Optimized framegrabber for the Cherenkov telescope array

Miguel Jiménez-López
Jorge Manuel Machado-Cano
Manuel Rodríguez-Álvarez
Maurice Stephan
Gianluca Giavitto
David Berge
Javier Díaz

SPIE.

Optimized framegrabber for the Cherenkov telescope array

Miguel Jiménez-López,[a,*] **Jorge Manuel Machado-Cano,**[b] **Manuel Rodríguez-Álvarez,**[a] **Maurice Stephan,**[c,†]
Gianluca Giavitto,[d] **David Berge,**[c,d] **and Javier Díaz**[a]
[a]University of Granada, Granada, Spain
[b]Seven Solutions, Granada, Spain
[c]University of Amsterdam, GRAPPA, Amsterdam, The Netherlands
[d]Deutsches Elektronen-Synchrotron, Zeuthen, Germany

Abstract. Our contribution presents a high bandwidth platform that implements traffic aggregation and switching capabilities for the Cherenkov telescope array (CTA) cameras. Our proposed system integrates two different data flows: a unidirectional one from the cameras to an external server and a second one, fully configurable dedicated to configuration and control traffic for the camera management. The former requires high bandwidth mechanisms to be able to aggregate several 1 gigabit Ethernet links into one high speed 10 gigabit Ethernet port. The latter is responsible for providing routing components to allow a control and management path for all the elements of the cameras. Hence, a simple, efficient, and flexible routing mechanism has been implemented avoiding complex circuitry that impacts in the system performance. As a consequence, an asymmetric network topology allows high bandwidth communication and, at the same time, a flexible and cost-effective implementation. In our contribution, we analyze the camera requirements and present the proposed architecture. Moreover, we have designed several evaluation tests to demonstrate that our solution fulfills the CTA project needs. Finally, we illustrate the general possibilities of the proposed solution for other data acquisition applications and the most promising futures lines of research are discussed. © *The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.* [DOI: 10.1117/1.JATIS.5.1.014001]

1 Introduction

Nowadays, there are many applications based on distributed data acquisition (DACQ) systems for both scientific facilities and industrial solutions. For the first case, there are important examples, such as experiments inside particle accelerators, for instance A Toroidal LHC ApparatuS (ATLAS) experiment;[1,2] telescopes, such as Large High Altitude Air Shower Observatory (LHAASO),[3] Square Kilometer Array (SKA)[4] and SKA's precursor Karoo Array Telescope (MeerKAT),[5] and applications in health sciences.[6] On the other hand, some industrial applications can be found in the framework of Internet of things or Smart Grids. All of these cases correspond with DACQ systems that share distributed sensor networks scattered over the facility. Moreover, the sensors generate a huge amount of data that must be routed to a central server to be processed. There, the topological design of the network is considered as one of the most difficult issues[7] to take into consideration for the DACQ system and, due to the high number of sensors and their activity, the network connections can suffer from congestion problems. Other contributions[8] propose software solutions to face it, however, the DACQ system bandwidth is reduced to avoid the congestion. A different approach that allows to make use of the full system performance is the development of high bandwidth data aggregation mechanisms in hardware. They are in charge of joining different slow data streams in a fast data output interface, for instance, from 1 Gigabit Ethernet (GbE) ports to one 10 GbE (10G) link. This data flow has a clearly defined direction because its main purpose is to send data from the devices to the core network for processing or storage. In addition to the previous described data path, many sensors also require a minimum configuration and management mechanism to guarantee and verify the proper system behavior. Therefore, some additional routing/switching components must be integrated into the system to enable the interconnection between any two nodes in the network. This represents an asymmetric network topology, where a significant amount of data and bandwidth is required in one direction, and in the other, only a small bandwidth is required but with fully configurable routing options. It enables one to develop a cost-effective solution specially well suited for DACQ applications like those in astrophysics facilities.

