JUNO Distributed Computing System

Xiaomei Zhang

Institute of High Energy Physics

On behalf of JUNO DCI group

CHEP2023, May 8

JUNO

- ❖ A multi-purpose neutrino experiment

  - Measure neutrinos (solar neutrinos, supernova neutrinos, atmospheric, geo ....) mass hierarchy and mixing parameters

  - Located at Guangzhou, China

  - Expect to take data in 2024

- ❖ JUNO-TAO is a satellite detector

  - Precisely measure reactor energy spectrum, improve sensitivity of JUNO on mass hierarchy study

- ❖ Data volume expected

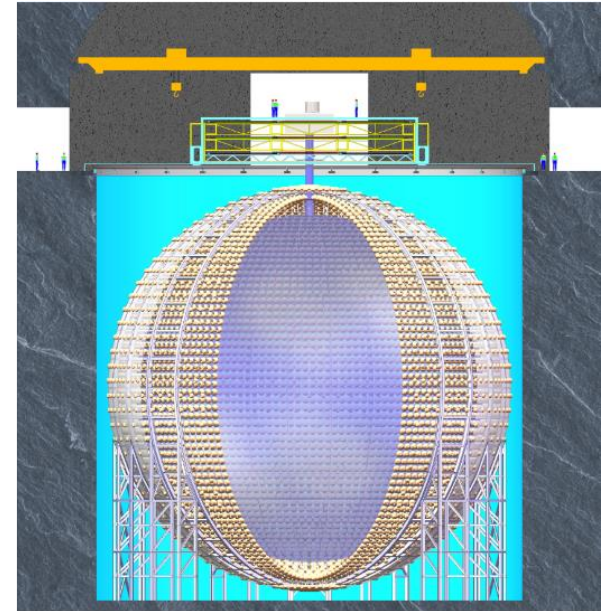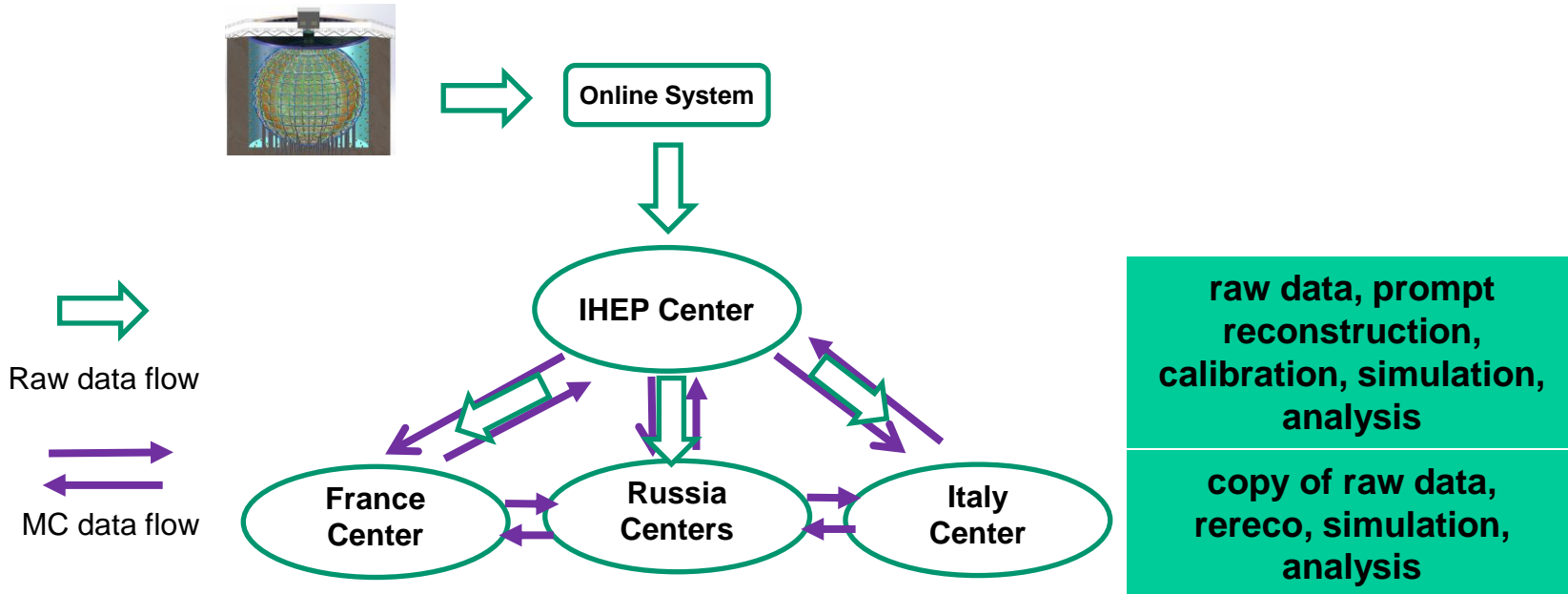 - Raw: 2.4PB/year (JUNO+TAO)

