# The KM3NeT data acquisition system - Status and evolution

*Tommaso* Chiarusi[1,2,*], *Emidio* Giorgio[2,**], and *Daniele* Zito[2,***] on behalf of the KM3NeT Collaboration

[1]INFN-Sezione di Bologna, Viale Berti-Pichat 6/2, 40127 Bologna, Italy
[2]INFN-Laboratori Nazionali del Sud, via S.Sofia 62, 95123 Catania, Italy

**Abstract.** In this contribution, we present the data acquisition system system of the KM3NeT neutrino telescopes, ARCA and ORCA, already operating while under construction at the bottom of the Mediterranean Sea. The DAQ system implements a modular and scalable hardware-triggerless streaming readout of the optical modules in the telescopes. Handling a raw incoming throughput up to hundreds of Gbps, it is integrated with the online alert system of the KM3NeT Multimessenger program. The system will evolve according to a new design, currently under validation, which will grant the extension of the telescope to the aimed cubic kilometer scale.

## 1 Introduction

The data acquisition (DAQ) system is an integral part of the KM3NeT infrastructure [2], responsible for collecting and storing data from the undersea Cherenkov detectors. Currently under construction in two different abyssal sites of the Mediterranean Sea, the ARCA and ORCA KM3NeT telescopes feature a common DAQ design which, although with different implementations, is designed to be operational since the first stages, scaling with the detector size. It includes also the control and monitoring systems of the off-shore installation as well as the on-shore infrastructure, which is made of the time-synchronization GPS reference, the networking fabrics, and the computing and storage resources. Although running with two independent DAQ installations, the two telescopes data taking are combined at higher level by the KM3NeT online Multimessenger system [3] In this contribution all the key elements of the DAQ system for the already running detectors are reviewed, together with the ancillary procedures to deploy, upgrade and maintain it. Eventually, the future DAQ paradigm, the *Full White Rabbit scenario* based on the White Rabbit paradigm [4] and currently under review, is presented.

## 2 The KM3NeT detector and the data streams

The KM3NeT detection elements such as the Digital Optical Modules (DOMs), the Detection Units (DU) and the Detection Unit-Base Modules (DU-BM), as well as the DOM and

---

*e-mail: tommaso.chiarusi@bo.infn.it
**e-mail: emidio.giorgio@lns.infn.it
***e-mail: daniele.zito@lns.infn.it

DU-BM readout electronics (the Central Logic Board - CLB) are reported in details in these proceedings ([1]) and in [2]. DOMs and DU-BMs are the nodes of the submarine branch of a global ethernet network infrastructure, which grants the communication between the detector and the related shore station, and provides the DOMs and DU-BMs with the distribution of the absolute time with a sub-nanosecond precision by means of the White Rabbit protocol. The data taking of each DOM is independent to the others, without any local trigger mechanism implemented in the DOM electronics. For this reason data from the DOM photomultipliers (PMTs) and the acoustic sensors are acquired according to a continuous streaming readout (triggerless) mode. Recorded data are packed in continuous sequences of *optical* and *acoustic* frames, each referring to a 100 ms time-slice duration. The digitisation of the PMT pulses after each photon-electron conversion leads to a 6-byte *hit* information, which includes the PMT identification, a hit-time reference and the duration of the pulse above a certain threshold. With PMT thresholds set at 0.5 photoelectrons, average ~6 kHz and ~8 kHz hit rates are measured for ARCA and ORCA, respectively, mostly due to the $^{40}$K decays and bioluminescence contamination. A 20 kHz acquisition veto per PMT is set to avoid the contamination by a too large optical background. The average measured DOM optical throughput is 12 Mbps,while the acoustic streaming accounts for roughly the half. Globally, the throughput from such fast acquisition data streams is ~350 Mbps per DU. The DAQ system is designed around a conservative throughput value which is at least a factor three larger than the measuerd one, targeting to O(100) Gbps per building block. Few precents of the global throughput is due to the ancillary instrumentations, sampled at a minor rate (from 0.1 to 1 Hz), the slow-control data exchanges and the White Rabbit stream.

