Journal of Physics: Conference Series

PAPER • **OPEN ACCESS**

The Trigger and Data Acquisition System for the KM3NeT-Italy neutrino telescope

To cite this article: T Chiarusi *et al* 2017 *J. Phys.: Conf. Ser.* **898** 032042

The Trigger and Data Acquisition System for the KM3NeT-Italy neutrino telescope

T Chiarusi [1], M Favaro [2], F Giacomini [2], M Manzali [2 3], A Margiotta [1 4] and C Pellegrino [1 4]

[3] INFN Bologna
[2] INFN CNAF
[1] Università degli Studi di Ferrara
[4] Università di Bologna

E-mail: `matteo.manzali@cnaf.infn.it`

Abstract. KM3NeT-Italy is an INFN project that will develop the central part of a submarine cubic-kilometer neutrino telescope in the Ionian Sea, at about 80 km from the Sicilian coast (Italy). It will use hundreds of distributed optical modules to measure the Cherenkov light emitted by high-energy muons, whose signal-to-noise ratio is quite disfavoured. In this contribution the Trigger and Data Acquisition System (TriDAS) developed for the KM3NeT-Italy detector is presented. The "all data to shore" approach is adopted to reduce the complexity of the submarine detector: at the shore station the TriDAS collects, processes and filters all the data coming from the detector, storing triggered events to a permanent storage for subsequent analysis. Due to the large optical background in the sea from 40K decays and bioluminescence, the throughput from the sea can range up to 30 Gbps. This puts strong constraints on the performances of the TriDAS processes and the related network infrastructure.

1. Introduction

The INFN's project KM3NeT-Italy, supported with Italian PON (National Operative Programs) fundings, is the inner core of the KM3NeT Cherenkov neutrino telescope for astrophysical searches using neutrinos as a probe. The detector consists of 8 vertical structures, called towers, instrumented with a total number of 672 Optical Modules (OMs) and its deployment is ongoing 3500 meters deep in the Ionian Sea in front of the south-east coast of Portopalo di Capo Passero, Sicily (Italy) [1][2]. A tower is a semi-rigid vertical structure composed by a sequence of 14 horizontal structures in marine grade aluminum named floors (Figure 1). Each tower is anchored to the seabed and kept vertical by an appropriate buoyancy on the top. Each floor, which is 8 m long, hosts six optical modules: two at either end and other two in the middle. Each optical module contains a 10-inch Photo-Multiplier Tube (PMT), a high-voltage supply circuit, a Front End Module (FEM) and an optical system for timing calibration.

The detection principle exploits the Cherenkov light from relativistic particles outgoing high-energy neutrino interaction within a fiducial volume around the telescope. In order to reduce the complexity of the underwater detector, the *all data to shore* approach is assumed, demanding to a Trigger and Data Acquisition System (TriDAS) [3] software running at the shore station. The collected data stream from all the towers is largely affected by the optical background in the sea [4], mainly due to the ^{40}K decays and bioluminescence bursts. Ranging up to 30 Gbps, such

Figure 1. Sketch of the floor structure.

a large throughput needs to be filtered and reduced before being stored permanently, for this reason the TriDAS software aim to apply a minimum of 1:1e4 data reduction when at regime. This puts strong constraints on the required TriDAS performances and the related networking architecture.

In the following sections we describe the final implementation of the acquisition system, the user and management interfaces and the testbed infrastructure. Finally, we present results of the scalability tests that demonstrate the system capabilities.

2. The TriDAS software

The TriDAS software has been developed to acquire and filter the data stream coming from the KM3NeT-Italy detector. Figure 2 shows the core acquisition components (TriDAS Core), the control components (WebServer and GUI) and their interactions with external services, presented in the next subsections. The C++ programming language is used for the development of the TriDAS Core programs, whereas the WebServer component is written in PHP and the GUI is based on the AngularJS framework. For the development of TriDAS, *git* has been chosen as version control system, due to its strong support for non-linear development that makes it ideal for projects that require an extensive collaboration. All the code is open-source and hosted on BitBucket [5], a freely available web-based service able to host *git* repositories. In the following subsections each component of TriDAS is briefly described.

