

Computing for the KM3NeT research infrastructure

Mieke Bouwhuis

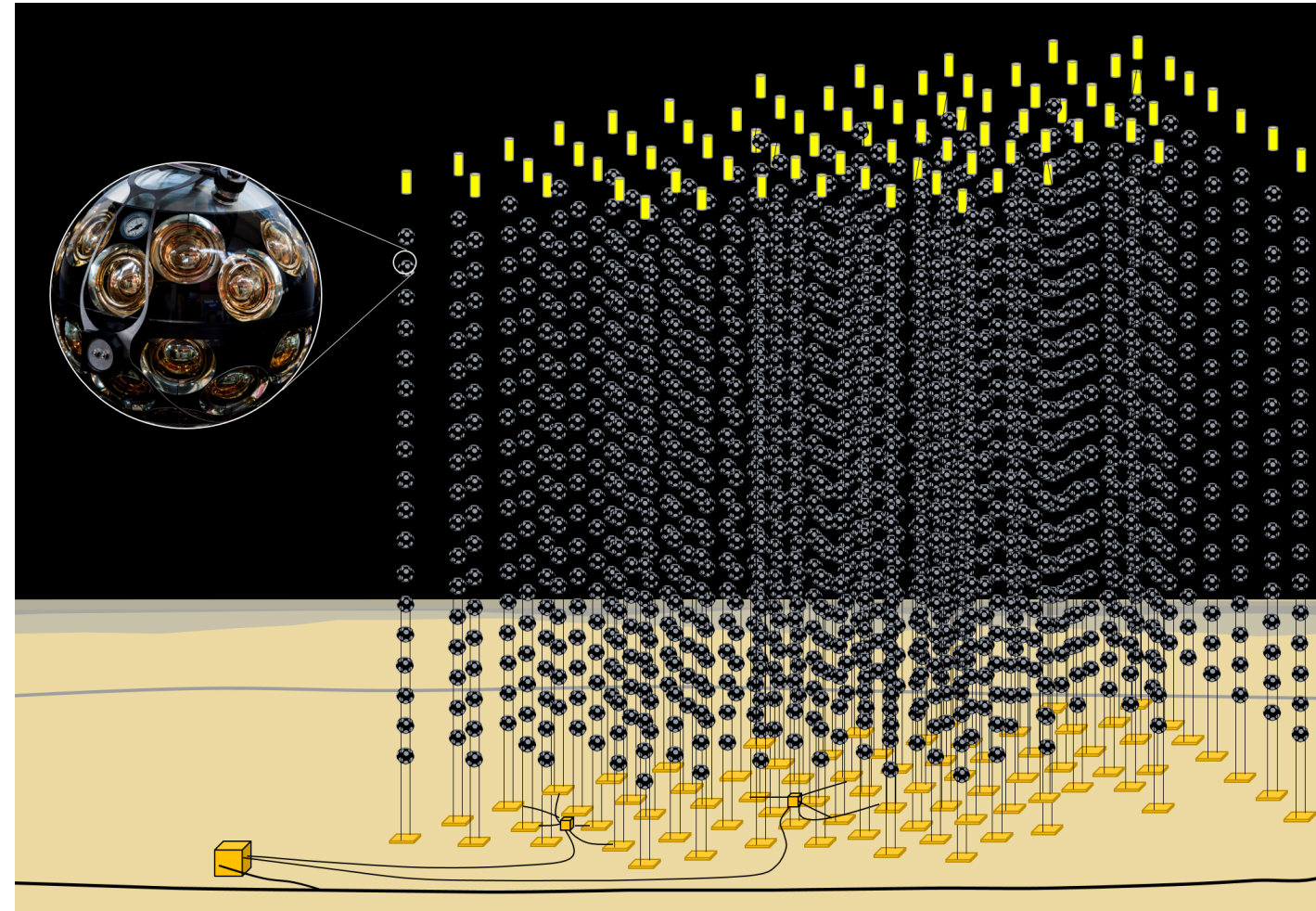
Nikhef (NL)

HEPiX Spring 2024 Workshop

The KM3NeT Research Infrastructure



- Distributed infrastructure in the Mediterranean Sea
 - FR: 10^7 m^3
 - IT: 1 km^3
- Modular telescope array
 - 345 vertical strings
 - 200.000 light-sensitive photo-multiplier tubes
- Today:
 - FR (ORCA): 15% operational
 - IT (ARCA): 12% operational