## 3 The DAQ system in the *Broadcast* scenario

The *Broadcast* scenario implies an asymmetric connection topology: a single downlink channel from the shore-station; one direct uplink per CLB back to on-shore resources ([5]). Such design is already active in the current installation of ARCA (21 DUs) and ORCA (14 DUs). It will serve both detectors up to the completion of the first deployment stage, i.e. 32 DUs and 48 DUs, respectively. The evolution of the connection design for the next deployment stages is reported in Section 7. For the asymmetry of the Broadcast scenario, the firmware of the White Rabbit switches (WRSs) used in the shore station and of the CLB has been customised, diverting from the the CERN Standard. Only the DU-BMs are in a true Master-Slave relation with the WRSs on shore, while the DOMs piggyback the incoming White Rabbit stream for synchronising, sending back to shore only the rest of the data streams to a fabric of standard (not WR) switches. Indeed, such standard high-throughput switches (various DELL® models) are configured with a dedicated implementation of the Software Defined Networking technology ([8]).

## 4 The DAQ processing chain

On-shore, the incoming optical and acoustics data streams are routed to the so called *Trigger and Data Acquisition System* (TriDAS), a multi-stage processing chain realised with C++ software applications. The first stage is the aggregation layer. A few concurrent processes called *DataQueue* collect the incoming optical and acoustic streams sent by a sector of the telescope, fragmented in various UDP datagrams, and reconstruct the frames recorded by the CLBs. The DataQueues route the frames downstream via TCP links. A valid reference for UDP and TCP protocols is [7]. The second stage is the filtering/processing layer, made of a number of concurrent processes called Optical and Acoustic *DataFilters*. Each Optical DataFilter receives from the DataQueues the entire set of CLB frames recorded during a

given timeslice. Different processes simultaneously work on data recorded during different timeslices. Each Optical DataFilter implements various causality-trigger algorithms, requiring precise time-space constraints for the recoreded hits, and quickly identifies possible tracks of neutrino induced muons or shower-like events. An additional trigger algorithm selects data that will be later scrutinised as candidate neutrino events from supernovae[9]. The reduction of the incoming stream after the triggers is smaller than $1 : 10^4$. On each Optical DataFilter, a circular buffer stores raw data which can be dumped and analysed off-line after external alerts. The trigger algorithms used in the Optical Data Filters are implemented in a KM3NeT C++ programming framework, called Jpp [10], which is also used for off-line analyses and Monte Carlo production, ensuring identical processing for both real data and simulations. The Acoustic Data filters reconstruct the detection-time, also referred to as time of arrival (TOA), of the sound emitted by various submarine beacons, each corresponding to a given sound-waveform, by means of cross-correlation algorithms. Computed for each DOM and DU-BM, the TOAs are at the base of the off-line detector positioning system. The last stage is the data-writing and storage layer. By means of a custom data-dispatcher of Jpp which operates as middleware, each Optical DataFilter asynchronously transmits the selected data to a single collecting process, the DataWriter, which writes the incoming stream into ROOT files [11] on local storage systems.The reconstructed TOA are written to disk directly by the Acoustic DataFilters, in a dedicated binary format. Two copies of the same recorded files are secured in two different remote repositories: CNAF-INFN, in Italy, and CC-Lyon, in France.

## 5 Management of the DAQ context

The number of the DataQueue, Optical and Acoustic DataFilter processes necessary to guarantee a smooth data-taking depends both on the global throughput of the detector and the complexities of the filtering algorithms. The larger the data-load to process, the more resources are allocated. This is achieved by distributing replicas of the same process on different high performing servers (>32 cores at 2GHz each, > 128 GB RAM), which are interconnected via a 25 GbE networking infrastructure. This approach is called the MUltiple SErvers with DIfferent Processes (MUSEDIP) scenario, whose modularity by desing allows for scaling with the situation complexity. The TriDAS processes are orchestrated together with the detector by another software process, the Control Unit, whose latest evolution is described in details in these proceedings ([6]). The DAQ framework, beyond the ARCA and ORCA cases, is also exploited in other contexts, such as the DOM, DU-BM and DU integration and validation setups. TriDAS and Control Unit processes are deployed via a dedicated framework called AIACE (Automatic Installation And Configuration Environment): based on the ANSIBLE technology, AIACE allows to define a set of environments, one for each DAQ installation, and then to proceede from scratch with the installation and configuration of all the devices with the requested applications. All software is deployed as Docker ([12]) images, activated in due containers by AIACE. This approach is extremely efficient for easily implementing and maintaining the MUSEDIP scenario in all the DAQ contextes.