2.1. Floor Control Module Server (FCMServer)

The FCMServers represent the interface of TriDAS with the data from the off-shore detector. They perform the read-out of data through dedicated electronic boards called NaNet3 [6], which is a custom FPGA-based readout board hosted on a FCMServer through a Peripheral Component Interconnect Express (PCIe) slot. Each NaNet3 is point-to-point connected to up to 4 floors through optical fiber links. The communication is bidirectional and allows to propagate the on-shore GPS time to the floor electronics and the read-out of the data that are stored in the host memory. Each FCMServer merges the data coming from each connected floor into a single TCP/IP data-stream in the form of *hits*, that are variable length structures. Each hit is made of one or more *DataFrames*, another variable length structure, with fixed maximum size. Each DataFrame contains the OM identifier, a GPS absolute timestamp, the total electric charge collected by the PMT and an array of charge signal samples. If the number of samples exceeds the limit for a single DataFrame, a new one is created by the electronics and marked as a subsequent fragment of the first frame. Upon successful TCP connection, each FCMServer sends the data-stream to the connected HitManager.

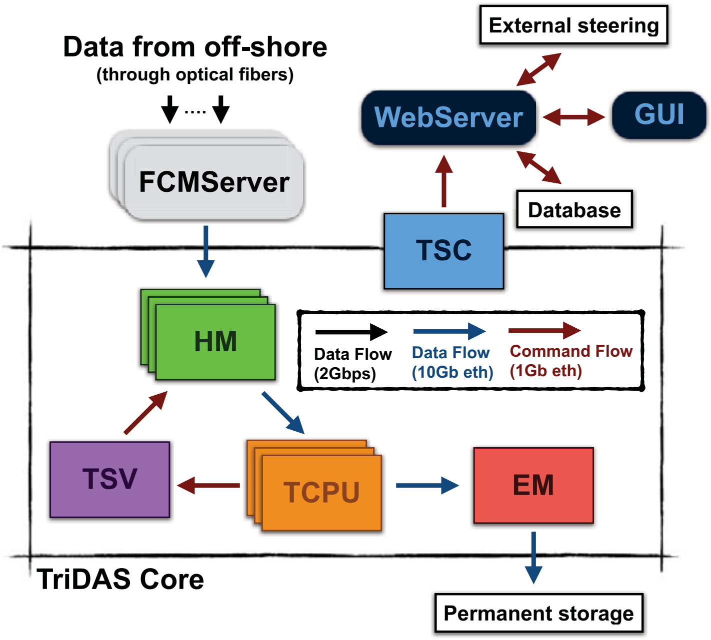

Figure 2. Scheme of TriDAS components and their interactions with external services.

2.2. HitManager (HM)

The HMs are the first aggregation stage for the incoming data-stream. Each HM is connected to a subset of FCMServers and it receives data from a portion of the detector called *Sector*. Hits are separated into arrays, one for each OM, and divided into temporally disjoint subsets called *SectorTimeSlices*. A *TimeSlice* is a temporal window of fixed duration, typically 200 ms. On completion of one of its SectorTimeSlices, each HM sends it to one TriggerCPU. The TriDAS SuperVisor assigns one TimeSlice to a precise TriggerCPU so that each HM sends its SectorTimeSlices, belonging to the same TimeSlice, to one TriggerCPU.

