

# JUNO computing

Giuseppe Andronico



# Activities

- Computing model; defining JUNO:
  - computing activities
  - Requirements and needed resources
- Distributed infrastructure
  - Data centres participating
  - Resources shared
  - Services and operations



Computing model

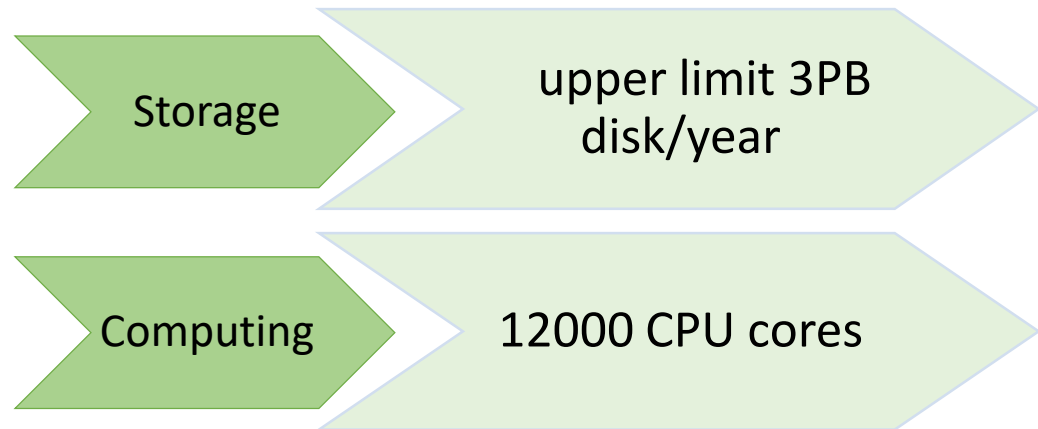


# Data volume

## Data type and volume

Data from elec	Order of 1000PB/year
Raw data	upper limit 2PB/year
Rec data	order of 20TB/year
Sim data	order of 100TB/year
Calibration data	order of 200TB + 2 TB/year
Ana data	order of TB/year

## Total resources requirements



**Base to plan the minimal amount of storage for JUNO experiment**



# Data storage

- Data copies:
  - main copy at IHEP
  - Backup in Europe
- Storage
  - Based on grid storage
  - User access by means of VOMS and JUNO VO
  - With book keeping and file catalog

# Reconstruction and analysis



Work in progress

Reconstruction indicative estimate (to be verified):

- Event rate 1kHz
- If
  - 1 sec of computing per event
  - 5000 CPU cores
- Then
  - 73 full days for SW version
  - Required 146 days/year (2 SW versions)

Analysis workflow is yet to be clearly defined



# Simulation

Some reference number found in DocDB  
Simulation times

- without optical photons propagation requires computing time  $\leq 10\text{s/event}$
- With optical photons propagation requires computing time till (in case of muons)  $O(\text{days})$

Producing a useful simulations library using 5000 CPU cores should requires 292 days (to be verified)



Distributed  
infrastructure



# Distributed computing components



## Sites layout

- How much data centers in the distributed infrastructure
- Their network connection

## Work flows

How tasks and jobs flow between data centers

## Data flows

Which and how data flow between data centers

## Data centers

- Capabilities
- Resources available

## Platform and Services

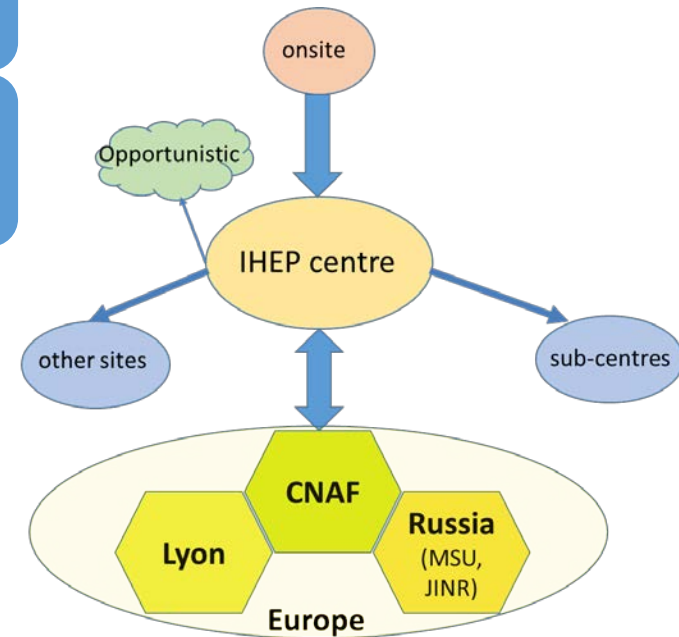
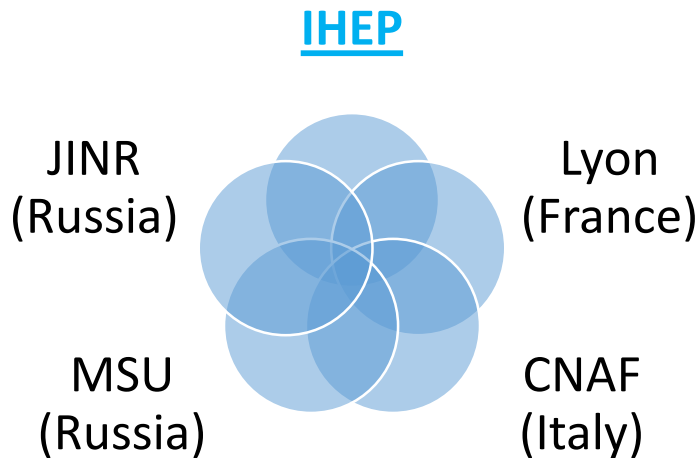
- Integrate resources in a distributed infrastructure
- Organize work and data flows
- Provide an unique interface to access to the resource



