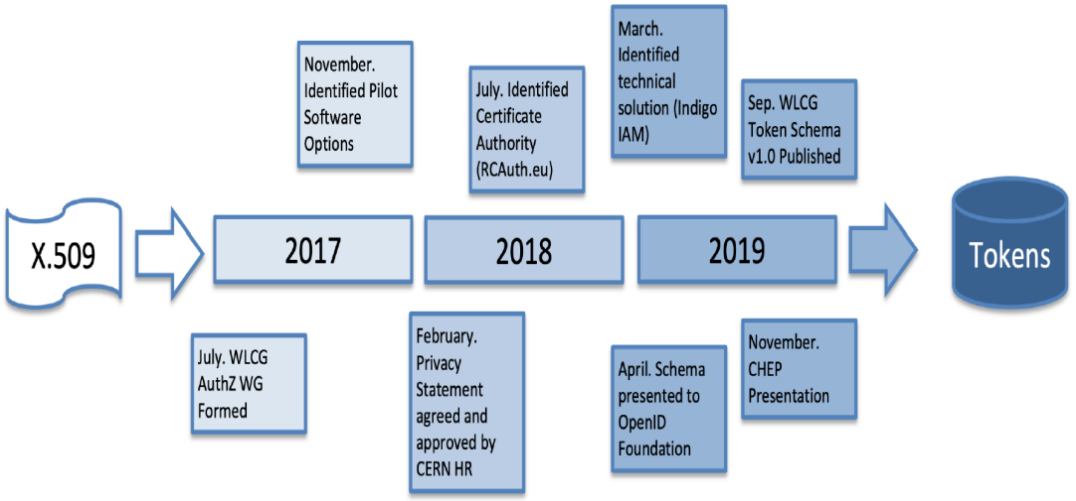
    File /cepc/user/z/zhuyf/test_001/sim/zz_sl0mu_down.e0.p0.00035_sim.slcio successfully downloaded

WLCG evolutions and follow-ups

WLCG evolutions

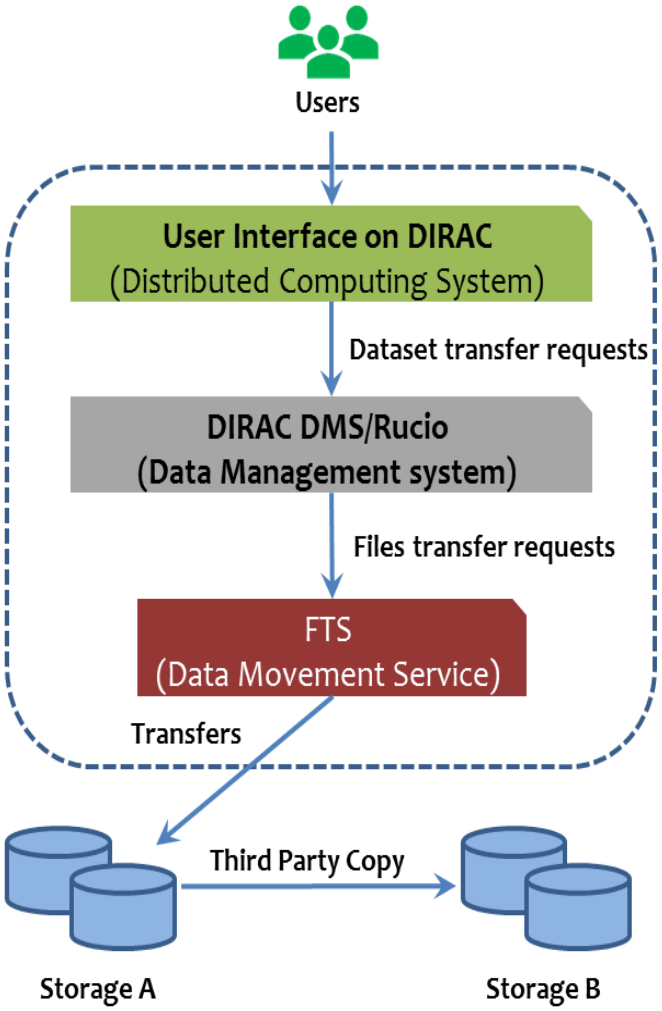
- CEPC distributed computing is using the WLCG middleware, which allow smooth communications with WLCG grid sites all over the world
- WLCG has evolutions on TPC protocols and AAI in recent years
  - **Webdav (HTTP)** is taking place of Gridftp as the next generation TPC (**T**hird **P**rotocol **C**opy) baseline protocol for data transfer
  - **AAI (Authentication Authorization Infrastructure)** adopted OAuth& OpenID industry standards, Token-based AAI taking place of x509-based one
- Works are being done to follow up the changes