2.3. TriggerCPU (TCPU)

The TCPUs are responsabile for the online analysis. Each TCPU receives from the HMs all the SectorTimeSlices that belong to the same TimeSlice, creating the so called *TelescopeTimeSlice* (TTS). This means that each TCPU has the snapshot of the whole detector during a specific TimeSlice. After its creation, each TTS is analyzed with a 2-steps trigger system:

- The Level 1 (L1) algorithms search for hit coincidences and charge excess along the whole TTS, identifying interesting portions of data, called *events*, made of the hits occurring in a time window 6 μs-wide centred in the trigger seed. The event window is increased if another trigger condition is satisfied within it.
- The Level 2 (L2) algorithms implement more complex conditions that operate on L1 events. If a L2 is satisfied, the event is marked to be saved on permanent storage.

Each TPCU node has 32 physical cores and it runs a single process that implements a parallel processing of the TTS by means of a fixed number of worker threads, defined by the run-setup file. Each worker thread retrieves a TTS from a shared queue and it sequentially performs the 2-steps analysis previously described. This approach leads to an event-level parallelism rather than an algorithm-level one.

IOP Conf. Series: Journal of Physics: Conf. Series **898** (2017) 032042 doi:10.1088/1742-6596/898/3/032042

2.4. Event Manager (EM)
The EM is the software component of TriDAS dedicated to the storage of the triggered data. It collects the events from all the TCPUs and writes them on the local storage in the form of a Post Trigger (PT) file. The PT file contains also the full run-setup and the positions and calibration information of the OMs. In this way, each PT file is self consistent with the bulk of informations needed for the first stage of offline reconstruction. Moreover, the replication of information in each PT file doesn't introduce a relevant overhead due its limited size (few KiloBytes) with respect to the maximum dimension of a PT file, that is configurable and it is usually in the order of GigaBytes of data.

2.5. TriDAS SuperVisor (TSV)
The TSV supervises the data exchange between HM and TCPU, taking note of the processed TelescopeTimeSlices. When a TCPU is ready to handle new data, it sends a token to the TSV, that assigns to that TCPU a new TimeSlice among those not yet processed. The computation time required by a single TTS can really differ from time to time, depending on the quantity and disposition of the hits. For this reason the protocol chosen for the token distribution follows a "first-come first-served" logic, in which the first TCPU that has sent a token is the first that will receive a new TTS to be analyzed.

2.6. TriDAS Controller (TSC)
The TSC is the software interface that permits to control the entire TriDAS environment. Its purpose is to organize and control the launch of each software, allowing a correct acquisition and real time analysis of the data. In order to achieve this functionality the TSC implements a simple hierarchical state machine with four states, as shown in Figure 3.

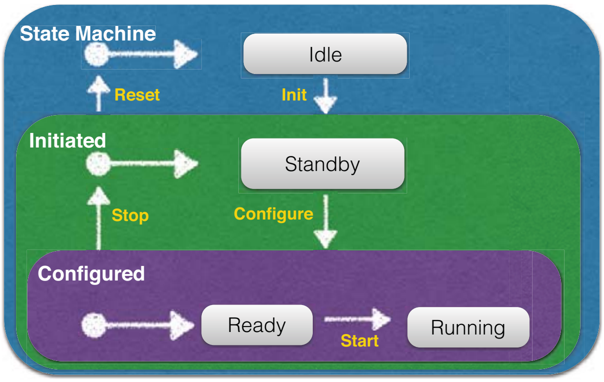

Figure 3. State machine of the TSC. The run-setup file is retrieved during the *init* transition.

2.7. WebServer
The WebServer is the unique entry point for operating the TriDAS Core. This component provides a set of RESTful APIs which allows to communicate with the TSC. Therefore, it provides user authentication based on hierarchical configuration implemented via different privileged groups. Despite the TSC is a local-single-client program, the WebServer can be contacted from several different concurrent users at a time. The WebServer allows to control the DAQ only one user at a time via an escalation procedure that permits users to acquire the control of the TSC. Finally, the WebServer can communicate instantly feedback and alarms through the use of WebSockets, implementing a real-time feedback to the users.

2.8. GUI
The GUI is a web application that graphically represents information provided by the WebServer and acts as control interface for the user.