KM3NeT ORCA and ARCA

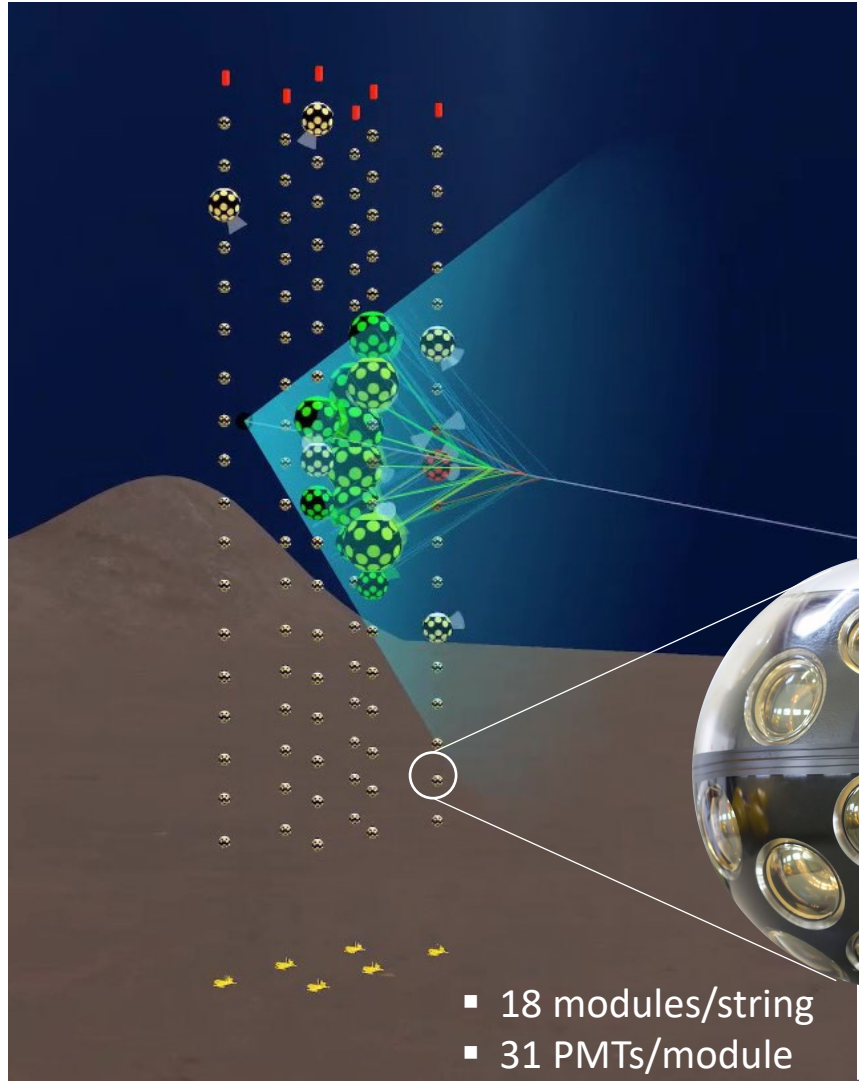


- ORCA (FR): for studying neutrino properties
- ARCA (IT): for studying astrophysical sources with neutrino astronomy
- Single collaboration of ~250 members
 - 45 partner institutes - 14 countries



- Same technology
- Shared computing and storage resources
- Same data processing software
- Common data formats



