PROCEEDINGS OF SCIENCE

Advancing KM3NeT Data Management: Harnessing Snakemake and Grid Computing

Anna Sinopoulou[a,*] on behalf of the KM3NeT collaboration

[a] *Istituto Nazionale di Fisica Nucleare - Laboratori Nazionali del Sud*
Via S. Sofia, 62, 95125 Catania CT, Italy

E-mail: anna.sinopoulou@ct.infn.it

The KM3NeT collaboration is constructing two underwater neutrino detectors in the Mediterranean Sea: ARCA, optimized for the detection of astrophysical neutrinos, and ORCA, for the determination of the neutrino mass ordering via the observation of atmospheric neutrino oscillations. As the size of the detector increases, significant data volumes are acquired thus necessitating efficient data processing and management strategies. To address the data handling demands, the KM3NeT collaboration has implemented a hierarchical computing model similar to that of the LHC experiments, featuring a multi-Tiered structure for data processing and distribution. This model enables the integration of diverse computing resources, thereby improving data handling efficiency. The KM3NeT collaboration has already integrated Snakemake, a contemporary workflow management system, to guarantee portability, reproducibility, and scalability across various computational environments. Snakemake optimizes and automates complex data processing tasks, enhancing flexibility and efficiency in data management. During the last year, efforts have been focused on integrating Grid Computing resources into the KM3NeT computing model to improve computational capabilities. This integration seeks to address the rigorous needs associated with detector calibration, simulation, reconstruction, and analysis. In this document, an overview of KM3NeT's data processing framework, focusing on the implementation of Snakemake and the integration of Grid Computing resources.

The European Physical Society Conference on High Energy Physics (EPS-HEP2025)
7-11 July 2025
Marseille, France

[*]Speaker

ISSN 1824-8039. Published by SISSA Medialab.

PoS(EPS-HEP2025)571

1. Introduction

KM3NeT [1] is a multi-site deep-sea research infrastructure hosting two next-generation neutrino telescopes in the Mediterranean Sea: ARCA (Astroparticle Research with Cosmics in the Abyss) and ORCA (Oscillation Research with Cosmics in the Abyss). ARCA, located off the coast of Sicily, approximately at 3500 m of depth, targets high-energy neutrinos ($E \gtrsim 1$ TeV) from astrophysical sources such as supernova remnants, active galactic nuclei, gamma-ray bursts, and of cosmogenic origin. ORCA, deployed at ~2500 m depth off the southern coast of France, is designed to study neutrino oscillations and determine the neutrino mass ordering using atmospheric neutrinos.

Both detector arrays share the same Digital Optical Module (DOM) technology [2]: each DOM is a 42 cm-diameter pressure-resistant glass sphere housing thirty-one 3-inch photomultiplier tubes (PMTs) that provide near-4π coverage with nanosecond timing precision. Each Detection Unit (DU) comprises 18 vertically spaced DOMs, anchored to the seabed and kept upright by a buoy at its top. Upon completion, ARCA will consist of 230 DUs and ORCA of 115 DUs. When finished and fully operational, the KM3NeT detectors will produce a large amount of data, thus necessitating a robust data processing and Monte Carlo (MC) production procedure. In this document, the KM3NeT offline computing model, the run-by-run simulation approach, and the current workflow implementation are described.

2. Computing in KM3NeT (Tiered Architecture)

KM3NeT employs an *all-data-to-shore* principle, in which all photomultiplier tube (PMT) signals above threshold are digitized offshore and transmitted unfiltered to the on-shore data-acquisition stations [3]. This approach preserves complete information from the detector, allowing comprehensive offline reconstruction, calibration, and re-processing. The continuous data stream reaches rates of up to ~ 5GB/s, requiring an efficient filtering and reduction strategy before storage. To manage these data and their derived products, KM3NeT implements a three-tier computing architecture (Fig. 1), designed to ensure scalability, modularity, and traceability across all processing stages.