Under this context, our research work is focused on telescope array systems. Its design and development are complex tasks with many aspects to be taken into consideration,[9] where some factors, such as energy efficiency,[10] have a key role. Our proposed system is poised to be used in the scientific project Cherenkov telescope array (CTA). CTA will be an observatory for gamma ray astronomy composed of more than 120 telescopes of three different sizes. Our contribution is focused on the small size telescopes (SSTs), which make up the bulk of the deployed telescopes (up to 70). A prototype for these cameras is the compact high energy camera (CHEC). They are responsible for recording images originating from gamma rays penetrating Earth's atmosphere. Each one has several photo sensor modules that capture and digitize the Cherenkov light. These

---

*Address all correspondence to Miguel Jiménez-López, E-mail: klyone@ugr.es

†Current address: German Aerospace Center (DLR), Bremerhaven, Germany

data must be transferred to a central and external server, also known as camera server, to be processed. The main concern is the high bandwidth needed for all data coming from the photo sensors that must be routed through a single 10G port according with the technological evolution of other scientific instruments[11] and the progress of the commercial wired and wireless networks with higher bandwidth every day.[12] Moreover, the camera server can send control packets to set up different elements in the camera and recover status information from them, this justifies the need of a routing mechanism to redirect each packet from the source to a specific module depending on its medium access control (MAC) address. Although these features are included in some time sensitive network devices and high performance switches,[13,14] they are very expensive. On the other hand, our proposed system is implemented using flexible field programmable gate array (FPGA) devices and represents a cost-effective solution. This simplification also reduces design failures associated with the reconfigurable hardware, providing a more dependable solution less prone to failures.

The solution developed is presented in this contribution that has the following structure: the CTA project is introduced and its requirements are briefly explained in Sec. 2; the proposed system for the CTA cameras is described in Sec. 3; the system validation and results are exposed in Sec. 4; and, finally, the main conclusion and the future work are discussed in Secs. 5 and 6, respectively.

2 Cherenkov Telescope Array

CTA[15,16] is an ambitious project, whose goal is to explore the universe in the gamma rays energy region (20 GeV to 300 TeV) and will have a sensitivity an order of magnitude better than current imaging atmospheric Cherenkov technique (IACT) infrastructures[17] for the same energy segment. To accomplish this task, it is divided into two telescopes arrays located at different regions: Paranal (Chile) and La Palma (Spain). The proposed locations for the CTA telescopes have been evaluated using Monte Carlo simulations[18] to check the impact of different factors, such as altitude, night-sky background, and local geomagnetic field. Each telescope array is composed of several telescopes and these are classified into three different types depending on the energy range to be measured: large size telescope, medium size telescope, and SST.

CTA employs the IACT to measure cosmic gamma rays by recording the few nanoseconds long Cherenkov light flashes, emitted in air showers initiated by these gamma rays in the Earth's atmosphere. The direction of the Cherenkov light cone, when recorded with multiple telescopes under different angles, allows for the measurement of the origin of the primary gamma ray in the sky, and the recorded light intensity is a measure of the primary gamma-ray energy. IACT telescopes consist of large tessellated mirrors that focus the Cherenkov light onto the camera with its photo sensors. These sensors are read out by fast electronics, which provide nanosecond sampling of the signals from the Cherenkov light front. Such a time precision can, together with the shape of recorded image, be used to distinguish gamma rays from charged cosmic rays that hit the atmosphere much more numerously, and thus contribute most to the background measured by IACT telescopes. Cosmic-ray air showers are on average broader, less symmetric, and have more irregular timing footprints. Moreover, precise time information results in optimum energy and direction reconstruction performance. Precise timing is therefore mandatory for CTA and the relative timing precision between different cameras is specified to be better than 2 ns on average with less than 1 ns root mean square jitter. The requirement for the absolute time precision with 1 μs is less stringent. Just to compare the magnitude of the synchronization requirements, in the other system such as the Hitomi satellite,[19] a 35 μs is demanded for proper operation, and in SKA telescope, a nanosecond range synchronization is also needed.[20]