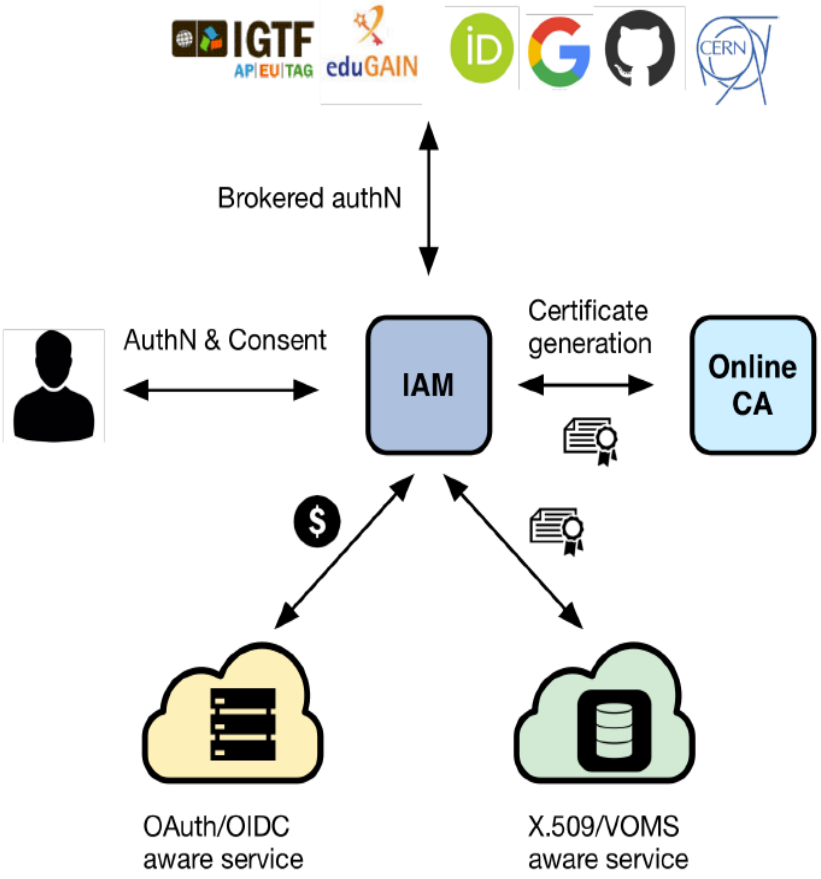
Status of migration to new TPC

- TPC protocol defines rules for grid SE to talk to each other during data transfer
- Porting to new TPC is important to ensure data transfer and sharing in grid env
- The whole DMS system has finished migration to new TPC
  - **DIRAC** upgraded to v7 beyond and allowed configuration of new TPC
  - **FTS server** upgraded to the latest version to allow transfers over new TPC
  - **EOS** enabled support of HTTP TPC
  - **StoRM** upgraded to get supports
- The transfer tests over HTTP TPC protocols have been done