- **Tier-0 – Shore Station: Real-Time Data Acquisition and Preprocessing.** At the shore stations, the raw digital streams are received directly from the detector via optical links. Real-time processing is applied to these streams in order to perform noise filtering and hit clustering, apply first-level event triggers based on temporal and spatial coincidences of PMT signals, and execute quasi-online calibrations for time offsets and positioning. This stage is crucial to obtain real-time monitoring of the detector conditions and data-quality assessment. Through this procedure, the data rate is reduced by several orders of magnitude to a manageable filtered data stream (~ 5MB/s) that is transferred to Tier-1 centers for permanent storage and offline analysis.

- **Tier-1 – Offline Data Processing and Storage.** Tier-1 computing centers, located at major national facilities (e.g., CC-IN2P3, France and CNAF, Italy), handle the bulk of the offline data processing and Monte Carlo production. Tier-1 also supports the central database which

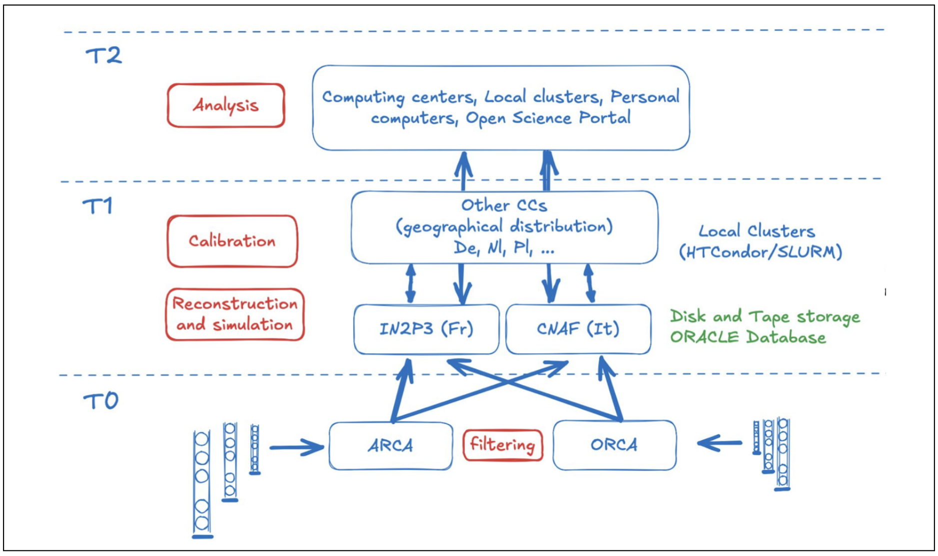

Figure 1: Schematic overview of the KM3NeT three-tier data-processing model.

stores detailed run configurations, detector status, and environmental measurements. These data are exploited to ensure consistency between data and simulation and allow for complete reproducibility of any data manipulation. Processed datasets, calibration constants, and simulation samples are archived and replicated within Tier-1 storage systems, ensuring data integrity and long-term accessibility.

- **Tier-2 – Analysis Layer.** At Tier-2 level, member institutes access processed data and simulation samples to perform physics analyses (oscillation studies, point-source searches, diffuse fluxes studies, etc.), make custom processing and user-defined reconstructions (e.g. tests for systematics uncertainties) as well as custom Monte Carlo studies. This tier level provides flexibility for collaboration-wide and individual analyses while ensuring full traceability of datasets and provenance of results.

3. KM3NeT Data Workflow Overview

The KM3NeT data-processing chain is structured in a series of Data Levels (DL0–DL6), each representing a distinct transformation stage from raw PMT signals to analysis-ready physics datasets. This layered workflow ensures full traceability of data products, reproducibility of processing, and efficient management of the experiment's high data throughput.

- **DL0 – Raw Hit Data Acquisition and Online Filtering (Tier-0)**. Digital signals ("hits") generated by the 31 PMTs within each Digital Optical Module (DOM) are transmitted in real time via optical links to the on-shore data-acquisition system. At the shore station, the Tier-0 computing resources perform the following processing of the raw data: low-level data integrity checks, real-time hit filtering to reject random noise and bio-luminescent bursts, and trigger algorithms to identify time- and space-coincident hits potentially produced by Cherenkov light from neutrino interactions. This first stage, also known as online filtering, reduces the raw data rate from several gigabytes per second to a manageable filtered stream, typically a few megabytes per second, while preserving all physics-relevant events.