Gamma-ray Cherenkov telescope (GCT) is a consortium to provide the SSTs as an in-kind contribution to the CTA observatory. The SSTs are designed to capture the energy range from
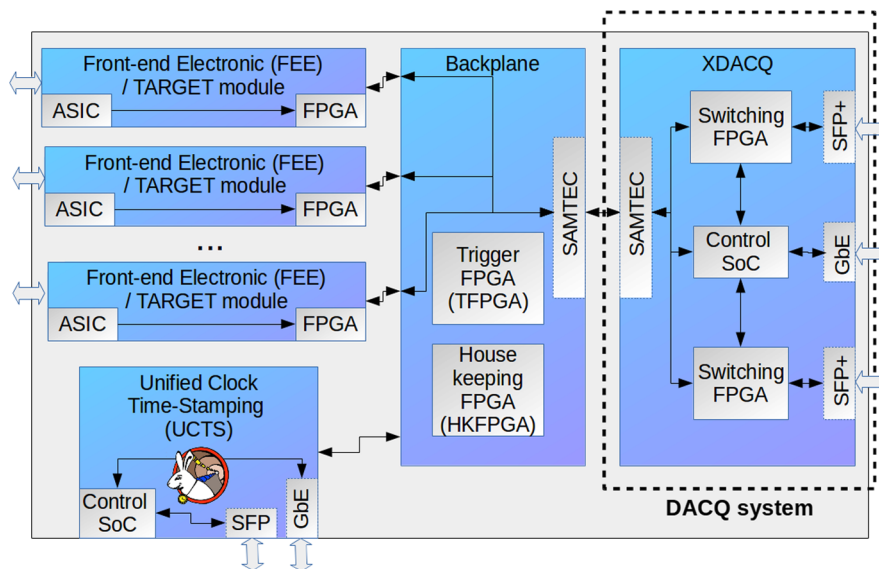
Fig. 1 CTA camera blocks. Schematic overview of the data path and trigger relevant components of the CTA camera. It is composed of TARGET modules, the backplane board, the UCTS board and the XDACQ board. The latter is the platform, where the data aggregation and routing capabilities must be implemented.

about 1 to 300 TeV. As mentioned previously, the prototype for SST is called the CHEC (Fig. 1), which is responsible for measuring and digitizing the sky stimulus and sending this data to a camera server in order to process it. The CHEC[21] is composed of 2048 pixels distributed in 32 front-end electronic (FEE) modules also known as TeV array readout electronics with GSa/s sampling and event trigger (TARGET) modules, the backplane board, two DACQ boards, the uniform clock and trigger time stamping (UCTS) board and auxiliary systems like cooling, calibration, and safety. Each FEE module contains a pixelated photodetector that is responsible for capturing the Cherenkov light information and transmit it to the backplane via the front-end buffers and the TARGET application specific circuits.[22] The backplane is a printed circuit board board that allows the communication between the 32 FEE modules and the DACQ boards. It is in charge of sending trigger patterns to the DACQ boards and triggers the UCTS board for absolute timestamps for the different types of camera triggers. Moreover, UCTS board is responsible for providing time synchronization by means of White Rabbit technology using a dedicated optical fiber network. Due to this, CTA cameras in the array are synchronized with a time accuracy better than 1 ns. In this contribution, we propose a solution to replace the two DACQ boards of CHEC with one single board, called eXtended DACQ (XDACQ). The XDACQ board receives serial data from the FEE modules through the backplane via two SAMTEC connectors and provides 36 GTX serial transceivers at 1 gigabits per second (Gbps). The XDACQ board implements a high-bandwidth data aggregation mechanism to transfer the FEE data and trigger information from the different 1 GbE links in the backplane to a camera server through a high speed interface based on 10G port. It also includes a routing mechanism to transmit packets from the camera server to the FEE modules. Moreover, the XDACQ board takes into consideration redundancy issues due to a second 10G small form-factor pluggable transceptor plus (SFP+) connector. The 10G technology has been required to transmit the high amount of data generate at the CTA telescope, as described in other contributions.[23,24]

3 XDACQ Data Aggregation/Switching System

In this section, the data aggregation and routing system requirements are presented and the proposed solution is described explaining its different components.