3. Preliminary tests
We set up a testbed at the INFN section of Bologna in order validate the TriDAS software in all of its parts. The testbed is called Bologna Common Infrastructure (BCI) and it is a scaled version of the official acquisition farm present at the shore station infrastructure of Portopalo di Capo Passero. The hardware technologies present at the BCI and its network topology are similar to those adopted for the official acquisition farm. Figure 4 presents a schematic view of the BCI and it introduces all the servers used in the tests. Hereby the legenda explaining the connection typologies:

- *(A) (B)* Private LANs for data transfer from FCMServers and simulators to HMs.
- *(C)* Private LAN for data transfer from HMs to TCPUs.
- *(D)* Private LAN for data transfer from TCPUs to EM.
- *(E)* INFN-Internal Public LAN. This LAN grants the control services and access to the servers.

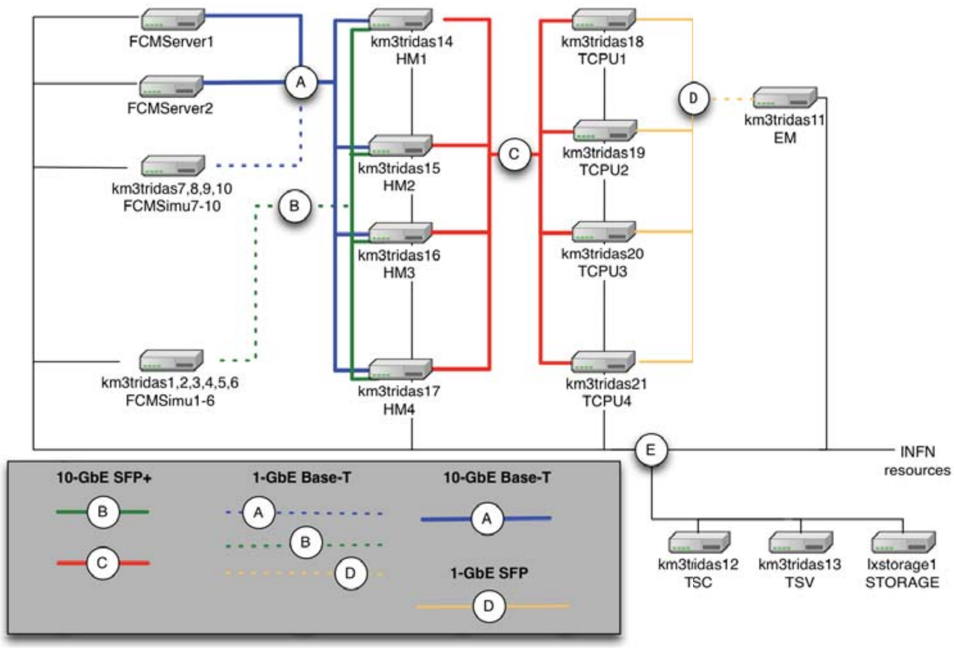

Figure 4. BCI network topology.

3.1. Scalability tests
In order to prove the scalability of TriDAS we performed different simulation tests at the BCI. The aim of these tests is to observe how the system behave increasing the data load (i.e. the number of towers). Due to the limited number of available nodes at the BCI we were able to simulate up to 4 towers. In order to reproduce the data throughput coming from the detector we developed a software called FCMSimu that is able to simulate the data stream coming from a single floor with a parametric rate of random hits. Preliminary studies performed at the Capo Passero deep-sea site have shown a background noise with an average hits rate of $\sim$ 55 kHz

and less than 10% of peaks over $\sim$ 100 kHz [4]. We decided to perform the scalability tests simulating an heavier background in order to validate the software in a critical scenario. For this reason we ran the FCMsimu processes with a constant hit rate of $\sim$ 100 kHz for each optical module.