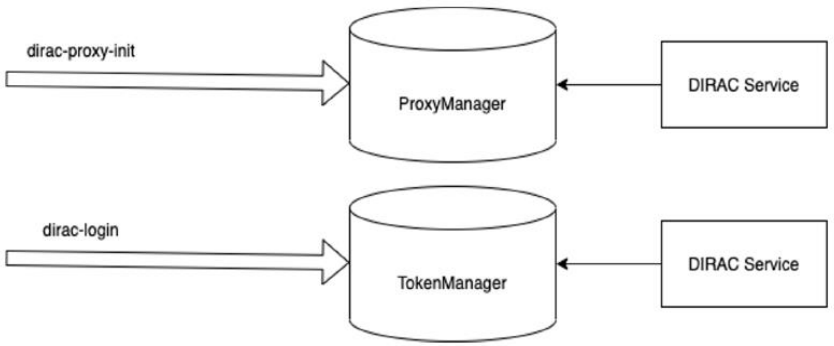
IAM-based new generation AAI service

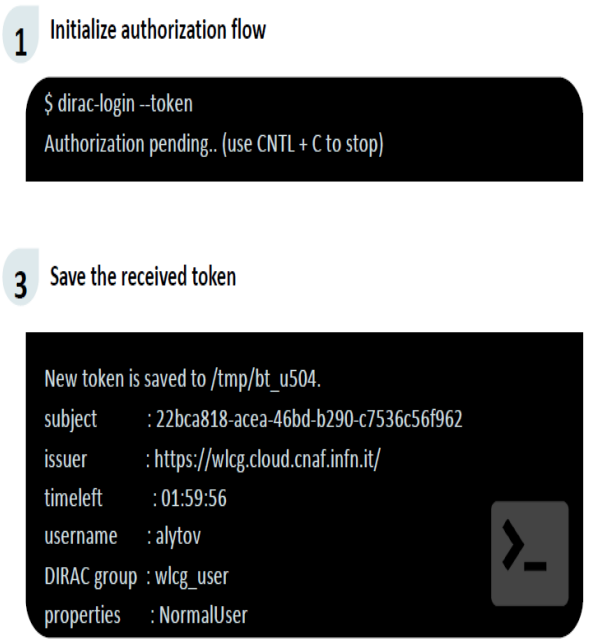
- AAI play a great role in controlling access to grid resources for security and priority
- IAM (INDIGO Identity and Access Management) is adopted to be core of the new token-based AAI in WLCG, taking place of VOMS which is core service of X509 AAI
  - Support multiple authentication mechanisms, standard OAuth & OpenID protocols…..
- IAM is To ensure smooth transition, IAM and VOMS will be kept working in parallel until the whole system moved to new AAI
- IAM also contains features to support this transition period
  - Interact with onlineCA to get user certificate
  - Integrate existing VOMS to produce VO-aware proxy

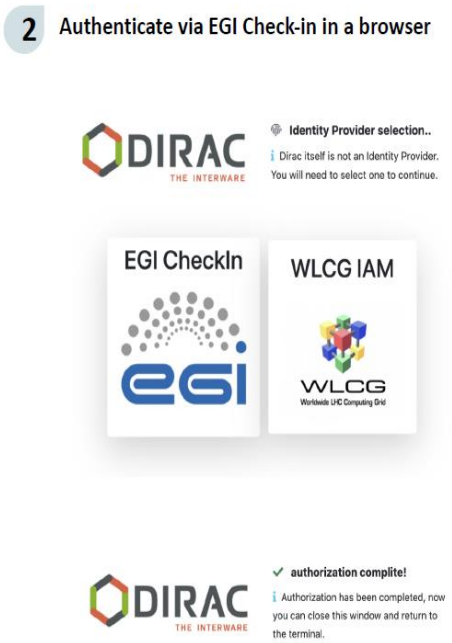


Status of migration to token-based AAI

- **IAM testbed has been set up in IHEP**
  - Keep close contact with CNAF IAM development team
- IHEP EduGain was connected to the IHEP IAM instance, allowing fast user registration
- The whole system is being upgraded and tuned to have a full support of token AAI
  - **DIRAC** plan to migrate to the version v8.1 beyond
  - **Site CE**
  - **Site SE**
  - **FTS**







R&D activities and outlook

Future challenges

- Current usage is small compared to when data-taking is started
- With growing usage in future, the challenges are foreseen in many aspects:
  - Increased resource requirements due to growing data volume
  - Diverse resource requirements due to software evolution
  - More complicated data provisioning with increased usage of diverse resources
  - More operation costs with the system growing big and complicated
  - .....
- To prepare for these challenges, we start considering
  - Heterogeneous and opportunistic resources integration
  - “Data Lake” model
  - Wider collaborations