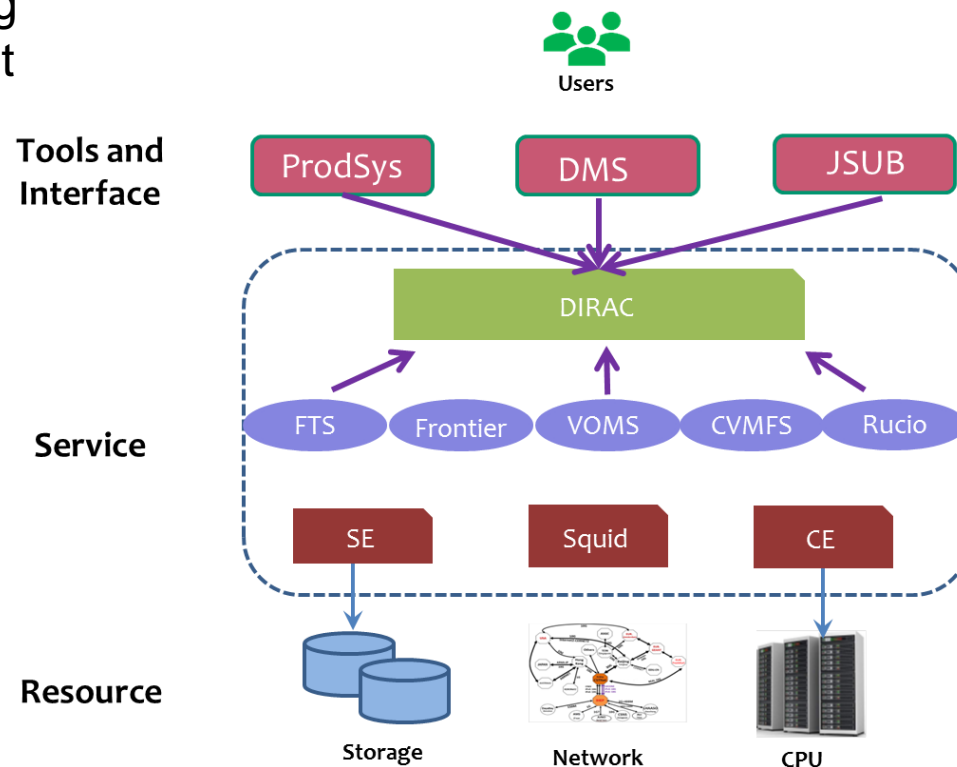 - MC+Rec: 600TB

Photo taken in 2023.2

Data centers and Computing model

raw data, prompt reconstruction, calibration, simulation, analysis

copy of raw data, rereco, simulation, analysis

- ❖ Five data centers: IHEP, CC-IN2P3, INFN-CNAF, JINR, MSU

- ❖ Raw data flows from Online to IHEP which then immediately distributes to other centers

- ❖ 1st Reconstruction and Calibration will run in IHEP

- ❖ MC Simulation, 2nd Reconstruction and Analysis are expected to run in all data centers

- ❖ Other centers provide a backup to JUNO data (CNAF/JINR 100%, IN2P3 1/3)

System Architecture

❖ JUNO distributed computing system was built to take care of data processing and data distribution in grid environment

❖ DIRAC is core of the system
  ❖ Organize heterogeneous resources
  ❖ Provide framework for workload management (WM) and data management (DM)
  ❖ Integrate necessary middleware and services

❖ Other WLCG services used
  ❖ VOMS/IAM, authentication and authorization
  ❖ FTS, file movement
  ❖ CVMFS, software distribution

❖ Experiment tools and Interface (details in later slides)
  ❖ JUNO-specific systems developed to meet the requirements of JUNO data placement and processing
 ❖ All codes were migrated to python3

