JUNO DCI: a status update

Giuseppe Andronico

INFN Sezione Catania

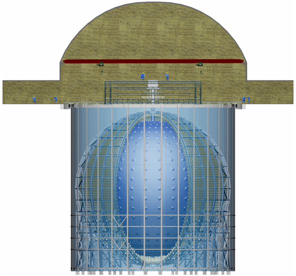
23rd May 2025

Contents

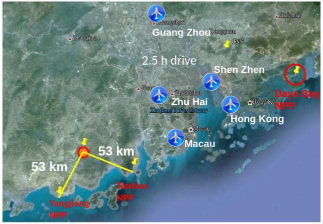
JUNO

- **Jiangmen Neutrino Underground Observatory**
- Neutrino physics experiment based on a 20kTon liquid scintillator detector in south China, near Jiangmen
- Several objectives; some are:
  - study mass hierarchy
  - measure oscillation parameters with unpaired precision (energy resolution $\frac{3\%}{\sqrt{E(MeV}}$)
  - Burst supernova neutrino events: $\sim$ 5000 events @10 kpc
  - Diffused supernova neutrinos if $\langle E \rangle > 15$ MeV
  - Geoneutrinos: determining U/Th ratio in earth
  - solar neutrino
  - atmospheric neutrino
  - proton decays
- **Detector is in commissioning phase: water filling already done, LS filling in progress**
- Data taking is running with commissioning data. Full data taking foreseen August this year
- Large collaboration: 72 institutions from 17 countries

JUNO CD



Location

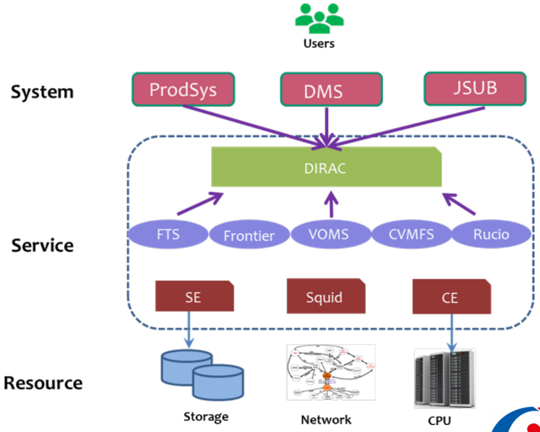


JUNO Computing and Data Requirements

- Calibration data are now merged in RAW data stream
- Data Volume:
  - Initial estimated RAW data: 2 PB/year (approx. 60 MB/s)
  - Updated estimates based on realistic implementation: $\sim$ 2.3 PB (77.2 MB/s)
  - Introduced RTRAW files (pre processed RAW files, in ROOT format, not reproducible), one to one with RAW files, about same size
- Reconstructed data $\sim$ 1.0 PB/year
- Total data (RAW+RTRAW+others) $\sim$ 6 PB/year
- Estimated Computing Power: $\sim$ 155 kHS23 to reconstruct 1 year of data in 1 year
- Requirement: manage expected data volume and processing needs across geographically distributed data centres

JUNO DCI Architecture - A Tiered Model

- The JUNO DCI is based on a tiered architecture.
- This organizational model covers offline activities across JUNO data centers to provide data access.
- It has been implemented to satisfy the computing model.
- The infrastructure is designed not only to aggregate resources but also to respond to a specific computing model.



Tier0 - The Central Hub (IHEP)

- IHEP (Institute of High Energy Physics) in Beijing, China, serves as the Tier0 center.
- Main Responsibilities:
  - Receiving and storing all RAW data directly from the JUNO detector site in Kaiping.
  - Holds copies of all other data types (MC, reconstructed, calibration).
  - Performs:
    - data quality monitoring, calibration analysis, conversion to RTRAW
    - the initial full data reconstruction and calibration.
  - Also supports simulation and user analysis.