- **DL1.1 – Run-Based Calibration**. Filtered data are stored into discrete runs of 3–6 hours, each associated with a well-defined detector configuration. At this level, the offline *per-run* calibrations—complementing the online ones already applied by the data filter at DL0—include the determination of PMT timing offsets and delays, the geometrical alignment obtained from the acoustic positioning system, and the monitoring of environmental parameters such as temperature, sea currents, and optical water properties. These calibrations ensure precise time and position reconstruction and are essential for correlating photon arrival times across the distributed detector volume [4, 5].
- **DL1.2 – Run-Based Monte Carlo Simulation**. In parallel to data processing, Monte Carlo simulations are generated for each run to reproduce the corresponding environmental and detector conditions, including the PMT calibration parameters applied during that data-taking period. This approach ensures that simulated events accurately reflect the detector response and can be directly compared to the reconstructed data, allowing also for systematic-uncertainty studies and detector-performance validation. Simulated samples include atmospheric muons, neutrinos, and optical background noise. The simulation chain sequentially encompasses event generation (MUPAGE [6], gSeaGen [7]), photon and particle propagation in the medium (KM3SIM [8, 9], JSirene [10]), and DAQ emulation ([11]).
- **DL2 – Event Reconstruction**. At the Tier-1 centers, calibrated data and simulated events are processed through advanced reconstruction algorithms. *Track-like* (muon-initiated events) and *shower-like* (hadronic/electromagnetic cascades) events are reconstructed using maximum-likelihood fits of photon arrival times to Cherenkov emission models. Outputs include reconstructed vertices, energies, and direction estimates [12]. At this stage, data are transformed from raw hits into fully reconstructed physics events suitable for higher-level analyses.
- **DL3 – Data Summary Trees (DST) Production**. Reconstructed events undergo quality filtering, classification, and reduction into compact Data Summary Trees (DSTs). These DSTs retain essential event-level observables, reconstruction parameters, and quality metrics while reducing storage size by more than an order of magnitude. DL3 datasets constitute the standard input for physics analyses within the collaboration.
- **DL4–DL6 – Aggregation and User-Level Data Products**. Multiple DL3 datasets are aggregated across many runs and detector configurations to form science-ready datasets (DL4–DL6). At these levels: events are combined into long-term data periods (e.g., detector seasons), standard selections for specific analyses (oscillation, diffuse flux, point-source searches) are applied, and derived products such as histograms, likelihood maps, or public data samples are produced. DL6 represents user-level, analysis-ready data, forming the basis for physics results, Monte Carlo comparisons, and public releases aligned with the *Open Science* policy adopted by the KM3NeT collaboration.