Tools and Interface

Service

Resource

Users

ProdSys   DMS   JSUB

DIRAC

FTS   Frontier   VOMS   CVMFS   Rucio

SE   Squid   CE

Storage   Network   CPU

ProdSys 1/2

- ❖ Implemented as a data-driven pipeline system, designed to
  - submit JUNO production tasks (simulation, re-reconstruction…) in grid env
  - manage workflow and dataflow in the tasks automatically
- ❖ Each JUNO production task is composed of several steps
  - Detector simulation (detsim), Electronics simulation (elecsim) , PMT Reconstruction (cal), Event Reconstruction (rec), Replication of output to destination sites
- ❖ All steps or part of them can be connected to each other with data to form a pipeline, chained and started through ProdSys
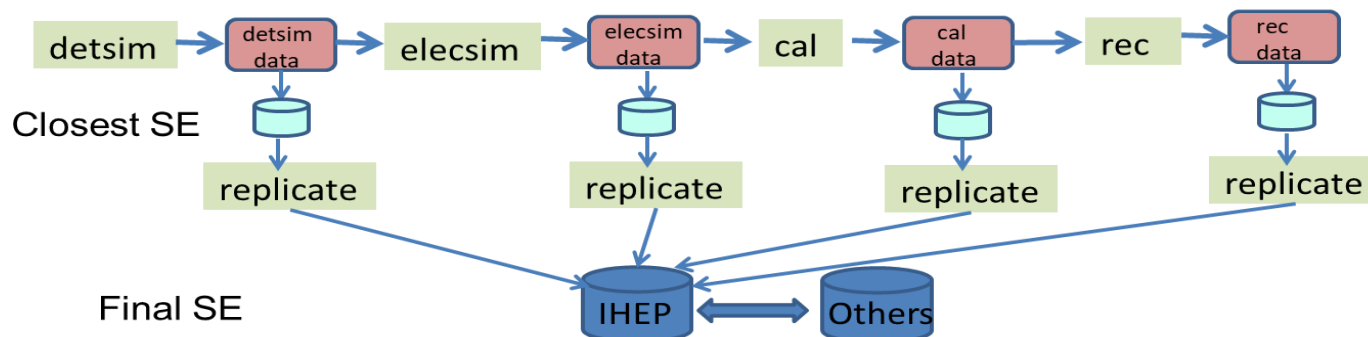
ProdSys 2/2

❖ **ProdSys is implemented based on DIRAC**

- **Frontend** Accept user requests and create JUNO workflow and dataflow

- **Transformation system (TS)** Transform JUNO workflow and dataflow into a pipeline

- **DIRAC File Catalogue (DFC)** Provide query of metadata and file status which is used to trigger the process

- Jobs and file transfers are submitted to **DIRAC WMS and DMS**

❖ Prodsys regularly used in JUNO MC simulation tasks

❖ Same mechanism will be used for JUNO 2nd reconstruction

- **Real data -> Cal -> Rec -> Destination SEs**
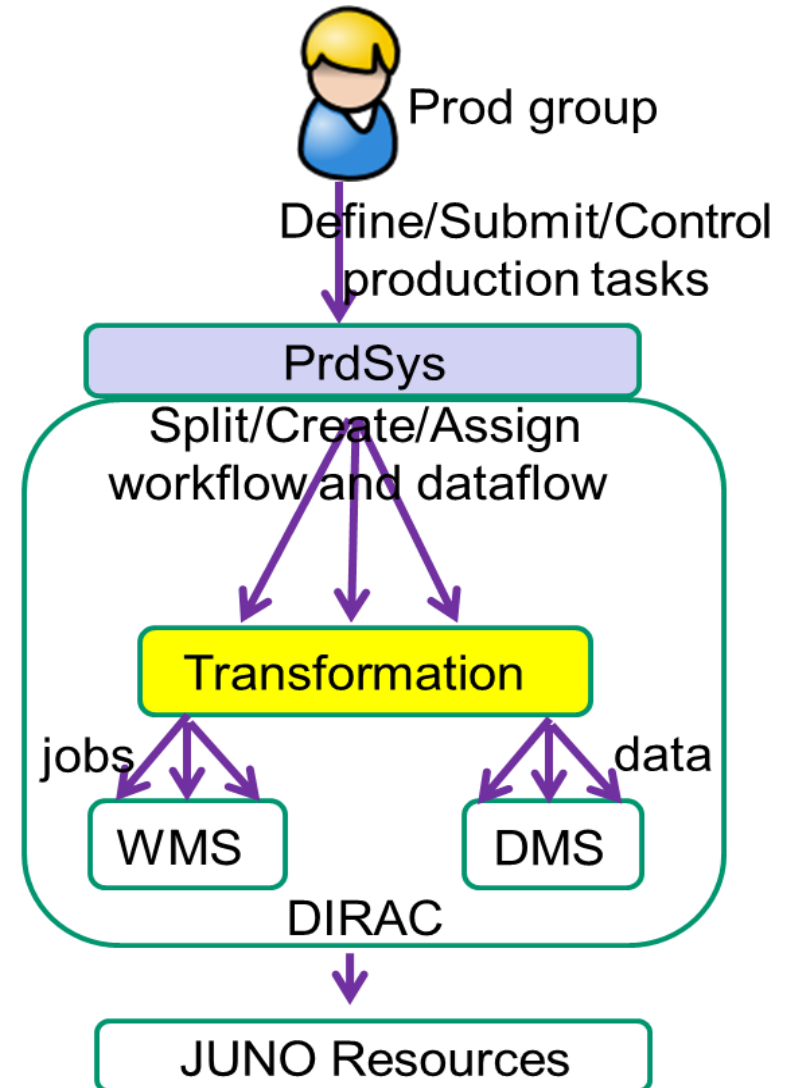
JSUB