Tier1 and Tier2 - International and Opportunistic Sites

- Tier1:
  - Some international data centers in Europe:
    - CC-IN2P3 (France).
    - INFN-CNAF (Italy).
    - JINR (Russia).
  - Store complete or partial replicas of JUNO data.
  - IN2P3 specifically holds 1/3 of the RTRAW data.
  - Main activities: re-reconstruction, simulation, user analysis.

- Tier2: Smaller computing sites, including cloud and cluster resources.
  - Provide supplementary CPU power.
  - Used primarily for simulation and user analysis tasks.
  - Some Tier2 sites may not have Storage Elements (SE).
  - a couple of candidate sites, not yet confirmed

JUNO DCI Resources

Present	Disk (PB)	Tape (PB)	Comp (kHS23)
IHEP	7.21	2.27	140
CC-IN2P3	0.24	3.38	15
INFN-CNAF	6.00	4.00	20
JINR	2.54	5.05	48

MoU 2025	Disk (PB)	Tape (PB)	Comp (kHS23)
IHEP	5.0	5.0	100
CC-IN2P3	0.2	3.0	15
INFN-CNAF	6.0	4.0	20
JINR	5.0	5.0	30

Present %	Disk	Tape	Comp
IHEP	45	15	63
CC-IN2P3	2	23	7
INFN-CNAF	38	27	9
JINR	16	34	22

MoU 2025 %	Disk	Tape	Comp
IHEP	31	29	61
CC-IN2P3	1	18	9
INFN-CNAF	37	24	12
JINR	31	29	18

IHEP: further 3PB disk storage on line end of June
JINR: further 2.5 PB disk storage on line in August

The Core Framework: DIRAC

- The JUNO DCI has been designed and built based on DIRAC.
- DIRAC serves as the core framework for DCI services.
- Integrates heterogeneous and remote resources, providing a unified platform.
- Offers fundamental services:
  - Workload Management (WMS): Job management (submission, scheduling, execution). Supports pilot-based scheduling strategies.
  - Data Management (DMS): Data management (placement, access).
- Core components maintained at IHEP

Data Management and Storage

- The DIRAC Data Management System is the core of DM.
- DIRAC File Catalogue (DFC): Central metadata and replica catalogue for all JUNO data. Provides a global view of data locations. All JUNO data are registered in the DFC. Also used for additional metadata to facilitate file localization.
- Storage Elements (SE): Data is stored on disk at participating data centers. Various SE technologies are used :
  - dCache (CC-IN2P3, JINR).
  - EOS (IHEP, JINR).
  - StoRM (CNAF).
- Tape Systems: A special type of SE for long-term preservation. Used for infrequently accessed files. Present in all sites.

Workload Management (WMS)

- Responsible for job submission, scheduling, and management.
- DIRAC provides the WMS framework.
- Capable of accepting and scheduling jobs on suitable resources, hiding complexity.
- Supports the management of massive tasks, each potentially with thousands of jobs.
- Supports single-core and multi-core jobs consistently.

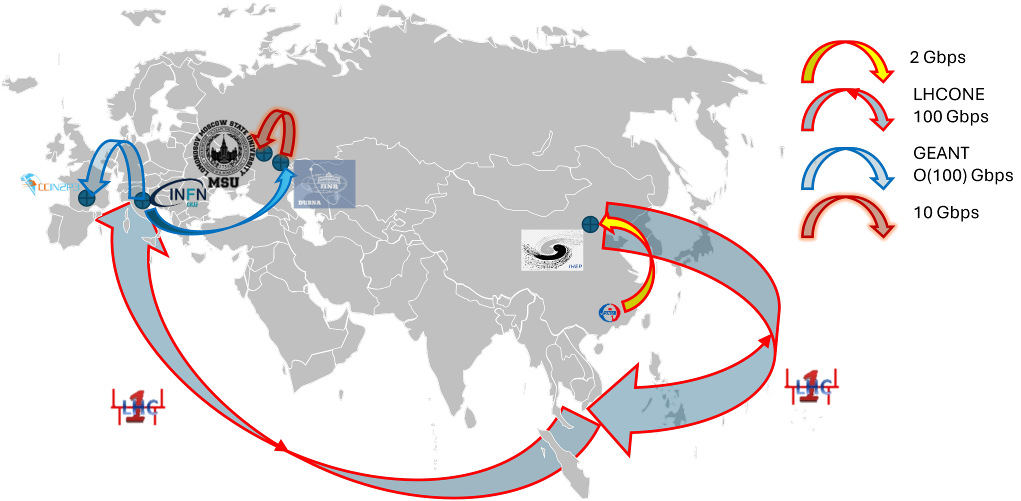
Authentication and Authorization (AAI)