PoS(EPS-HEP2025)571

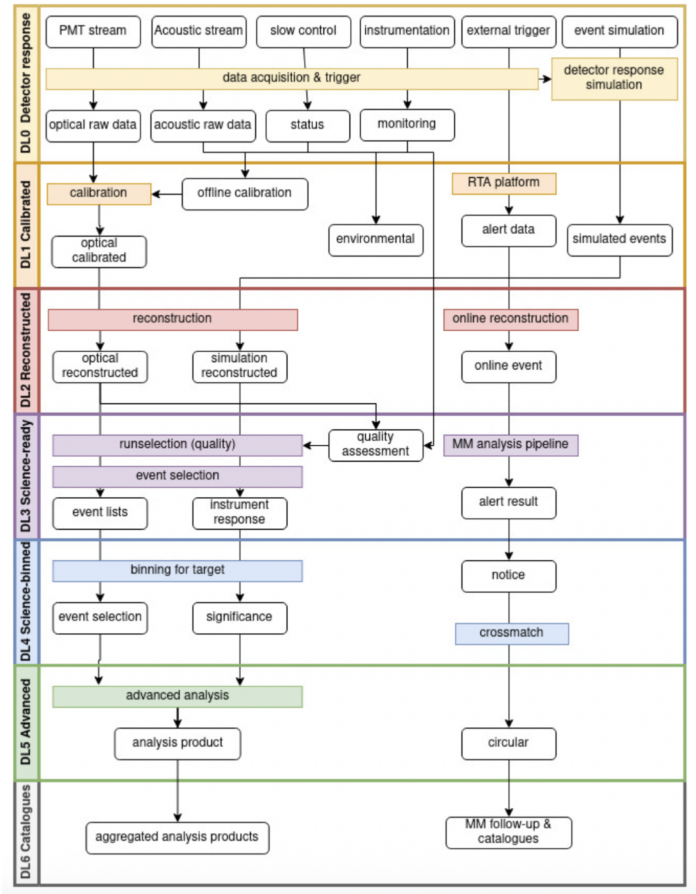

Figure 2: Block diagram of the KM3NeT processing chain for data and MC.

4. The KM3NeT run-by-run data-processing workflow

4.1 Snakemake: the backbone of KM3NeT data pipelines

The increasing volume and complexity of KM3NeT data, combined with the need to exploit multiple distributed computing sites, led to the adoption of Snakemake [13, 14] as the core workflow management system of the experiment. Snakemake provides a unified and scalable framework to orchestrate data and Monte Carlo (MC) processing, covering both calibration, reconstruction, and simulation, while ensuring reproducibility and efficient use of computational resources.

Snakemake allows for portability, running identically on local workstations, Tier-1 centers, and Grid [15] or HPC environments without modification. It is scalable, automatically constructing a directed acyclic graph (DAG) of dependencies to identify and execute parallel tasks efficiently and allows for reproducibility as each rule explicitly defines its inputs, outputs, and execution command, eliminating hidden scripts and manual preprocessing. Snakemake also offers fine-grained resource control, allowing rules to specify the number of CPU threads, memory, and time limits, which are translated into submission directives compatible with SLURM, PBS, or HTCondor. Finally, its modular structure facilitates code reuse and maintainability. Rules and configurations are shared across workflows, while containerization (Apptainer [16]/Singularity [17]) guarantees consistent environments across sites.

In the Snakemake implementation each data-taking run handled independently. This *run-based processing chain* enables concurrent execution across sites, improves CPU utilization, and simplifies recovery, as only failed tasks are re-executed. Resources are mutualized across runs, with early merging of intermediate outputs (e.g. photon-propagation files) to reduce duplication. Disk usage is optimized through automated cleanup once data are safely stored. Unlike older sequential pipelines, Snakemake removes strict run-order dependencies, executing jobs as soon as inputs become available, resulting in smoother scheduling and shorter turnaround times.
All processing steps run in containerized environments, encapsulating reconstruction software and dependencies to ensure bit-level reproducibility. Container images are distributed via CVMFS [18] or site registries, eliminating manual installation on compute nodes. Each rule declares its resource requirements which the scheduler enforces when submitting jobs to the batch system. To streamline submission across heterogeneous sites, the workflow defines a set of logical resource classes (for example, light-simulation, reconstruction, or merging). These classes group rules with similar computational profiles and are mapped to the appropriate partitions or queues at each site (e.g., short, long, or high-memory queues). This abstraction allows the same workflow configuration to run efficiently on different infrastructures without site-specific code changes. An integrated benchmarking and monitoring system records performance metrics and visualizes workflow status in real time, allowing operators to monitor progress, identify bottlenecks, and optimize future campaigns.

4.2 Entering Grid Computing: Future Developments in KM3NeT Data Processing

In order to meet the evolving needs of the increasing detector configurations, KM3NeT aims at transitioning toward a Grid-enabled distributed computing infrastructure. This shift will allow the experiment to leverage additional resources across multiple sites while maintaining workflow reproducibility and data integrity.