- ❖ JSUB – a lightweight user job submission tool, developed in python
  - Ease process of physics analysis and small number of simulation for JUNO users
  - Automatically take care of life cycle of **user analysis** in grid env
- ❖ Main common function packages in JSUB
  - Job splitting and submitting, Job management, dataset operation, backend, UI
  - User Steering file is written in YAML
- ❖ Main features
  - Extensible with multi-experiments and multi-backends
  - Support fast submission with DIRAC parameter job submission feature
  - Support flexible splitters with multi parameters to split tasks into subjobs
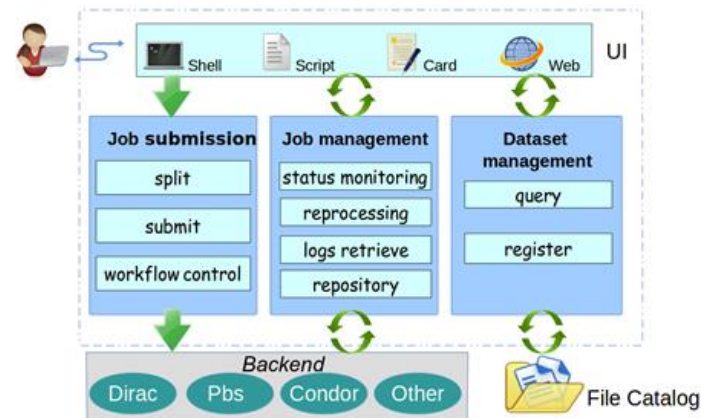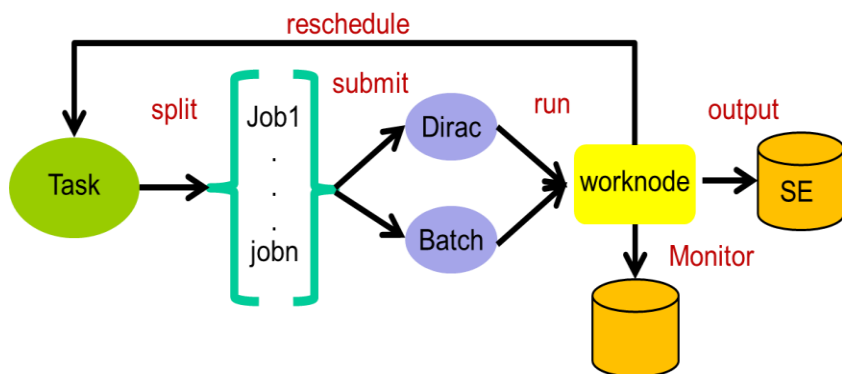
Data management

❖ **User Interface** with interactive console and commands

- Provide a global data view

- Create and manage dataset

- Submit and manage transfer requests

❖ **DIRAC Data Management System (DIRAC DMS)**

- DFC: metadata and replicas catalogue

- Request Management System and Transformation System: split dataset into file transfers and arrange in queue