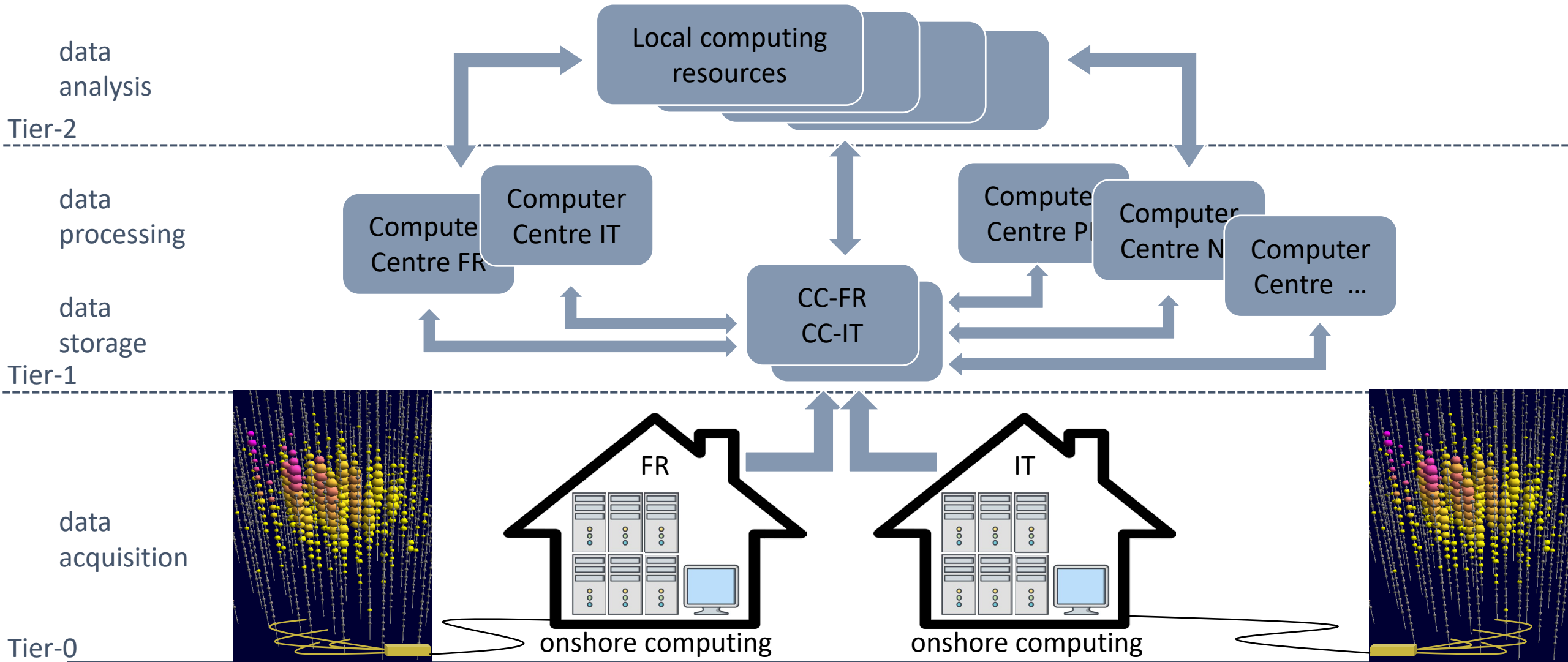


- 18 modules/string
- 31 PMTs/module

- Detection of Cherenkov light produced by relativistic particles
- Photo-multiplier tubes (PMT) measure the arrival time of the light at the quantum level
- PMT singles rate 10 kHz
- 1 string 300 Mbps
- Full infrastructure 100 Gbps (345 strings)