Grid computing unifies geographically-distributed computing centers under a common infrastructure. Sites may differ in hardware, operating systems, and batch systems, yet remain accessible to the user through a single interface. Authentication is based on digital certificates or token-based credentials, and data are accessed via distributed storage federations rather than local shared filesystems. Jobs are submitted through middleware services rather than directly to a cluster queue, and scheduling, data staging, and monitoring are handled transparently by orchestration tools. This evolution can be summarized as a transition from:

- One-site, homogeneous clusters $\rightarrow$ Multi-site, heterogeneous Grid nodes;
- Shared local storage $\rightarrow$ Distributed storage with data federation;
- Direct job submission $\rightarrow$ Middleware-based orchestration;
- Password authentication $\rightarrow$ Certificate or token-based access.

These changes introduce new challenges, such as heterogeneous environments, varying site policies, and diverse schedulers but also enable a leap in scalability and flexibility for computing model of the experiment.

Integration with the EGI Grid via DIRAC

To achieve seamless interaction with distributed computing sites, KM3NeT is integrating the workflow system with the European Grid Infrastructure (EGI) [19] through the DIRAC [20] platform. DIRAC acts as an intelligent layer between the Snakemake-managed production system and the underlying Grid middleware, and it handles job submission, monitoring, data transfer, and resource access across all participating sites. This integration enables automatic balancing of workloads, transparent recovery from job failures, and consistent tracking of job metadata across heterogeneous environments.

Distributed Data Management with Rucio

To complement distributed job execution, KM3NeT is deploying a prototype Rucio [21] service for data management and replication. Originally developed for the ATLAS experiment, Rucio provides a rule-based system for dataset registration, placement, and consistency checking. Snakemake workflows are being extended to interface directly with Rucio, enabling automatic data retrieval for input files and replication of output products across Tier-1 and Tier-2 centers. This ensures that every dataset required for reconstruction, calibration, or MC validation is accessible where computing resources are available, without manual intervention.

Finally, to maintain portability and reproducibility across this heterogeneous infrastructure, KM3NeT continues to rely on CVMFS and Apptainer (Singularity) containers. This dual approach guarantees that the same validated environment is used for every task, whether it runs on a local cluster, Tier-1 Grid node, or cloud resource.

5. Conclusions

The adoption of a run-by-run simulation strategy and a Snakemake-based, containerised workflow marks a significant step forward in managing the growing data and simulation volumes of KM3NeT. The system supports reproducible, scalable, and portable processing across multiple centres and has already enabled key physics results, such as the observation of an ultra-high-energy neutrino [22]. The event KM3-230213A (observed 13 Feb 2023) — a muon with estimated energy $\sim$120 PeV — was analyzed on reconstructed datasets produced via this workflow, with matched-condition simulations used for systematics. This computing backbone is central to the efforts of the collaboration towards the search for cosmic neutrinos and the study of neutrino oscillations and will serve the evolving needs of the experiment as deployment expands. This computing workflow guarantees that calibration will be able to process and analyze large volumes of data across sites.

References

[1] S. Adrián-Martínez et al., "Letter of intent for KM3NeT 2.0," J. Phys. G: Nucl. Part. Phys. **43**, 084001 (2016), https://doi.org/10.1088/0954-3899/43/8/084001.

[2] S. Aiello et al., "The KM3NeT multi-PMT optical module," J. Instrum. **17**, P07038 (2022), https://doi.org/10.1088/1748-0221/17/07/P07038.

[3] The Online Data Filter for the KM3NeT Neutrino Telescopes,xThe KM3NeT Collaboration, 2025, 2506.05881, arxiv, astro-ph.IM, https://arxiv.org/abs/2506.05881

[4] De Benedittis, A., on behalf of KM3NeT Collaboration. KM3NeT/ORCA Calibration Procedures and Capabilities. Phys. Sci. Forum 2023, 8, 44. https://doi.org/10.3390/psf2023008044