Heterogeneous resources 1/2

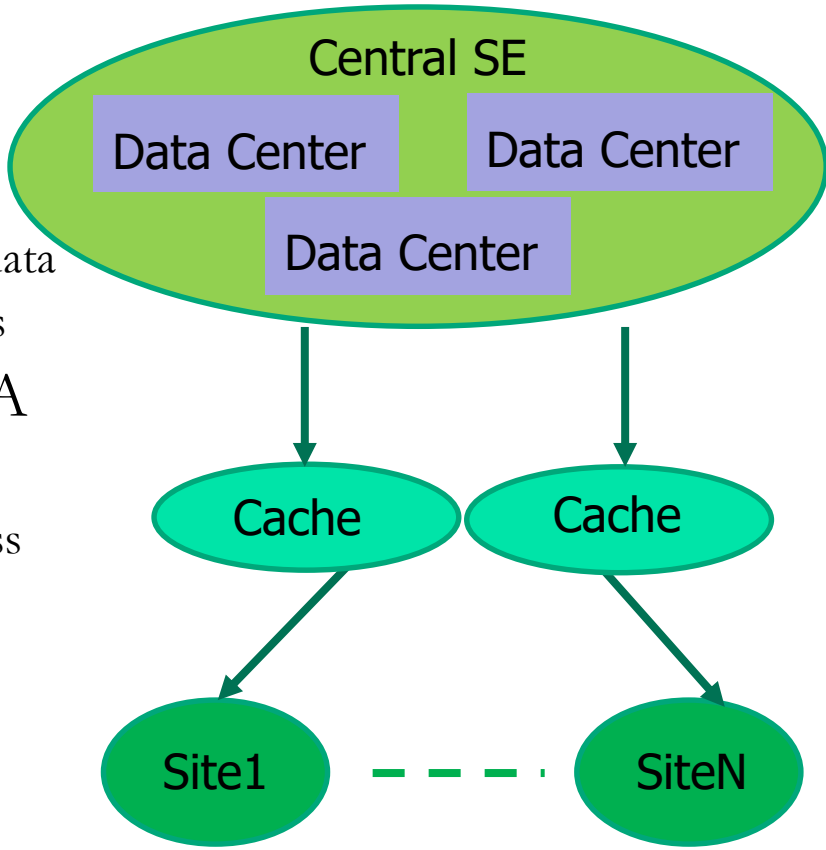
- Hardware and software evolution is expanding resource heterogeneity
  - More hardware type emerging: CPU (x86_64, Power, ARM…), GPUs, FPGA…
  - Heterogeneous platforms (CPU + accelerators) became common used by HEP software, eg. ML training and inference
- Based on DIRAC infrastructure, we started studying on integrating more diverse resources in the system
  - ARM resources
    - ARM computing cluster with 100 worker nodes (~10K CPU cores) set up in IHEP
    - In 2022 CEPC distributed computing integrated these resources smoothly
      - Thanks to the flexible infrastructure design of DIRAC (diracOS)
  - GPU resources
    - Aim to build a "grid" of GPU, together with CPU resources
    - Integration of GPU with CPU resources is successful
    - Need further solutions on outbound network connection of some GPU farms

Heterogeneous resources 2/2

- Commercial and private clouds
  - The platform is able to seamlessly integrate with various kinds of cloud, allowing CEPC application to use cloud resource in an elastic way
  - Tests have been done with several type of clouds for CEPC simulation jobs
    - Tests with Amazon AWS China region cloud looks promising
- Supercomputing centers
  - LHC experiments has much efforts on exploiting HPC resources in supercomputing center in the past years
  - Different supercomputing centers have their own policy
  - Main bottleneck: No outbound network, not CVMFS available, VPN access….
  - Started investigating supercomputing centers, still need to spend more efforts

