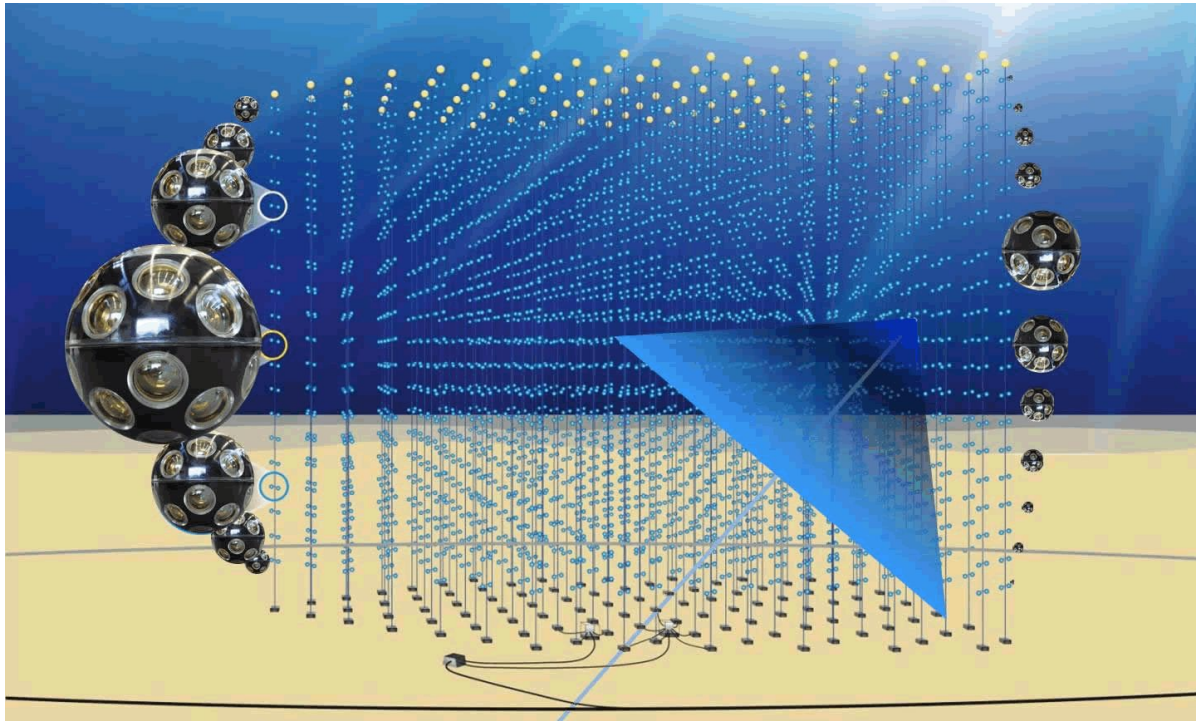# Use (Cases) of KM3NeT for WP5 batch processing

WP5 meeting 9th February 2022

J. Schnabel (FAU) for the KM3NeT collaboration

The KM3NeT experiment

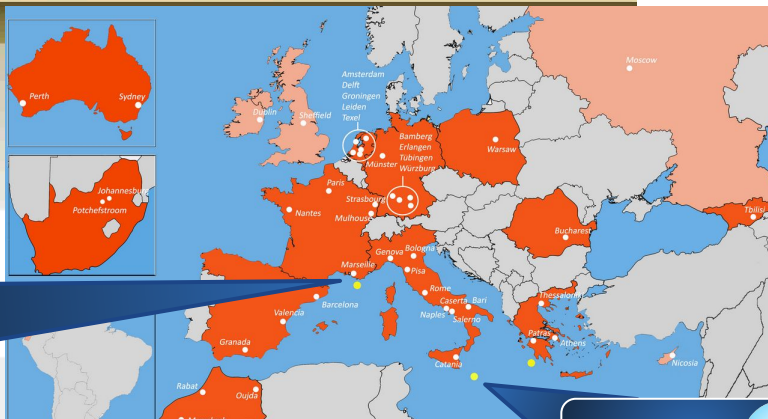**Water Cherenkov detector for high-energy neutrinos**
- Multi-PMT sensor modules
- Building blocks (BBs) of 115 DUs (lines)

Science goals
- astrophysics (ARCA)
- neutrino oscillations (ORCA)

Under construction
- 1 DU installed in ARCA
- 6 DUs installed in ORCA
- more to come this year

Where to use batch processing?