- Interface to available file transfer tools

❖ **Rucio is under evaluation**

❖ **Data Movement Service – FTS**

- Take care of file transfers

8
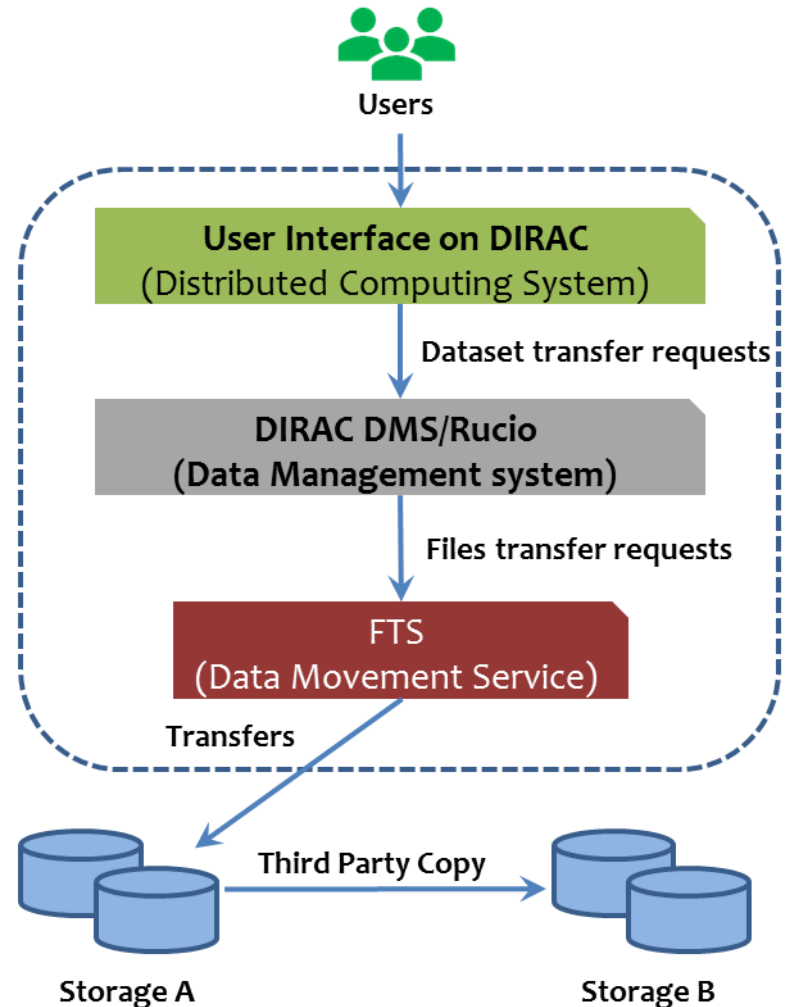
Raw data transfer system

❖ Aim to take care of raw data distribution to data centers

- Receive data information from Message Queue to trigger the whole process

- Register data in DFC

- Replicate data to data center and register in DFC

- Archive in tape and register in DFC

- Validate data and monitor status

❖ It consists of four modules, implemented based on DIRAC DMS

- Data Receiver, File Register, Block Creator, Block Transfer

- Transfers and validation are based on blocks which are grouped by data receiving date
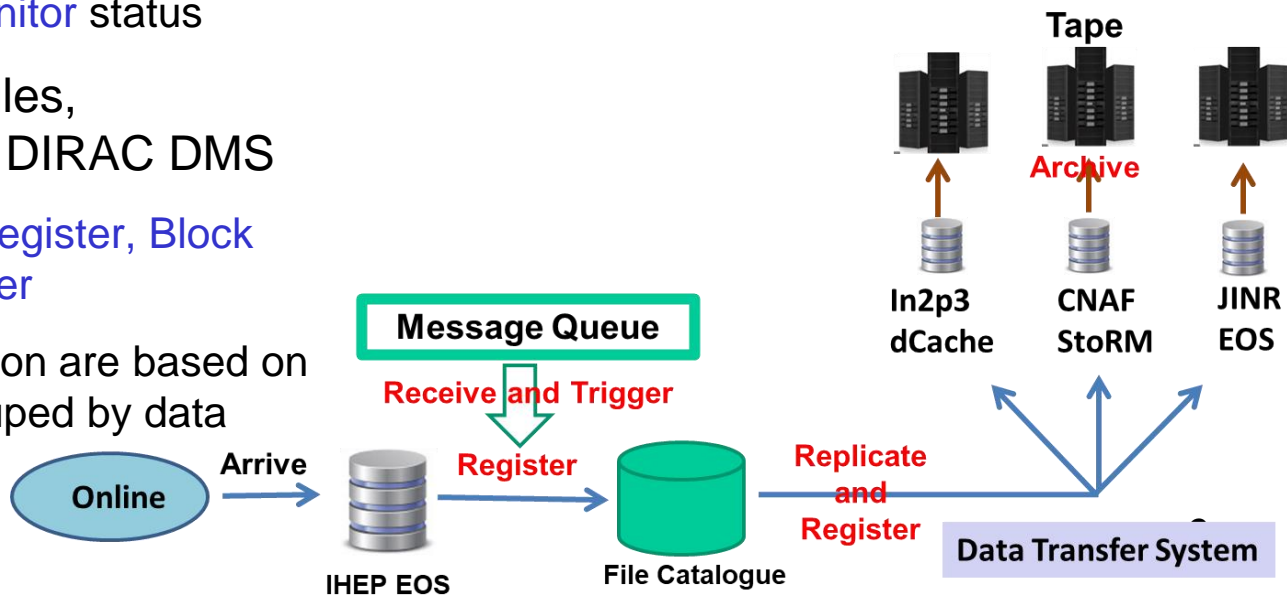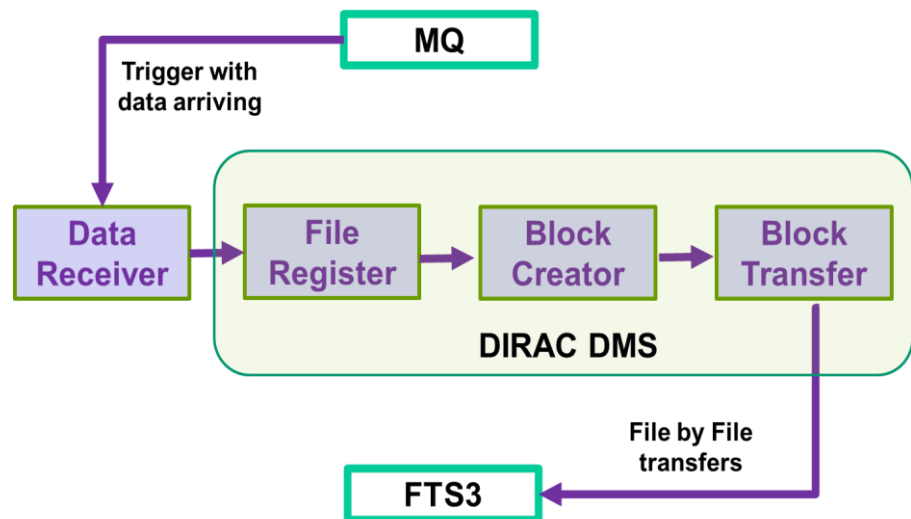
Offline Condition DB access