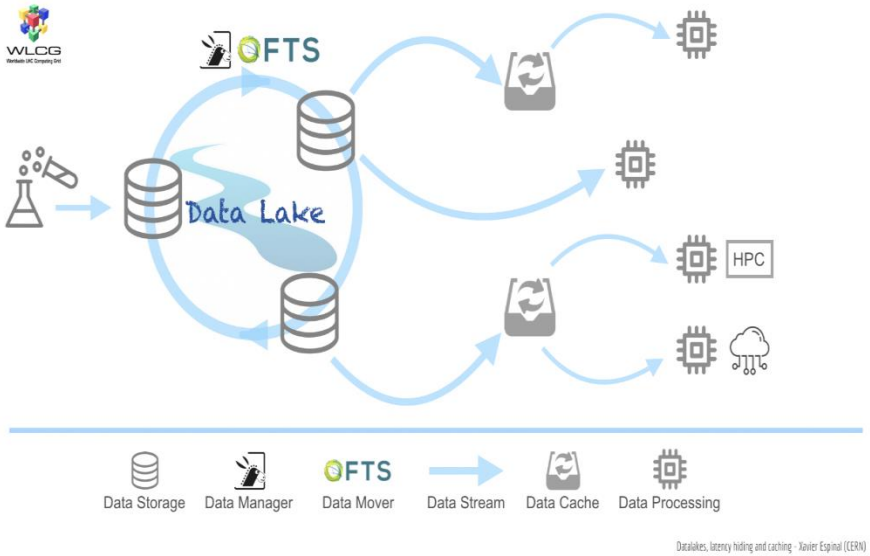
“Data Lake” model

- For future scaling up, no problems to handle 100 times jobs in this platform
  - DIRAC is able to handle >100k jobs in peak
- Data provisioning could be a challenge
  - Single-point central storage could be a weak point
  - Diverse resource need more flexible strategies to access data
  - Operation costs for a SE can’t be affordable to small sites
- **“Data Lake” model** from the WLCG DOMA group is expected to be a solution
  - Storage is federated among big data centers for robustness
  - Cache layers provide light stateless storage for sites
  - More efficient: Data is only replicated when it is needed
- Advanced data management system and data access policy are needed
  - Rucio and XCache could be a choice to achieve that



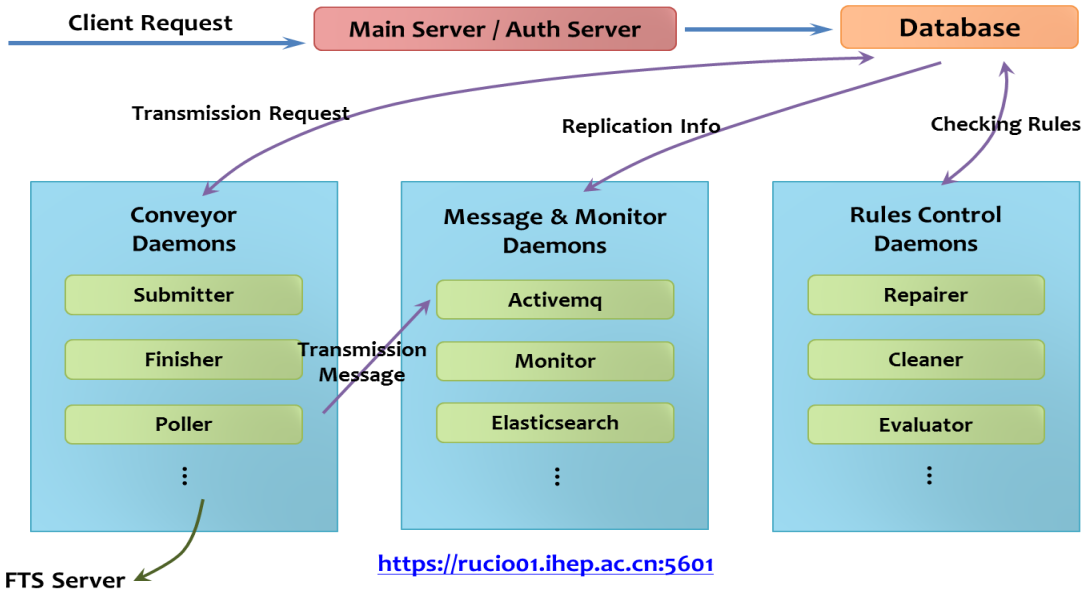
Rucio

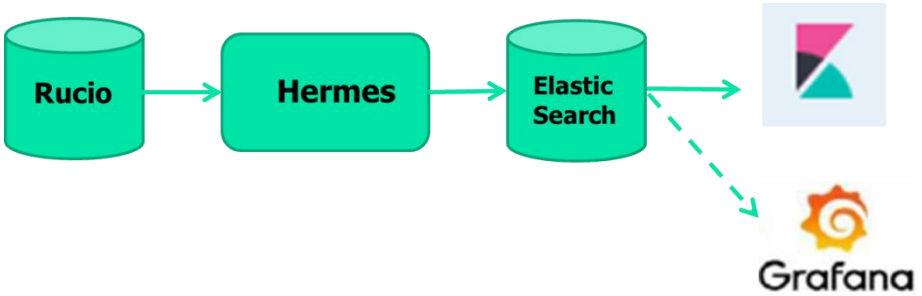
- Rucio is a data management system which can provide the functionalities needed to manage SEs, data and data flow globally
- Rucio is developing into a common standard for scientific data management, and widely evaluated and used by many experiments
- In the "Data Lake" model, Rucio is being considered to federate storage system in big data centers and provide a single entry point to outside