3.1 System Requirements

The CHEC requires a very specific data aggregation and routing mechanism to implement a communication between the different 32 FEE modules and the camera server. They generate packets when they detect any interesting event that must be sent to the camera server. These packets contain the sampled waveform of the photo sensors and are Jumbo frames of up to 9000 bytes. The target event rate requirement is 600 to 1200 Hz with 2 to 10 packets per FEE module, and the demanded bandwidth goes from 2.6 up to 5.1 Gbps. Therefore, a 10G port is able to cope with these needs and it can provide more bandwidth if needed in future applications. An important consideration about data bandwidth requirement is that packets from different FEE modules arrive SAMTEC connector at the same time. Under these circumstances, the instantaneous data bandwidth is higher than 10G port capacity. For this reason, buffering

mechanisms must be implemented in the DACQ system in order not to discard any packet.

The packets go into the DACQ system through the 1 GbE connections in the SAMTEC connectors. Then, they are aggregated to reach the common higher bandwidth interface (10G SFP+ port). This path must be ready to receive a high bandwidth transaction from different FEE modules at the same time and the system must have enough memory to implement buffering mechanism. Moreover, the main functions of the camera server are to control all camera subsystems and to collect and store the digitized data coming from the camera photo sensors. The first function is called slow control, which requires routing functionalities in the XDACQ board. The second function is the DACQ that exploits the aggregation system of XDACQ board and imposes high data bandwidth requirements. In addition to that, the aggregation system also implements redundancy mechanisms using its 10G SFP+ ports. During regular operations, only one of them is active, whereas the other one is configured as backup. If the other port becomes active, or if the user manually selects it, their roles are inverted, and the uplink is not interrupted.

Due to the specific CHEC requirements, such as interface connectors, very compact design, the amount of the 1 GbE links and the asymmetric data flows, it is very difficult to find a commercial device to be used for the DACQ and aggregation system. For this reason, we propose the XDACQ board as a specially designed solution to be integrated in the CHEC camera. It has a very powerful hybrid architecture based on a Zynq system-on-chip (SoC) and two FPGA devices. It enables the utilization of hardware/software codesign framework to decide which system features should be faced using hardware components and which need the software flexibility. The data link aggregation mechanism requires high bandwidth and memory buffers that are not easily afforded in software. Then, this subsystem must be implemented using hardware intellectual property (IP) cores to fulfill the CHEC requirements. The other important feature of the system is the routing mechanism. It is responsible for redirecting each packet based on its destination MAC address. These packets are for controlling and monitoring purposes, then the communication at high bandwidth is not required. Therefore, it is mainly implemented in a simple routing table unit (RTU) IP core configured by software.

3.2 Hardware

The XDACQ board is a platform, shown in Fig. 2, developed specifically for the CHEC. This board has a Zynq (xc7z015clg485-1) SoC and two Kintex 7 Ultrascale (xcku040-ffva1156-1-c) FPGA devices. The former includes an advanced RISC machines (ARM) Cortex-A9 dual core processor and a FPGA chip with 74000 logic cells, 3.3 Mb random access memory (RAM), and 4 high speed transceivers. It is responsible for controlling and monitoring the entire DACQ system. The latter are advanced FPGA devices with 530250 logic cells, 21.1 Mb RAM, and 20 high speed transceivers each. They must aggregate all the traffic from the FEE modules (1 GbE interfaces) to the high bandwidth interface (10G interface) and must allow one to route control packets in the opposite direction. Moreover, the XDACQ board has two SFP+ ports, two SAMTEC sockets with 18 1 GbE interfaces each, a control serial peripheral interface (SPI), three universal serial bus connectors for debugging the different FPGA devices, and a control standard 1 GbE port for the Zynq SoC.
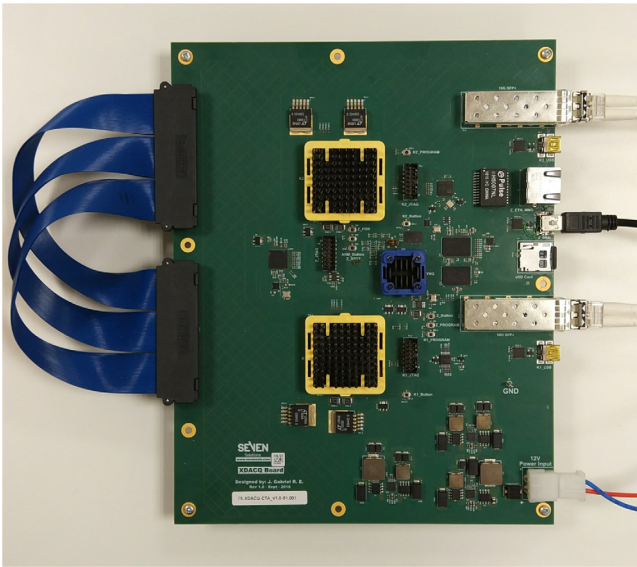
Fig. 2 The XDACQ board. The XDACQ board contains two different kind of FPGA devices: Kintex Ultrascale one and on-chip Zynq one. The former provides buffering enough memory for the data transfer mechanisms. The latter is in charge of controlling and monitoring the entire system and to route the trigger information to the Kintex Ultrascale chips. The SAMTEC connectors to the left (here interconnected for test purposes, see text) are nominally used for the downlink to the further camera electronics.