All data to shore

KM3NeT Computing Model

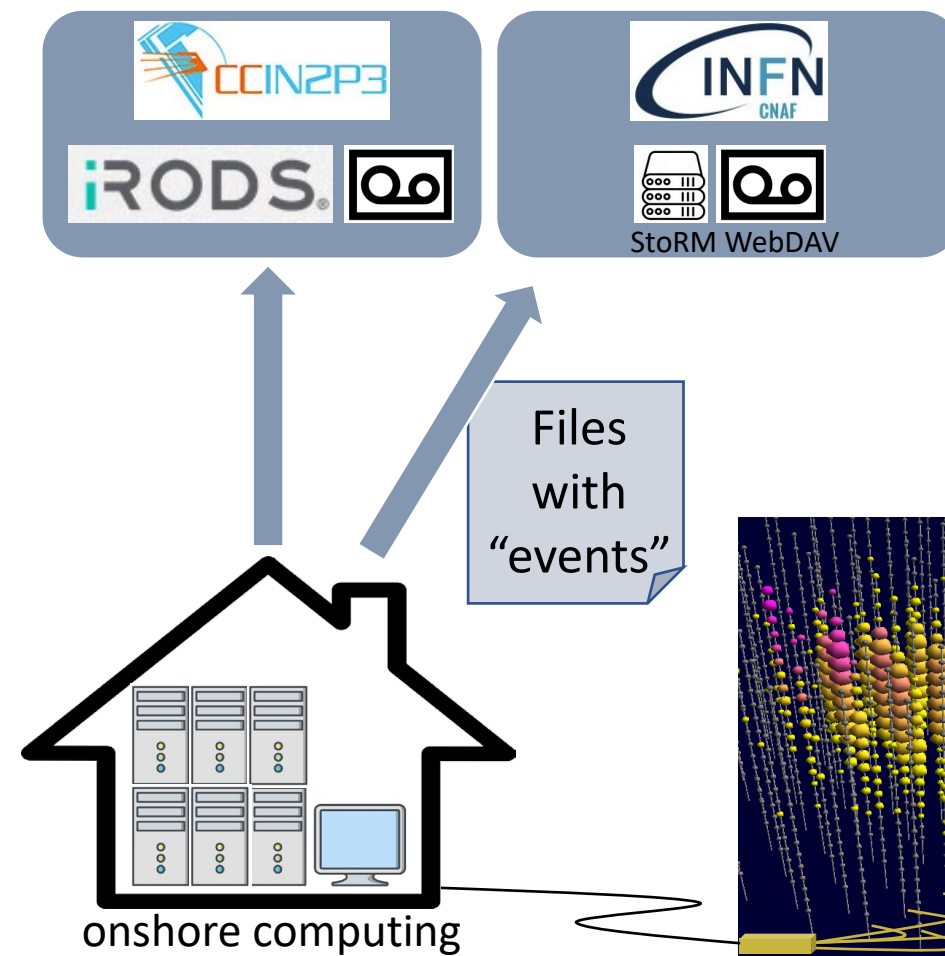


KM3NeT Computing at the Tier-0 level

Real time data filtering



- All data to shore
- Onshore computing: finding causalities between
 - Arrival time of the photons
 - Positions of the detection modules
- Filters out “events”
 - Event rate: 200 Hz
 - Reduction factor: 10^4
- Nightly copy of event files to CC-IN2P3 and CNAF