# Data centers

From EU data centers required a special status

Presently confirmed data centers:



# Resources



Centre	Actual share		HPC	GPU	Global	
	Storage	Core			Storage	Core
IHEP	500 TB	888	Yes	Yes	15 PB	18000
CC-IN2P3			Yes	Yes	25 PB	35000
CNAF	230 TB	250	Yes	Yes	38 PB	40000
JINR	100 TB	100	Yes	Yes	14.6 PB	10328
MSU	40 TB	32	No	No		

# Services

## User management

- Global:
  - Certification Authorities (CA)
  - Virtual Organization Management Service (VOMS)
  - JUNO-VO

## Data and Storage

- Global
  - DIRAC
  - File Catalog
  - Book Keeping
- Local: Storage Element (SE)

## Job management

- DIRAC



# User authentication

- Resources access using grid authentication mechanism
  - Certification authorities
    - making IHEP CA available for who do not own a CA
  - VOMS
- SSO to access web resources (DocDB, Wiki, ....)
  - Federating existing AAI from JUNO partners
  - Making available IHEP AAI for others



# Status of Distributed computing

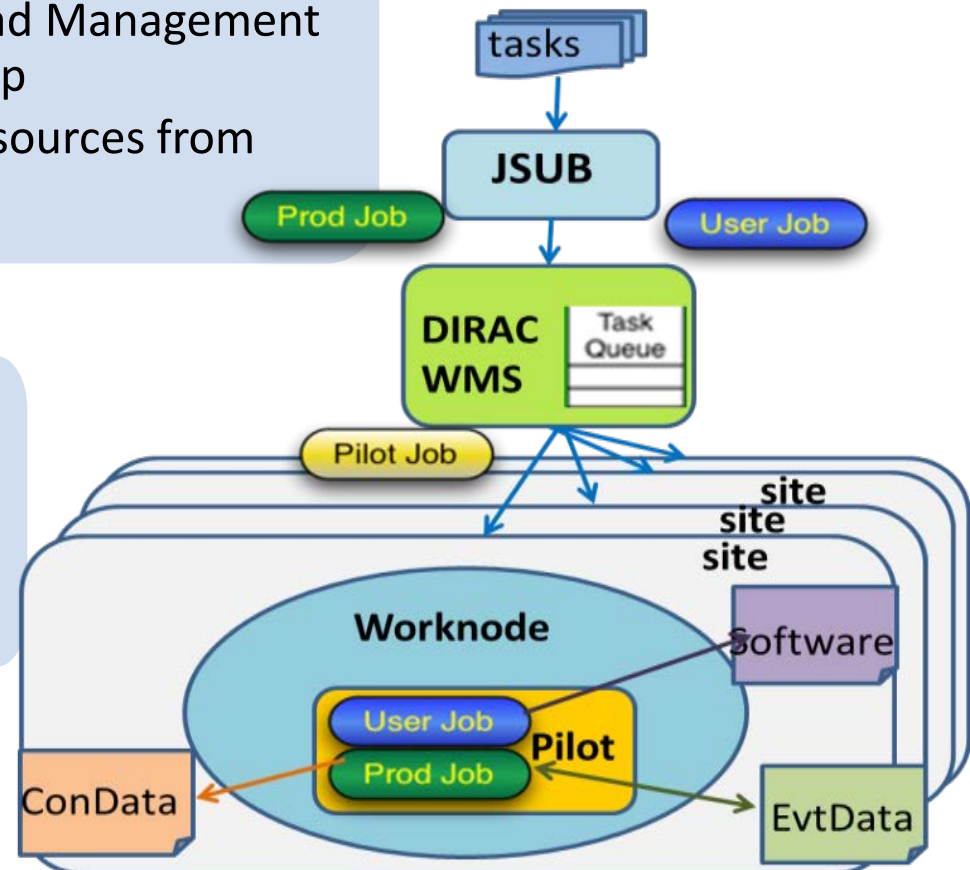
Prototype for the system is ready

## Job management

- Dirac-based Workload Management System already set up
- Ready to connect resources from other data centers

## Storage

StoRM-based central SE already set up to accept data from CNAF





# Data preservation

## Long preservation data

- How much time data are to be preserved
- Need to preserve environment
  - Software version
  - Needed libraries with correct versions
  - Operating System version
  - File types and specs

# Thank you



Any question?