3.3 Gateware

The XDACQ FPGA firmware (gateware) is schematically presented on the block design of Fig. 3. It is divided into two parts: the Zynq gateware and Kintex Ultrascale gateware.

3.3.1 Zynq gateware

Its design is composed by five subsystems controlled by the on-chip ARM processor (Fig. 4):

- **Kintex program subsystem**: This system is in charge of programming the Kintex Ultrascale FPGA devices with the high bandwidth routing design.
- **Backplane control subsystem**: This system controls the house-keeping and trigger FPGA devices in the backplane through a master SPI interface.
- **Kintex communication subsystem**: This system meets two functions, remote memory mapped access to Kintex FPGA devices from Zynq ARM processor and trigger packet transmission to the aggregation system. A hub core is used to add extra information, also known as headers, together with data in order to distinguish between trigger information and memory mapped read/write access commands. A splitter core is responsible for reading the header information and routes data. An Aurora 8b/10b core is used to instantiate serializers–deserializers
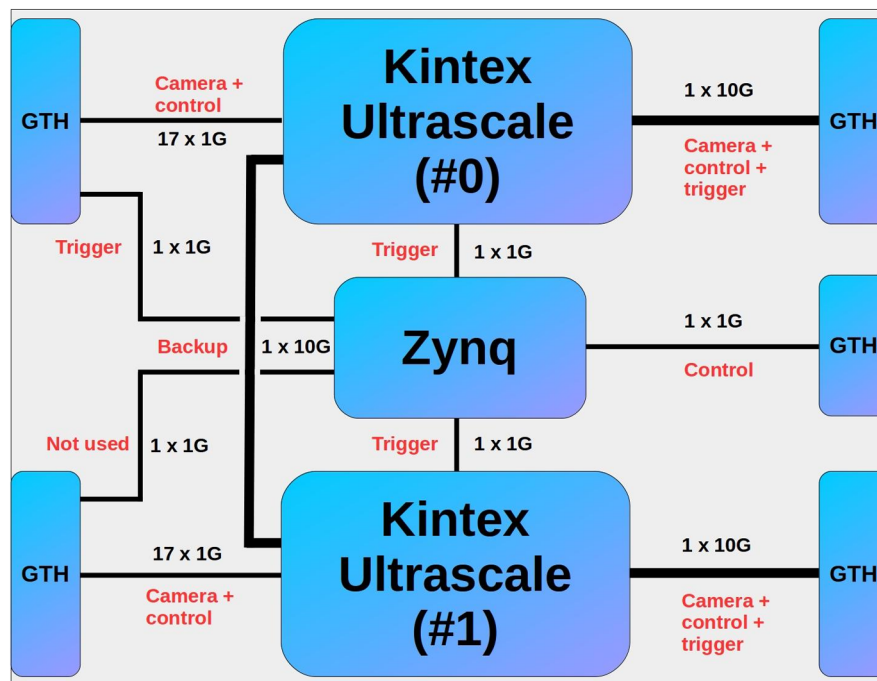
Fig. 3 The XDACQ FPGA architecture. Some routing and aggregation mechanisms must be provided in order to process the different packets. Some of them come from the FEE modules and must be aggregated and redirected to the 10G SFP+ interface. In addition to that, control packets can reach the 10G SFP+ interface and must be routed to a specific FEE module. Moreover, the Zynq device is able to