[5] The calibration units of KM3NeT, Le Breton, R. and Billault, M. and Boutonnet, C. and Champion, C. and Colonges, S. and Cosquer, A. and Creusot, A. and Henry, S. and Ilioni, A. and Keller, P. and Lagier, P. and Lahmann, R. and Lamare, P. and Lesrel, J. and Lindsey Clark, M. and Royon, J. and Riccobene, G. and Samtleben, D. and Van Elewyck, V. on behalf of the KM3NeT collaboration, 16, ISSN=1748-0221, http://dx.doi.org/10.1088/1748-0221/16/09/C09004, 10.1088/1748-0221/16/09/c09004, 09, Journal of Instrumentation, IOP Publishing, 2021

[6] G. Carminati, M. Bazzotti, S. Biagi et al., "MUPAGE: A fast atmospheric muon generator for neutrino telescopes based on parametric formulas," arXiv preprint (2009), https://arxiv.org/abs/0907.5563.

[7] S. Aiello et al., "gSeaGen: The KM3NeT GENIE-based code for neutrino telescopes," Comput. Phys. Commun. **256**, 107477 (2020), https://doi.org/10.1016/j.cpc.2020.107477.

[8] CERN, "GEANT program manual," CERN Program Library W5013 (1993).

[9] A.G. Tsirigotis, A. Leisos, S.E. Tzamarias, "HOU Reconstruction & Simulation (HOURS): A complete simulation and reconstruction package for very large volume underwater neutrino telescopes," Nucl. Instrum. Meth. A **626-627**, S185-S187 (2011), https://doi.org/10.1016/j.nima.2010.06.258.

[10] M. de Jong, E. van Campenhout, "The probability density function of the arrival time of Čerenkov light," arXiv preprint (2023), https://arxiv.org/abs/2305.19626.

[11] A. Albert et al., "Monte Carlo simulations for the ANTARES underwater neutrino telescope," J. Cosmol. Astropart. Phys. **2021**, 064 (2021), https://doi.org/10.1088/1475-7516/2021/01/064.

[12] Melis, Karel and Heijboer, Aart and De Jong, Maarten, "KM3NeT/ARCA Event Reconstruction Algorithms", doi = "10.22323/1.301.0950", PoS, 2017,"ICRC2017", 950

PoS(EPS-HEP2025)571

[13] F. Mölder, K.P. Jablonski, B. Letcher et al., "Sustainable data analysis with Snakemake," F1000Research **10**, 33 (2021), https://doi.org/10.12688/f1000research.29032.1.

[14] J. Koester, "Snakemake 8.19.0 Documentation," (2014-2024), https://snakemake.readthedocs.io/en/stable/.

[15] Foster I, Kesselman C, Tuecke S. The Anatomy of the Grid: Enabling Scalable Virtual Organizations. The International Journal of High Performance Computing Applications. 2001;15(3):200-222. doi:10.1177/109434200101500302

[16] Apptainer Team, "The Apptainer software," https://apptainer.org/docs/admin/main/index.html.

[17] Kurtzer GM, Sochat V, Bauer MW (2017) Singularity: Scientific containers for mobility of compute. PLoS ONE 12(5): e0177459. https://doi.org/10.1371/journal.pone.0177459

[18] CERN, "The CVMFS system," https://cernvm.cern.ch/fs/.

[19] S. Andreozzi et al., "The European Grid Infrastructure (EGI)", Journal of Grid Computing 9 (2011) 273–286. DOI: 10.1007/s10723-010-9176-0

[20] A. Tsaregorodtsev et al., "DIRAC: A community grid solution," J. Phys.: Conf. Ser. **119**, 062048 (2008), https://doi.org/10.1088/1742-6596/119/6/062048.

[21] Barisits, Martin et al., "Rucio: Scientific Data Management", Computing and Software for Big Science, vol. 3, 9/8/2019, 1, 11,issn = "2510-2044", "10.1007/s41781-019-0026-3",https://doi.org/10.1007/s41781-019-0026-3, "Springer International Publishing"

[22] S. Aiello et al. (The KM3NeT Collaboration), "Observation of an ultra-high-energy cosmic neutrino with KM3NeT," Nature **638**, 376 (2025), https://doi.org/10.1038/s41586-024-08543-1.