KM3NeT Computing at the Tier-1 level

Offline data processing

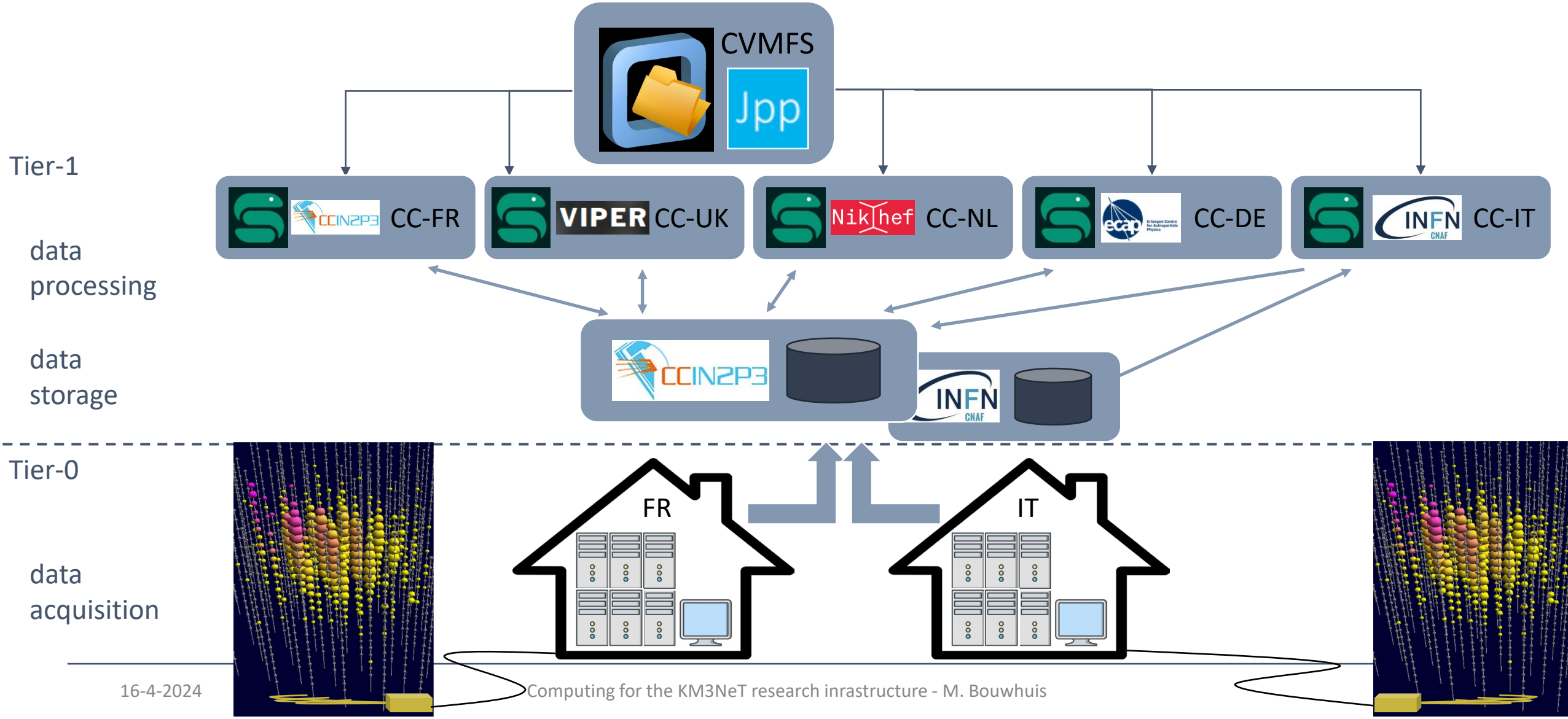


- Calibration of raw events
 - PMT quantum efficiency calibration
 - PMT gain calibration
 - Position calibration of detection modules
 - Orientation calibration of detection modules
 - PMT time calibration
- “Reconstruction” of calibrated events, determination of
 - neutrino direction
 - neutrino energy
 - neutrino type