For each test we measured the time needed to analyze a TTS of 200 ms. The design of TriDAS allows to execute concurrently several TCPU processes on different nodes: the setup adopted for these tests foreseen to use 4 TCPU nodes, each one able to elaborate 20 TTS in parallel, for a total of 80 parallel TTS. This means that the system is able to run as long as the time needed to analyze a single TTS is less than the duration time of a TTS multiplied by the number of TTS parallely processed by the system; in the described setup every TCPU node requires at most 0.2 s $\times$ 80 = 16 s for processing a TTS. Scalability tests were performed scaling from 1 to 4 towers and even in the worst case the mean time required to analyze a TTS is below the maximum time allowed by this setup, being 15.34 s against 16 s. Figure 5 shows the computation times measured divided by the number of TTS parallely processed for each test performed.

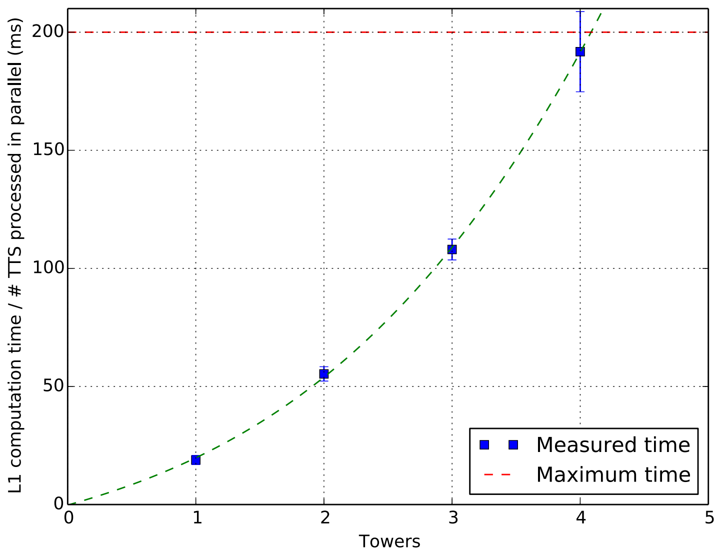

Figure 5. TTS computation time as function of the number of Towers

4. Conclusions

An on-line Trigger and Data Acquisition System (TriDAS) has been designed and implemented to handle the data coming from the KM3NeT-Italy underwater neutrino detector. Different validation and scalability tests have been performed running the software on a dedicated testbed that is a scaled version of the official acquisition farm. Each software component has been successfully validated and the scalability tests have shown that TriDAS is stable and matches the requirements. More investigations will be carried out, thanks to the expansion of the testbed hosted at the BCI that will allow to test the system against a 8 towers detector.

References

[1] S. Aiello et al., *Measurement of the atmospheric muon depth intensity relation with the NEMO Phase-2 tower*, DOI: 10.1016/j.astropartphys.2014.12.010, Astroparticle Physics (2015).

[2] T. Chiarusi, M. Spurio, *High-energy astrophysics with neutrino telescopes*, DOI: 10.1140/epjc/s10052-009-1230-9, The European Physical Journal C (2010).

IOP Conf. Series: Journal of Physics: Conf. Series **898** (2017) 032042 doi:10.1088/1742-6596/898/3/032042

[3] C. Pellegrino, et al., *The trigger and data acquisition for the NEMO-Phase 2 tower*, DOI 10.1063/1.4902796, AIP Conference Proceedings (2014).
[4] M. Pellegriti et al., *Long-term optical background measurements in the Capo Passero deep-sea site*, DOI: 10.1063/1.4902780, AIP Conference Proceedings (2014).
[5] TriDAS web site, https://bitbucket.org/chiarusi/tridas.
[6] R. Ammendola et al., *NaNet3: The on-shore readout and slow-control board for the KM3NeT-Italia underwater neutrino telescope*, DOI: 10.1051/epjconf/201611605009, EPJ Web of Conferences 116, 05008 (2016).