Status of Rucio study in IHEP

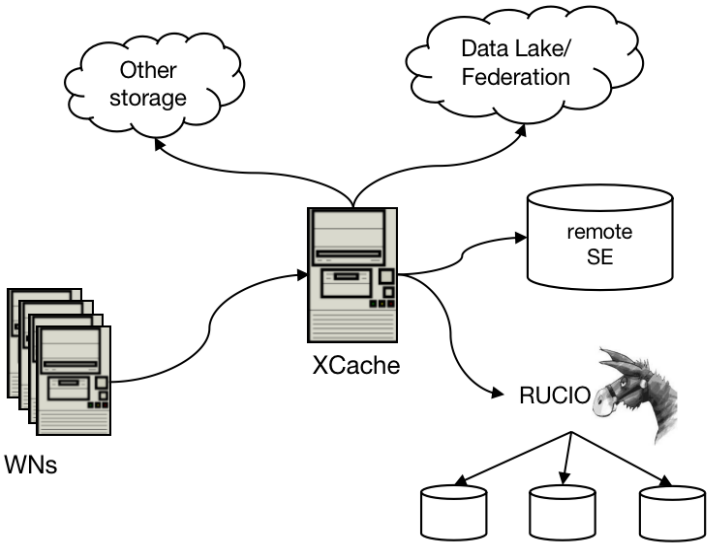
- Rucio testbed has been deployed with containers in IHEP
  - Main server and Authority server
  - Run independent daemons for transmission and monitoring
- Various SEs registered
  - ➔ StORM, EOS, dCache
- Connected to IHEP FTS server
- Monitor has been set up
  - Hermes daemon in Rucio collects info of data and transfers
  - Info is sent to ES and shown in Kibana, or later in Grafana
- Seamless integration of DIRAC and Rucio is under consideration
  - In DIRAC highly extensible infrastructure, Rucio embeded as DIRAC services