Input to all analyses

KM3NeT Computing at the Tier-1 level

Current implementation

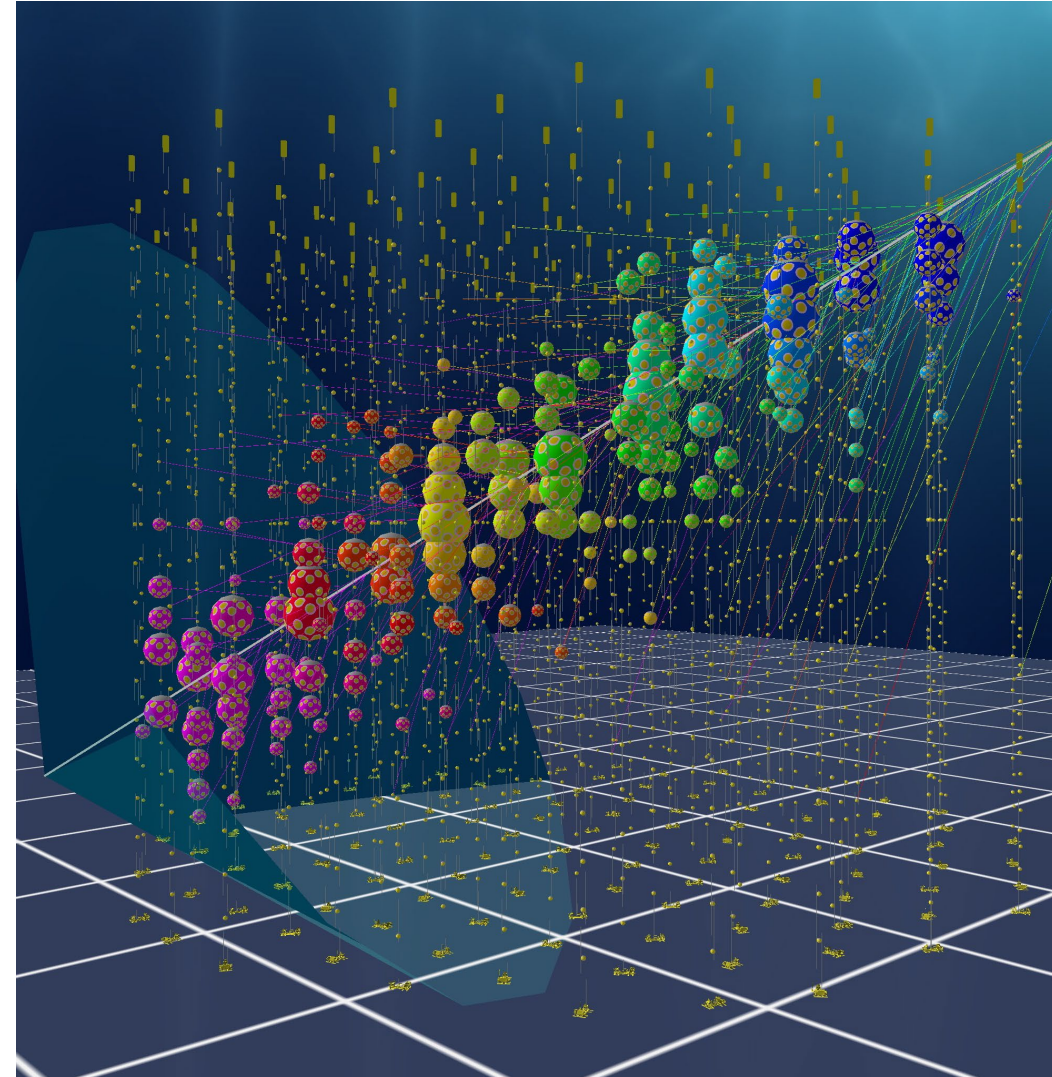


KM3NeT Simulations

Integrated with data processing



- Events are
 - Atmospheric muons 200 Hz
 - Atmospheric neutrinos $10\text{-}100 \text{ day}^{-1}$
 - Cosmic neutrinos $10\text{-}100 \text{ year}^{-1}$
- Determination of the detection efficiency, resolution and background
- Using the measured environmental conditions
- Simulation includes:
 - neutrino interaction / cosmic ray production, propagation of particles in the water, production of Cherenkov light, propagation of Cherenkov light in the water, detection of light by the PMTs
 - data filtering, calibration, reconstruction
- Simulations 10x data volume
- Integrated with the data processing





- Currently all data processing output is stored at CC-IN2P3
- Ideally no dependence on a single site
- Ideally the computing load and storage are distributed across multiple sites
- Grid resources available via different KM3NeT partners
- Comply with the principles of FAIR
 - Findable, accessible, interoperable and reusable

Towards a sustainable solution

For KM3NeT data management and data processing



- Currently deploying a prototype Rucio instance for data management of distributed data



- Using the CERN-hosted FTS service



- Testing workflow management on the KM3NeT workflows with the EGI-hosted Dirac instance for distributed data processing



- Proceeding to integrate the services with the Indigo IAM service hosted by CNAF