## 6 The Monitoring and the Multimessenger processing branches

The selected events, with some complementary information, are mirrored to both the monitoring system and the alert system for the Multi-Messenger program. The monitoring system is a collection of web appplications written in Python language which shows the activity status of the detector parts, plots of the the trigger algorithms efficiency, the synchronisation status and the quality of the data taking. The Multi-messenger alert system is a further branch of

online contex, separately active in both ARCA and ORCA shore stations, performing a more accurate analysis of the triggered events, assigning them a quality score. High score events are sent to a central processing facility, presently running in the ORCA shore station, where a final analysis step is performed by searching for signals from supernovae or other transient phenomena. This is done by combining the selected data with external alerts as well as releasing alerts via the GCN [13], the SNEWS 2.0 [14] and other systems. In order to send alerts with the shortest delay, the full DAQ processing plus the online seraches are finalised within 15 seconds from the reception of the first raw data on-shore.

## 7 Evolution of the DAQ system

The design goal of two building blocks for ARCA cannot be achieved within the Broadcast scenario. In fact, in order to guarantee all the additional uplinks for the CLBs yet to be deployed, it would be necessary a number of optical fibres larger than what is available with the already implemented sea-floor network. For this reason, new aggregation layers will be implemented directly in the DUs: a pair of White Rabbit switches will be placed into each DU-BM, called *Wet*-White Rabbit switches (WWRS), after the redefinition of the form factor of the switch backplane to fit the DU-BM space. The two WWRS will connect all the CLBs of one DU, allowing also *cold* redundancies, and provide 2x 1Gbps ethernet uplinks to shore. Such an innovative approach, called *Full White Rabbit* scenario, implies various design changes. The main ones affecting the DAQ concern both the detector and the shore station: the new DOM-CLBs will become pure White Rabbit Slave units, with a new dedicated firmware. On-shore, standard White Rabbit switches will be used for implementing the front-end switch fabric (called *Dry* WRS). Additional advantages of this new design will be the reduction of connections on-shore by at least one order of magnitude with respect to the Broadcast scenario; the direct support by the international White Rabbit community (no customisation is foreseen in the White Rabbit firmware); a simplified time-calibration procedure for the CLBs embedded in the standard White Rabbit protocol. Despite the differences, Broadcast and Full White Rabbit parts of the detector will contribute together to the data taking, and their data streams will be combined at the TriDAS and Control Unit levels. At present, the Full White Rabbit approach is under production-readiness review. The new DUs following this design are expected to be produced by the end of 2023.

## References

[1]  Simone Biagi in these Proceedings
[2]  S. Adrián-Martínez et al., J. of Physics G: Nucl. and Part. Phys, **43**, 084001 (2016)
[3]  W. Assal et al., J. of Instrumentations, **16**, C09034 (2021)
[4]  White Rabbit. `http://www.ohwr.org/projects/white-rabbit`
[5]  S. Aiello et al., arXiv:2210.13328, accepted for publication on JINST
[6]  S. Aiello et al., Computer Physics Comminications, **256**, 107433 (2020)
[7]  L. Dostàlek, A. Kebalova, *Understanding TCP/IP* (Packt Publishing, 2006)
[8]  T. Chiarusi, E. Giorgio, EPJ Web of Conferences, **207**, 06009 (2019)
[9]  S. Aiello et al., Eur. Phys. J. C, **82**, 317 (2022)
[10]  The JPP framework, `https://sftp.km3net.de/documentation/Jpp/v17.2.0/`
[11]  The ROOT system, `https://root.cern.ch/`
[12]  The Docker system, `https://www.docker.com/`
[13]  The Gamma-ray Coordinates Network, `https://gcn.gsfc.nasa.gov/`
[14]  The SuperNovae Early Warining System, `https://snews.bnl.gov/`