- From beginning based on Virtual Organizations (VOs) and VOMS.
- JUNO is working to move towards token-based authentication.
- Both storage and computing resources can be contacted with tokens, but DIRAC does not yet fully support them.
- Created an INDIGO IAM instance, hosted at CNAF.
- VOMS has been enabled on IAM as a gradual transition step towards tokens.
- Deployed on INFN cloud a voms-importer to keep IAM-VOMS in sync with standard VOMS, until the swap.

Network Infrastructure

- A good network connecting all data centers is a prerequisite.
- JUNO utilizes:
  - National Research and Education Networks (NRENs).
  - connection between China Accademy of Science and GEANT at 100 Gb/s since September 2023
  - International networks such as LHCONE and GEANT.
- All JUNO data centers are connected to LHCONE, and JUNO has been recognized as an LHCONE community.
- Network tests across key sites (IHEP, CC-IN2P3, CNAF, JINR) confirmed high throughput and full NIC utilisation.
- The IHEP–JINR link remains suboptimal; data transfers are temporarily performed via CNAF.

JUNO International Network



Monitoring and Operations

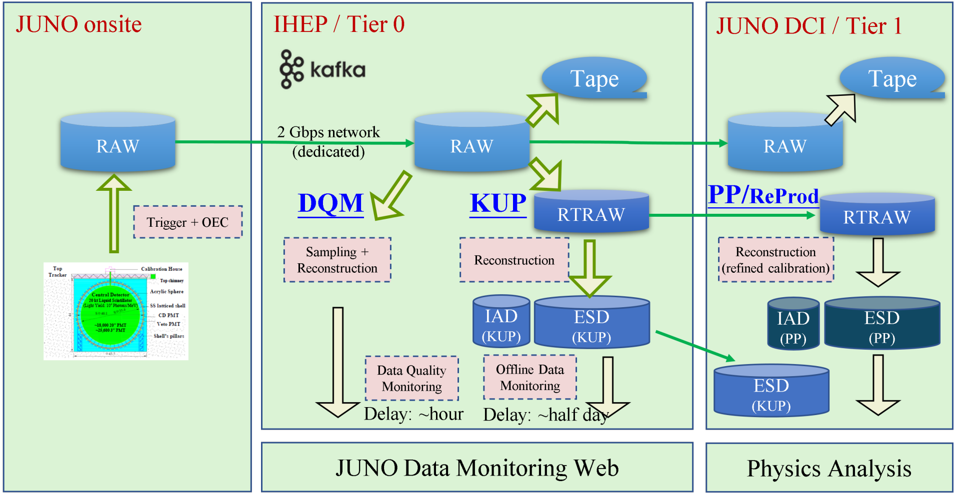
- A comprehensive monitoring system is used to track the status of the DCI.
- Includes Grafana-based dashboards.
- Monitors:
  - IHEP T0 status
  - DCI sites status
  - network performance (using perfSONAR)
  - data transfers (FTS monitoring)
  - job execution.
- Communication of issues and operations: agreed to use the European Grid Initiative (EGI) helpdesk platform.
- Accounting is managed by the Computing and Steering Group, based on internal monitoring, with the idea of comparing with EGI tools.