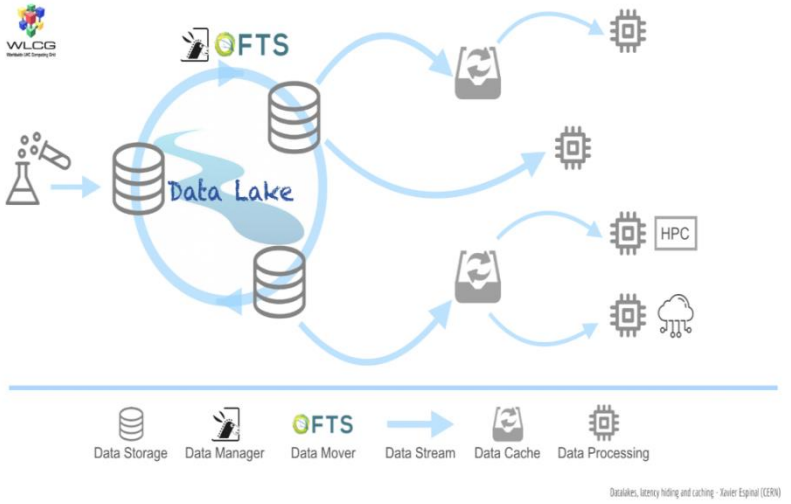




XCache

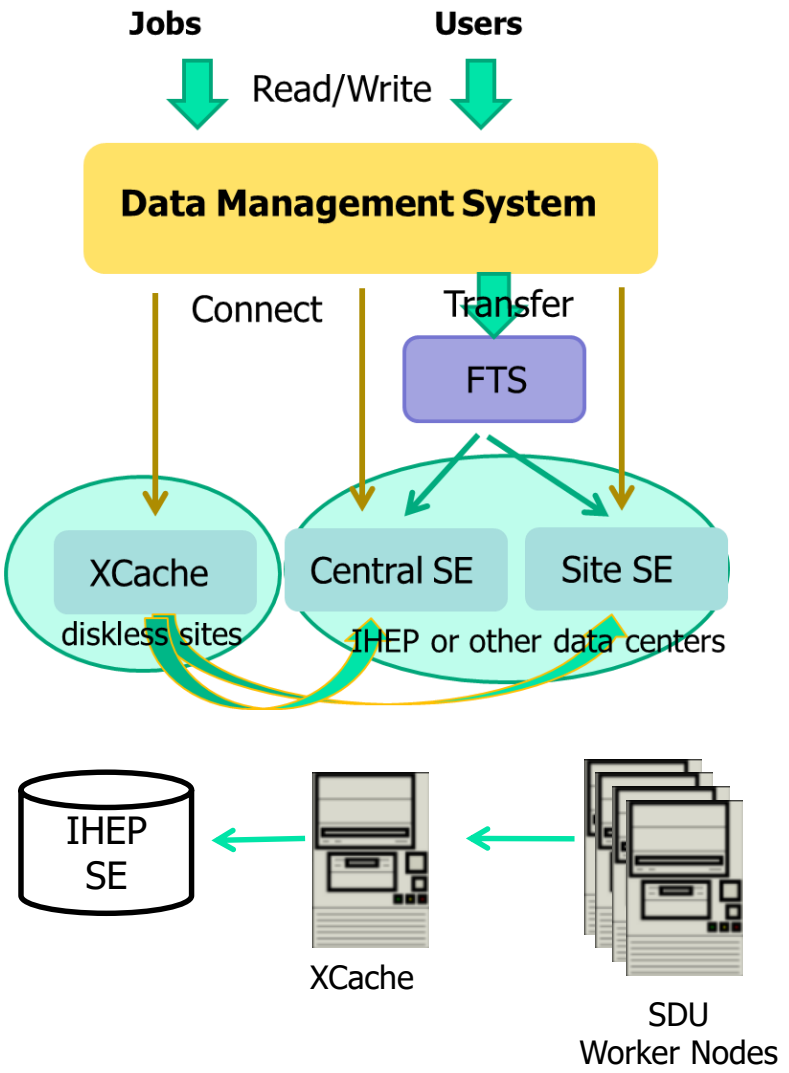
- XCache(XRootD Cache system) is a squid-like cache system, supporting ROOT and HTTP protocol
- With plugin architecture, XCache is clusterable, easy to scale up
- XCache can achieve highly efficient remote data access with high cache hit
- In the “Data Lake” model, XCache is considered for building cache layer
  - Act as volatile SE for the sites without enough manpower for a decent SE
  - Can also help reduce burden of central SE and network traffic





Status of XCache study in IHEP

- XCache server has been set up
- Register XCache into DIRAC as a volatile SE for small sites
  - Setting data source from decent SEs
- The job submission tool is configured to allow XCache-aware access to data when jobs arriving in diskless sites
  - JSUB checks with DIRAC and uses XCache when downloading input data
- Tests were done at SDU, a typical site
  - Jobs running in SDU can automatically read data through XCache which can use IHEP SE as data source
  - Improvements of data access with XCache were well seen from the tests