❖ JUNO uses MySQL to store condition data

❖ Frontier/Squid infrastructure has been set up in grid env

  ● Help avoid high load in central DB and speed up access to condition data

❖ Frontier server was deployed in IHEP and JINR, connected to DB

❖ Tests done with jobs has proved system is functioning

  ● Show > 10 times better with cache access than direct DB access

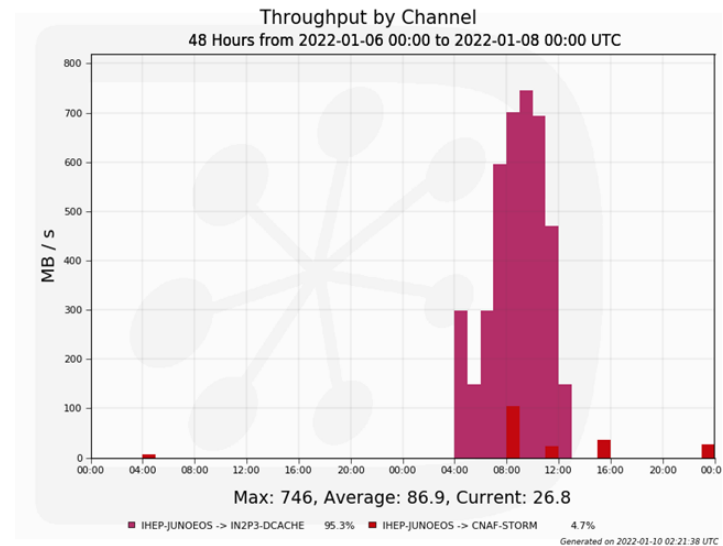❖ Tests to simulate production-like environment will be done this year
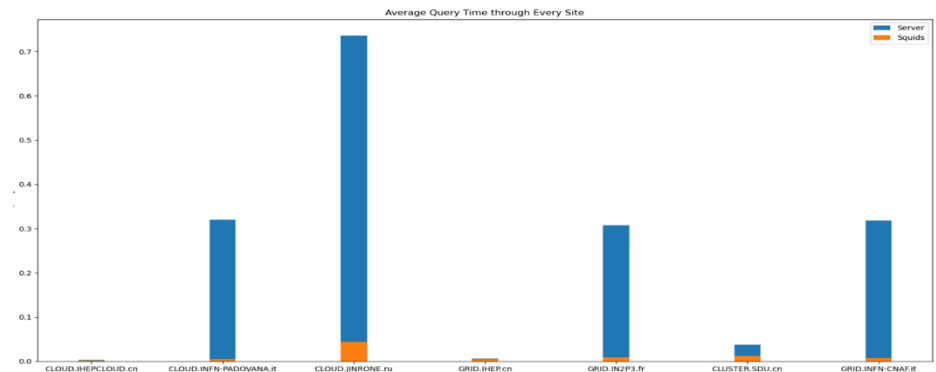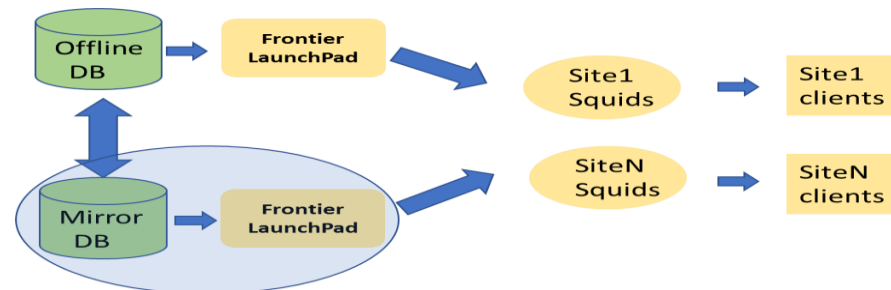
HTTP TPC

❖ **All SEs has completed migration from Gridftp to HTTP TPC**

❖ **TPC daily monitoring** has been developed using ES+kibana

❖ Pressure tests has been done with DIRAC DMS+FTS

- Maximum speed can reach limitation of network bandwidth (IHEP EOS -> IN2P3 dCache)

- Problem found on communications between StoRM and EOS, updating to EOS5 to solve