Main data processing (internal)

- complex workflows
- high-volume data
- multi-site computing

Under development

- defining workflow environment (nextflow, other workflow description)
- investigating DIRAC (in Concordia with CTA) for batch processing
- aiming for RUCIO for file management

target of the ESCAPE environment: high-level data → low requirements of batch processing
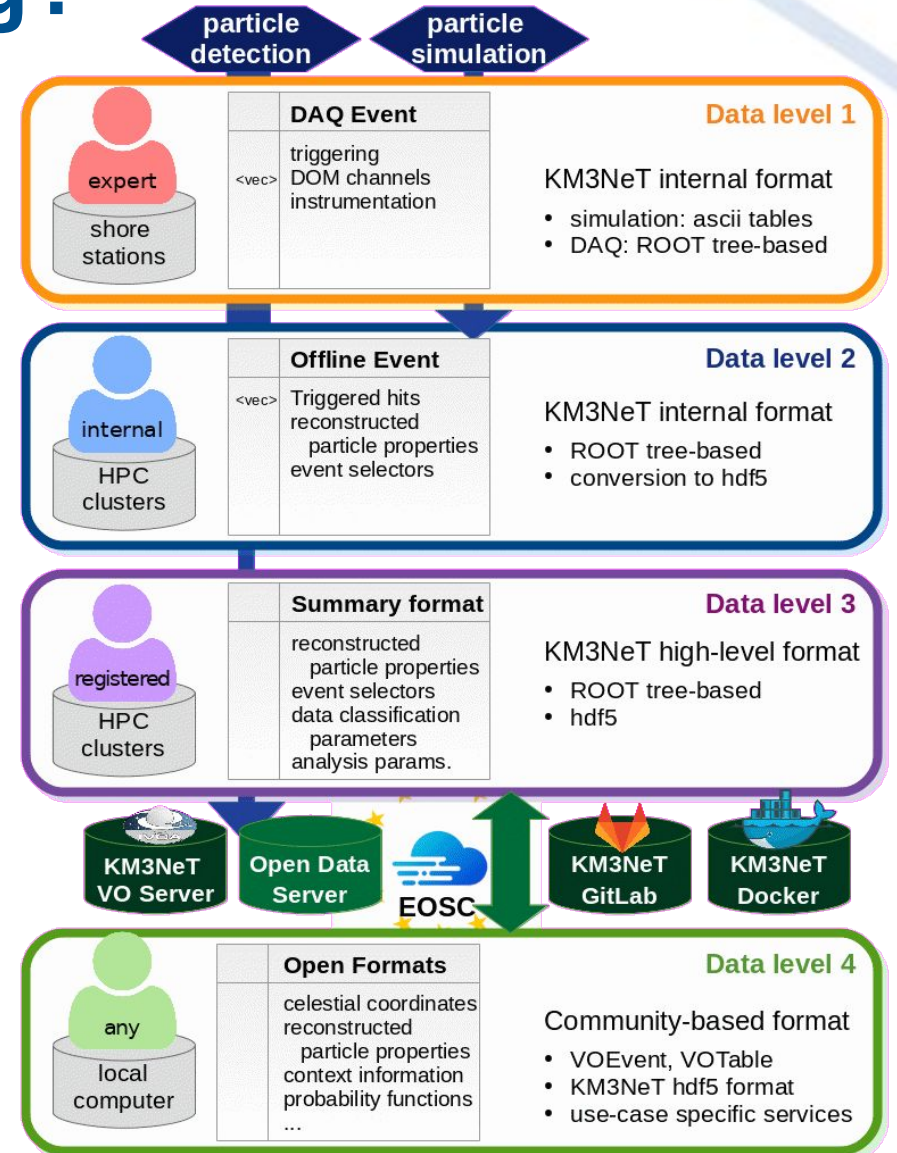
Current processing @CCLyon

- Software available in containers and through module system

- For data releases (simulation and/or data)
  - Data read from shore station, daily synch to computing center (DAC21 use case 1)
  - Calibration and reconstruction of data performed in processing campaigns (DAC21 use case 2)
  - Workflow management currently a (highly sophisticated) arrangement of batch scripts
    - Testing workflow description language (nextflow)
  - Parameters for releases defined in by experts
  - reprocessing and monitoring mostly custom made
    - Investigating use of DIRAC

- Current interface to ESAP: Reading processed files through Jupyter notebook, convert to open data format (DAC21 use case 3)

Main questions

- Is batch processing for full data release workflows in/out of scope?

  - Quite complex and customized

  - Chances for common developments?

- Can consider using batch processing for further work with high-level data (e.g. short dedicated simulations, ML applications, IRF generation based on containerized software)

  - Could define common use cases for this scenario

  - not necessarily requires batch processing