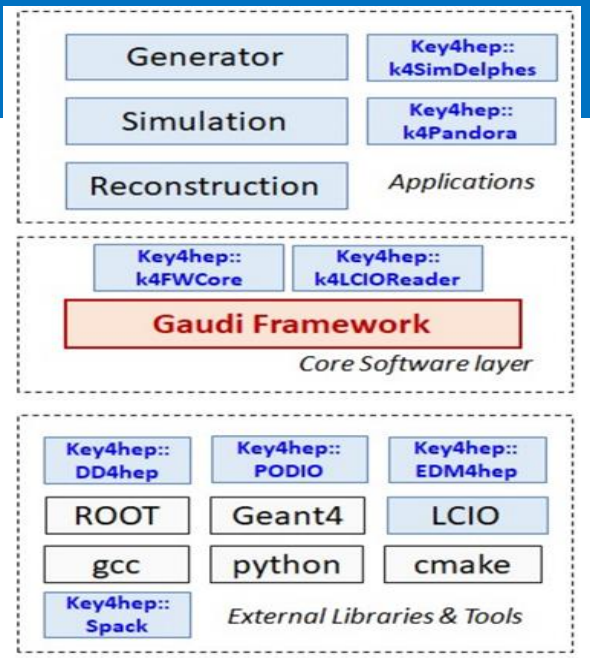
Collaborations for sustainability

- The future challenge is not only in resource and technologies, but also sustainability
  - Grid is a big system, need a lot of development and operation efforts
- The European Strategy of Particle Physics and WLCG collaborations have proposed on establishing a scientific computing infrastructure, in which
  - HEP communities can collaborate and share services, expertise and operations to reduce costs
  - Some co-operations have been established

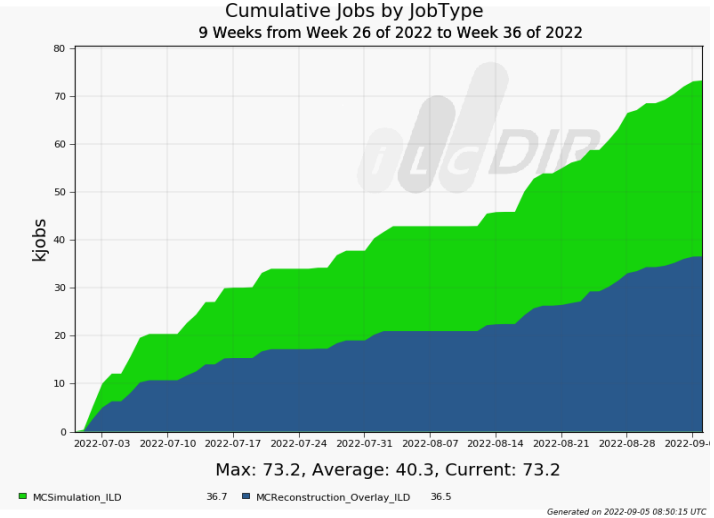
From S.Campana, Computing - challenges and future directions (ECFA 2021)

CEPC collaborations

- CEPC shares the same challenges and concerns on software and computing with other experiments, especially ILC, FCC, CLIC
- **CEPC computing is seeking opportunities to get closer to the global collaborations on common infrastructure and tools**
- For software, CEPC has joined Key4hep and Edm4hep
  - ➔ CEPCSW is fully integrated in common software stack
- For computing, we are
  - ➔ Having common needs on software and operations
    - Workload management and Data management
    - Job submission, data transfer, database access, monitoring…….
  - ➔ Using same systems
    - Core Framework – DIRAC
    - Data transfer tools – FTS
  - ➔ Sharing resources through Grid
    - In 2022, 640 CPU cores in IHEP shared by CEPC and ILC
- We are trying to do more in the following years



From Xingtao Huang, 2021 CEPC workshop



Summary

- CEPC already had a basic distributed computing platform for the detector R&D
- Following WLCG evolutions on TPC and AAI, the platform is adapting to these changes
- CEPC computing is expected to face challenges in data-taking
- But CEPC computing is not alone in these challenges, seeking wider collaborations with other future experiments

Thank you!

User job submission

- **CEPC users can use IHEPDIRAC client to submit jobs**
- **DIRAC interface for job submission**
  - *source /cvmfs/dcomputing.ihep.ac.cn/dirac/IHEPDIRAC/bashrc*
  - *dirac-proxy-init –g cepc_user*
  - *dirac-wms-job-submit yourjob.jdl*
- **JSUB for job submission**
  - Help manage life cycle of an analysis task in an automatic way and make a large number of job submission more convenient
  - *source /cvmfs/dcomputing.ihep.ac.cn/dirac/IHEPDIRAC/bashrc*
  - *source /cvmfs/dcomputing.ihep.ac.cn/frontend/jsub/activate.sh*
  - *Jsub create yourjob.yml; jsub submit yourjob.yml*

Example of JDL files:

```
Executable = "Application.sh";
Arguments = "";
InputSandbox = {"Application.sh"};
StdOutput = "std.out";
StdError = "std.err";
OutputSandbox = {"std.out","std.err"};
```