RAW Data Flow and Processing Pipeline

1. RAW Data Acquisition: RAW data from the JUNO detector is initially transferred to the IHEP data center. Transfer managed by SPADE.
2. Processing RAW data at IHEP: first step is Data Quality Monitoring (DQM). In parallel, RAW data are used to generate RTRAW (Reconstructed Raw Data), ESD (Event Summary Data), in future also IAD (Indexed Analysis Data).
3. Storage and Distribution:
    - RAW data are registered on DCI and moved directly to tape for long-term preservation, both locally and on some DCI sites
    - RTRAW are registered on DCI and copied to disk storage (locally and on some DCI sites) for a limited period (max 2 years), then moved to tape.
    - ESD are copied to disk storage, locally and on DCI sites.
    - The Data Transfer System manages the replication of RAW, RTRAW, and ESD.
4. Distributed Processing: Re-reconstruction (or PP, Physical Production), simulation, and user analysis are performed mainly on DCI.
5. Data resulting from activities are copied and stored within the DCI .

Data Flow Schema

Schema



Automated System for RAW Data

- KUP, an automatic system based on Kafka messaging, has been developed to automate the process.
- Receives messages from SPADE when data arrive on the IHEP EOS disk.
  - DQM is notified of the arrival of new files.
  - Also, the process to convert RAW in RTRAW is notified, and then to have the first reconstruction producing ESD and IAD.
  - Registers RAW, RTRAW, ESD, IAD data in the DFC.
  - Replicates data from IHEP EOS to data centres disks and tapes in parallel.
- Independent systems for RAW, RTRAW and ESD.

Condition Data Access

- Condition data (event rates, detector status, etc.) are stored in a central MySQL database at IHEP.
- The FroNTier/Squid infrastructure is used for efficient access.
- Provides caching of condition data at local sites, reducing the load on the central DB.
- Accelerates data access for remote jobs.
- Tests have shown performance improvements of over 10 times compared to direct access.

JUNO-Specific Tools

- Several specialized tools developed for JUNO needs:
  - ProdSys: Production system based on the DIRAC Transformation System (TS). Manages large-scale MC simulation and data production workflows automatically and data-driven. Chain processing steps (DetSim, EleSim, PmtRec, EvtRec) based on data availability in the DFC.
  - JSUB: Lightweight and user-friendly Python tool for submitting user analysis jobs. Simplifies the process for users in the grid environment. Can split tasks into sub-jobs and submit thousands of jobs to DIRAC. Link: https://github.com/jsubpy
  - Raw Data Transfer System: Automatic system (see Slide 18).

Validation and Production Campaigns

- Several data challenges have been conducted to evaluate DCI functionality and performance.
- Tested data replication, network capacity, tape system functionality, service performance.
- Official MC production campaigns have been run on the system since 2020.
- Processed petabytes of MC data, transferred successfully.
- Example: First data production campaign (Feb 2020) generated >16TB of data, with 100% success for detsim workflows.
- A data challenge in 2020 evaluated DM and network with 20TB of data, simulating RAW data transfer from IHEP to Tier1, with the network available at the time, limited to 10 Gb/s. Average speeds between 1Gb/s and 3Gb/s, total 8Gb/s, nearly reached the full network potential available.
- More than 2 million jobs executed in the 2022 Mock Data Challenge (MDC) using ProdSys.

Summary and Next Steps

- The JUNO DCI is a well-defined system, based on consolidated WLCG middleware and a tiered architecture.
- Designed to meet the significant data and computing requirements of the JUNO experiment.
- Tests and validation are ongoing, including large-scale and pressure tests.
- Efforts are underway for multi-core support and the transition to token-based AAI.
- The data management system and reconstruction pipelines have been tested and are currently used in commissioning.
- Commissioning is also used to fix some final point in view of the full data taking start, foreseen in August this year.
- Actively working on metadata.

Thank you