send some control packets to FEE modules. It is possible due to the Aurora 8b/10b protocol that allows one to share a high speed link to send control packets and write directly to the Kintex registers using Advanced eXtensible Interface (AXI) commands. The XDACQ board also includes an advanced backup mechanism between the two Kintex Ultrascale FPGA devices.
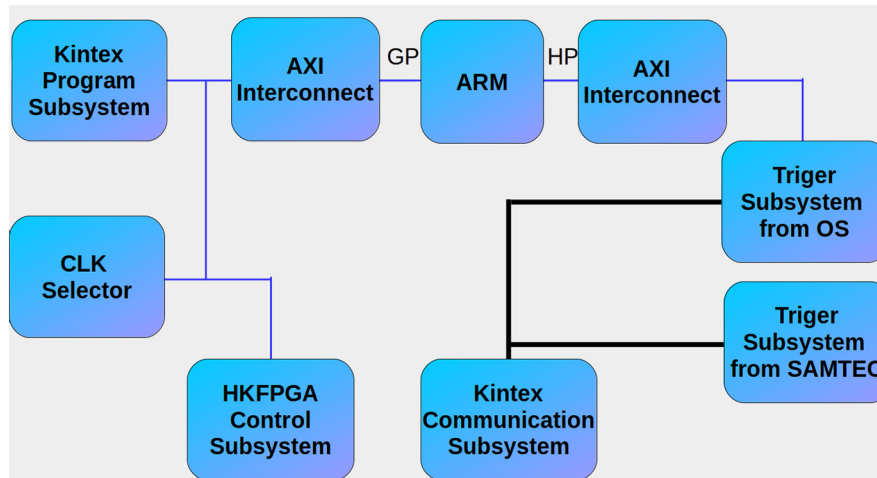
Fig. 4 The Zynq gateware design. The main functionalities are the Kintex Ultrascale devices FPGA programing, reference clock selection mechanism, the trigger redirection capability, and the communication modules with the backplane and the Kintex Ultrascale FPGA devices.

(SerDes) and convert a bus transaction in a packet through the GTP interface.

- **Trigger subsystems**: Both systems receive trigger data, which is sent to both Kintex Ultrascale 10G interfaces. One of them receives data from the operating system and the other one receives data from SAMTEC connector.

- **CLK selector**: General purpose input output to select Kintex Ultrascale devices input clock between internal or external clock.
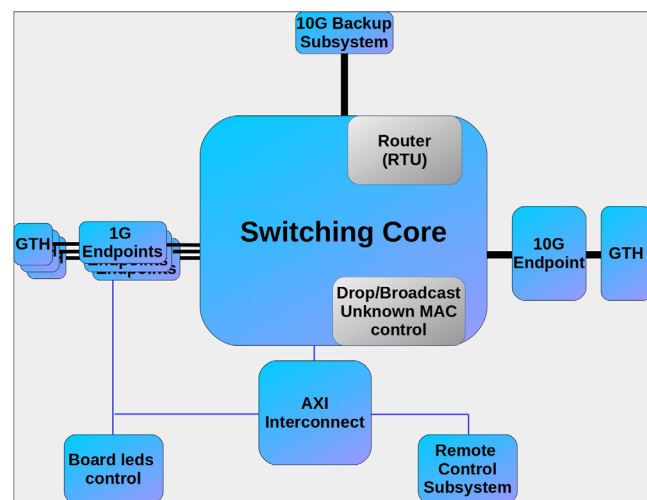
3.3.2 Kintex Ultrascale gateware

It is designed to accomplish two different tasks: link aggregation and packet routing. The Kintex Ultrascale FPGA device has 17 1 GbE links from the SAMTEC socket coming from backplane. These are used to transfer data packets from FEE modules to the camera server and, at the same time, exchange control and status information through the 10G SFP+ (or 10G interface if backup is used).