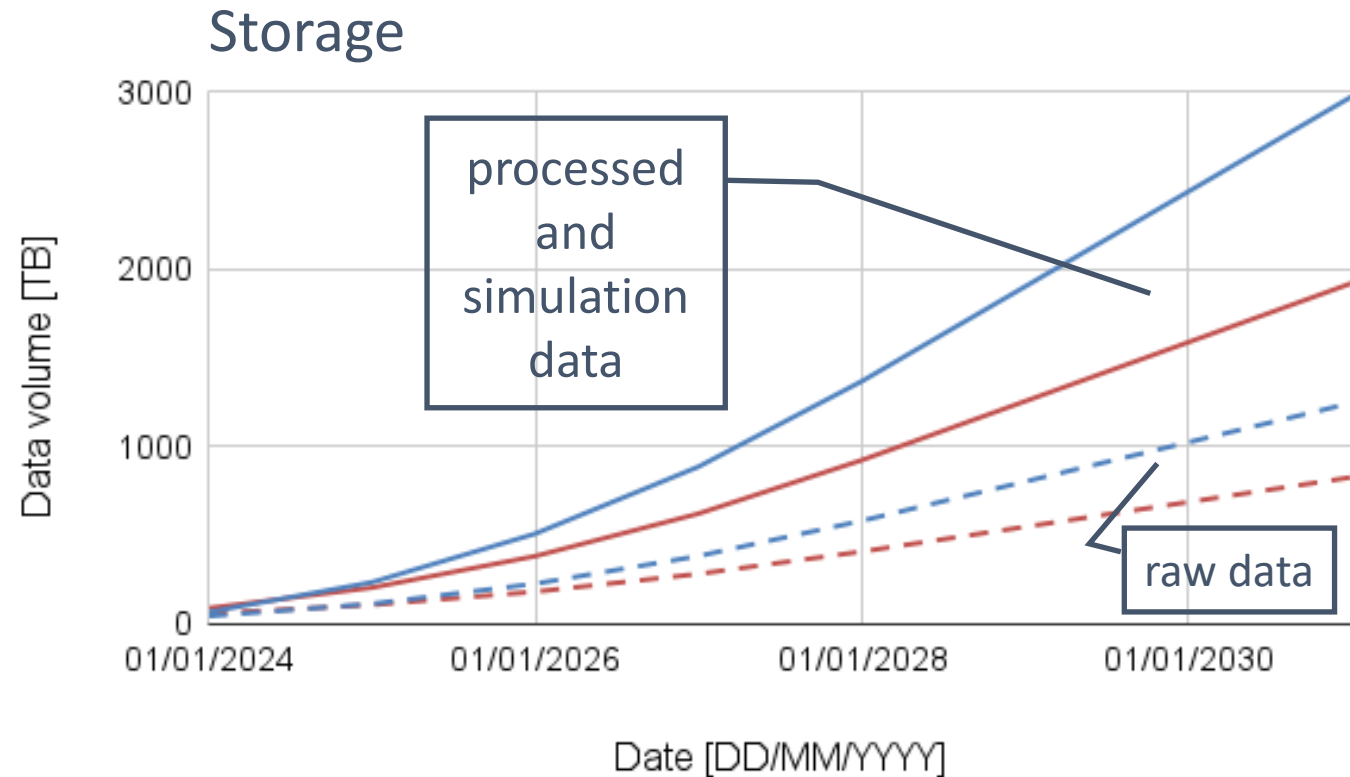
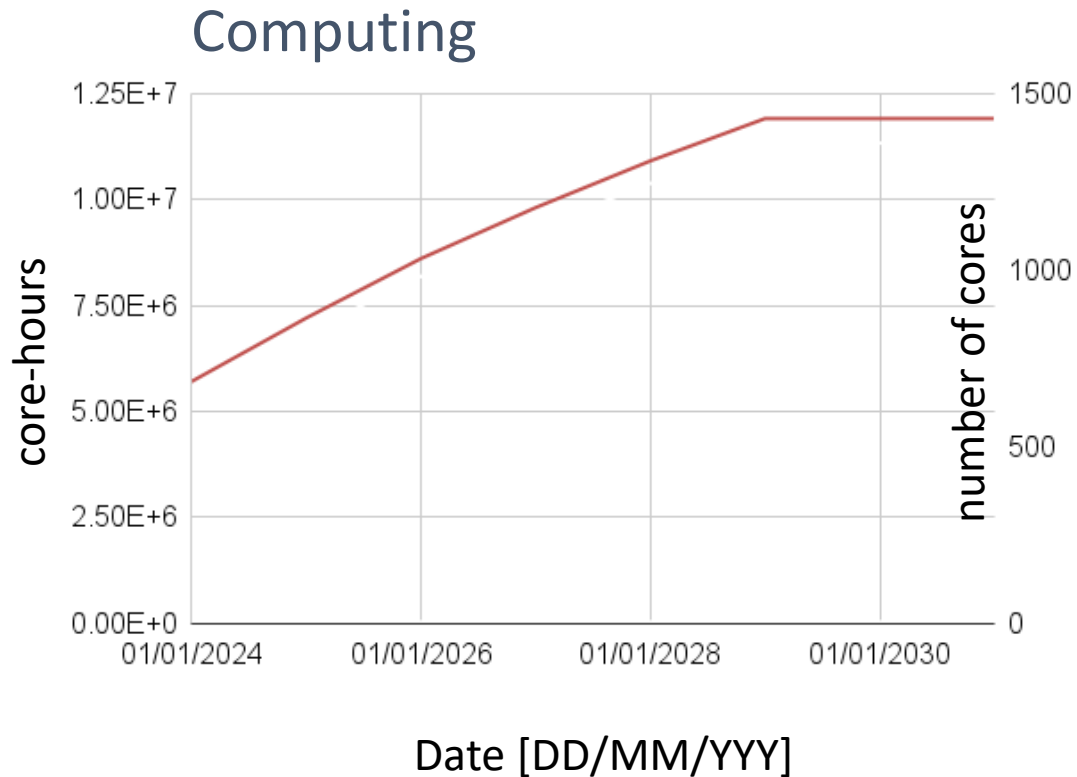


- Software distribution via CVMFS

- Fixed latency processing
 - multi-messenger astronomy
 - prevent data pileup

KM3NeT Tier-1 requirements

Towards the full infrastructure



- ARCA (IT)
- ORCA (FR)



- KM3NeT under construction
 - Data taking ongoing 24/7
 - Analyses progressing
- KM3NeT needs a data processing and data management solution for the next 15-20 years
- KM3NeT is adopting Grid solutions
- Full infrastructure: 500 TB/y and fixed latency data processing