❖ More pressure tests are planned this year

Details for https://junoeos01.ihep.ac.cn → https://ccdcacli303.in2p3.fr

Throughput by Channel
48 Hours from 2022-01-06 00:00 to 2022-01-08 00:00 UTC

Max: 746, Average: 86.9, Current: 26.8

IHEP-JUNOEOS -> IN2P3-DCACHE 95.3%   IHEP-JUNOEOS -> CNAF-STORM   4.7%

Generated on 2022-01-10 02:21:38 UTC
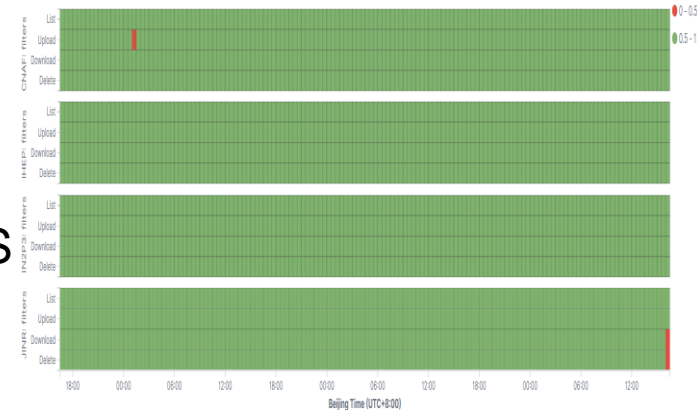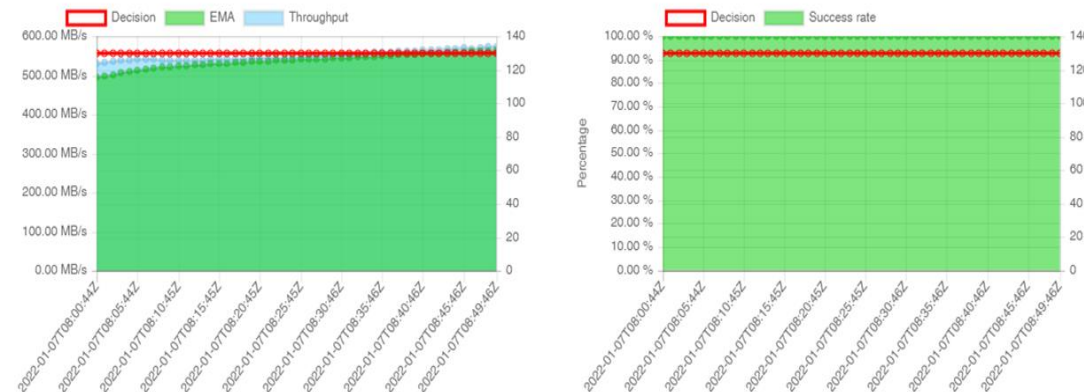
Token-based AAI

❖ Migration from **X509-based AAI to token-based AAI** is on-going for JUNO DCI

❖ Status:

- **IAM service has been set up**: https://iam-juno.cloud.cnaf.infn.it/login

- **IHEP SSO is connected**

- **Connections to eduGAIN are in progress**
  - CNAF and IHEP are working
  - IN2P3 and JINR in testing

- Some site CEs and SEs already supported token

- More to do:
  - Push IAM service to be in production in parallel with VOMS
  - Migrate DIRAC to the version supporting both certificate and token
  - Complete CEs and SEs support of token

Welcome to **juno**

Sign in with your juno credentials

👤 Username

🔒 Password

Sign in

Forgot your password?

Or sign in with

eduGAIN

INFN

中国科学院高能物理研究所
Institute of High Energy Physics
Chinese Academy of Sciences

System in production 1/2

- ❖ ~5M Jobs are submitted and run with ProdSys since beginning of 2022
  - ● ~5.18MHS06 Normalized CPU time
- ❖ DM has been used for massive file registration and file transfer
  - ● 1.4PB data transferring
  - ● In DFC, ~2PB data and 16M files registered and visible to users

Cumulative Jobs by Site

69 Weeks from Week 52 of 2021 to Week 17 of 2023

Max: 4.99, Min: 0.01, Average: 1.78, Current: 4.99

■ GRID.JINR-CONDOR.ru	2.1	□ CLUSTER.CSNS.cn	0.1	□ CLUSTER.SDU.cn	0.0
■ GRID.IHEP.cn	1.5	□ CLOUD.INFN-PADOVANA.it	0.1	■ GRID.JINR.ru	0.0
■ GRID.INFN-CNAF.it	0.7	■ CLOUD.IHEPCLOUD.cn	0.0	■ ANY	0.0
■ GRID.IN2P3.fr	0.5	□ GRID.MANCHESTER.uk	0.0		