The main design for the Kintex Ultrascale device is divided into three subsystems:

- **Switching subsystem**: This system is the most important one, and it is responsible for routing and link aggregation mechanisms. Its internal architecture is explained in more detail later.

- **Remote control subsystem**: This system is the counterpart of the Kintex communication subsystem in the Zynq device. It is responsible for receiving trigger information and handling AXI commands from Aurora 8b/10b core. Then, trigger data are routed to the aggregation system to reach the 10G port while the AXI transactions provide memory-mapped access to Kintex Ultrascale registers.

- **10G backup**: This system covers the functionalities related to the backup configuration and communication between both Kintex Ultrascale devices. Its main goal is to provide communication between both Kintex Ultrascale devices when the main SFP+ interface is used and a redundant bidirectional interface to reach

the camera server through the backup SFP+ interface in case of failure of the main one. It can be set in manual mode and the user can decide if packets will be routed to the 10G main SFP+ or the 10G backup SFP+, or in automatic mode, in which packets will be routed to the 10G backup SFP+ interface if the 10G main SFP+ is down.

The main Kintex Ultrascale design is shown in Fig. 5. It contains several 1 GbE subsystem cores, one for each channel of the SAMTEC connectors, two 10G subsystem modules for the SFP + ports, a switching core, and an Aurora 8b/10b component that implements the communication between the Kintex Ultrascale and Zynq FPGA devices. The switching core is a complex module that is responsible for implementing the aggregation and routing capabilities, and it has two data flows. The first one receives data from 17 ports in the SAMTEC socket. Data arrive to 17 first-in first-out (FIFO) queues while an AXI4-Stream

Fig 5 The Kintex Ultrascale gateware design. It contains several gigabit endpoints for the 1 GbE links in the SAMTEC connectors, a switching core that aggregates and routes packets, two 10G endpoints (one of them for backup configuration) and some modules to control the light-emitting diodes (LEDs) in the board and to receive information from the Zynq FPGA device.

(AXIS) switch core gets data from them and sends it to the backup router, which decides to send packets to the 10G SFP+ interface or to the 10G backup interface. The other data flow gets data from 10G SFP+ interfaces to two FIFO queues. Both queues are connected to the router core though an AXIS switch. The main logic block of the router core is the RTU module that uses data registers to storage data words in a three-stage pipeline while the MAC catcher logic finds the destination MAC address into the content addressable memory. Then, it opens one of the possible output channels and appends out-of-band signaling information to allow other components to route the specific packet properly.

3.3.3 Software

The XDACQ software runs in the ARM processor inside the Zynq device. The ARM architecture contains all the elements required to deploy a Linux-based system. The Linux operating system (OS) enables the use of standard applications and eases the software development. Some software modules have been developed and are briefly described:

- **Xilinx Ethernet subsystem configuration**: It is responsible for configuring the necessary registers to enable transmission, reception, and Jumbo frames.
- **Statistic driver**: Linux driver to show the interface statistics through ifconfig shell command.
- **RTU configuration**: Its main goal is to load the routing configuration file into the RTU core when Linux OS starts up.
- **Backup configuration**: It is in charge of enabling/disabling backup and set automatic/manual mode.
- **Clock input configuration**: It enables the clock selection between internal and external sources.

In addition to the custom software modules, Linux common services, such as Secure SHell (SSH), file transfer protocol (FTP), and even a web server have been integrated in the OS environment. However, external access to these services is limited to the copper Ethernet interface.