Generated on 2023-04-30 08:20:01 UTC

Transfered data by Channel

56 Weeks from Week 52 of 2021 to Week 04 of 2023

Max: 1.41, Min: 0.04, Average: 0.98, Current: 1.41

■ GRID.JINR-CONDOR.ru -> JINR-EOS	0.3	■ IHEP-JUNOEOS -> IN2P3-DCACHE	0.0
■ JINR-EOS -> IHEP-JUNOEOS	0.1	■ DIRAC.Client.cn -> IN2P3-DCACHE	0.0
■ ihep.ac.cn -> IHEP-JUNOEOS	0.1	■ DIRAC.Client.cn -> IHEP-STORM	0.0
■ GRID.IHEP.cn -> IHEP-JUNOEOS	0.1	■ DIRAC.Client.local -> IHEP-JUNOEOS	0.0
■ DIRAC.Client.cn -> IHEP-JUNOEOS	0.1	■ DIRAC.Client.cn -> CNAF-HTTPSTORM	0.0
■ IHEP-JUNOEOS -> JINR-EOS	0.1	■ CLUSTER.CSNS.cn -> IHEP-JUNOEOS	0.0
■ GRID.INFN-CNAF.it -> CNAF-STORM	0.1	■ IN2P3-DCACHE -> IHEP-JUNOEOS	0.0
■ GRID.IN2P3.fr -> IN2P3-DCACHE	0.1	■ CNAF-STORM -> IHEP-JUNOEOS	0.0
■ IHEP-JUNOEOS -> CNAF-HTTPSTORM	0.1	... plus 180 more	

Generated on 2023-01-30 06:38:58 UTC

System in production 2/2

❖ Raw data transfer system has started testing with commissioning data

❖ Plan to use FTS Monitoring dashboards for transfer

  ● Infrastructure is set up
    ▪ FTS->ActiveMQ->logstash->ES->Kibana/Grafara

❖ DIRAC Accounting will be used for history view
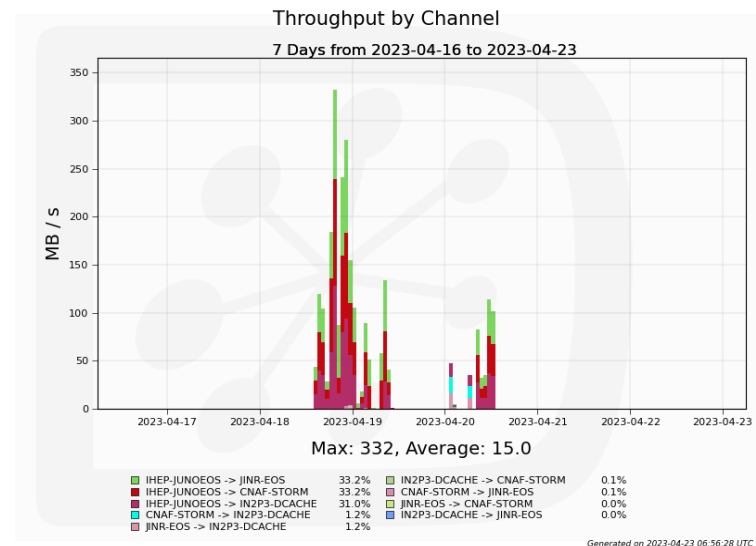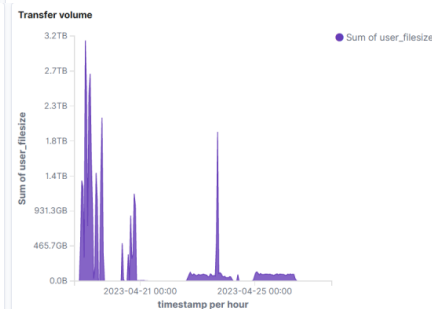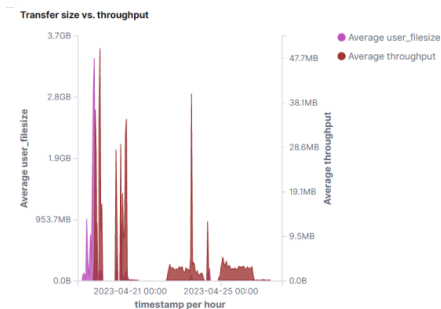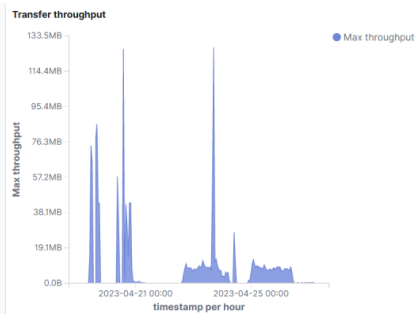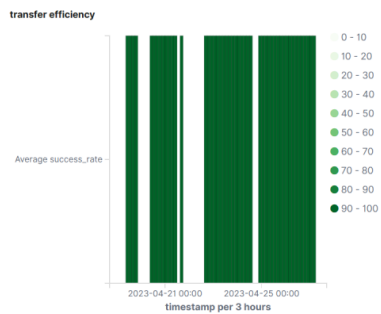
Summary and Plan

❖ As a medium-size experiment, JUNO has successfully set up distributed computing system using existing WLCG middleware with quite limited manpower

❖ The system meets the requirements of JUNO computing model, successfully used in some of JUNO production activities

❖ Pressure testing on production-like environment will be carried out soon to prepare for data-taking in 2024

❖ Thank you!