4 System Validation and Results

In this section, we provide some tests to demonstrate that the developed system fulfills the CHEC requirements. The first part shows the resource utilization meanwhile the second one evaluates the system performance.
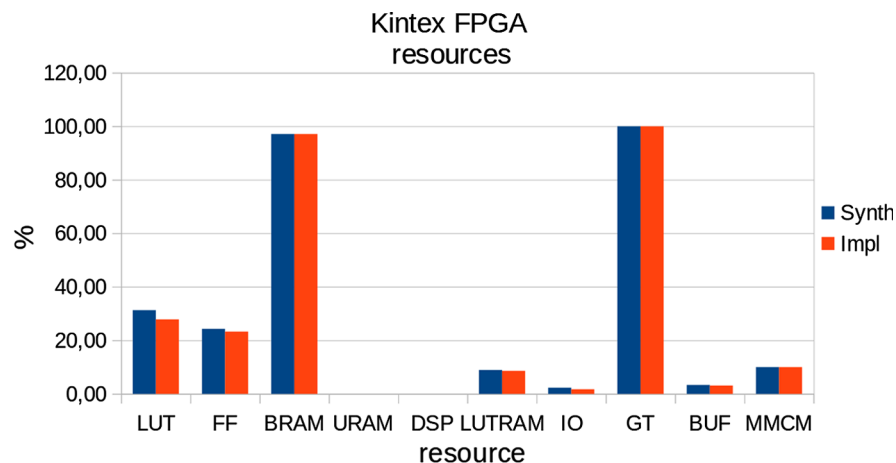
4.1 Resource Utilization

The system requires two different implementations for the Kintex Ultrascale and the Zynq FPGA devices. Figure 6 presents the resource utilization for the Kintex Ultrascale FPGA devices. It demands several block RAM (BRAM) blocks to generate the FIFO components for the data aggregation and routing cores. All the gigabit transceivers are also used in this design. However, the overall utilization is not so high because there are many available look-up table (LUT), flip flop (FF) and LUT as RAM (LUTRAM) blocks that are the basic building components for the programmable logic devices.

The Zynq FPGA device is responsible for controlling and monitoring the XDACQ board and therefore, it presents different resource needs than the Kintex Ultrascale's ones (shown in Fig. 7). The most used resources are the gigabit transceivers for the high speed external communication and the phase-locked loop/mixed-mode clock manager blocks for the clock generation. Nevertheless, some free logic elements are available for future developments.

4.2 System Performance Evaluation

The system evaluation is a crucial, yet not trivial task whose main goal is to obtain the system bandwidth and latency. It requires one to test all interfaces of the XDACQ board: the 1 GbE ones in the SAMTEC connectors and the 10G SFP+ ports. To accomplish this task, there are different ways depending on the equipment that we use. The first option to perform this evaluation is the conventional personal computer (PC) utilization. The main issue is that these equipments normally do not have a high number of interfaces and it makes hard the exhausting test for all the XDACQ interfaces. The second choice is to make use of a specific switch or router. However, this alternative is very expensive as it has to fulfill the CTA
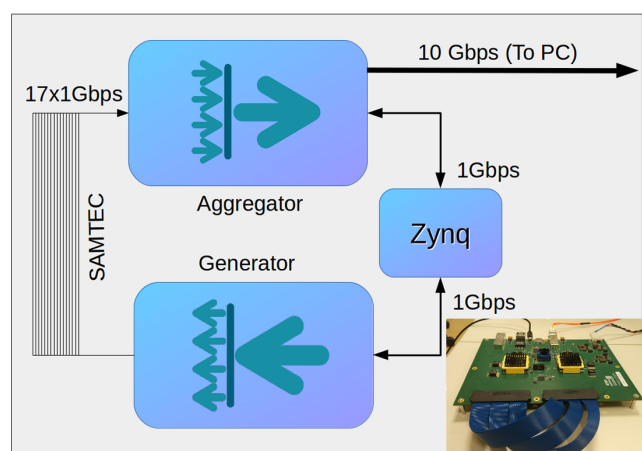

Fig. 6 Kintex FPGA resource utilization report. It shows that all the gigabit transceivers are used and practically all the BRAM available for the FIFO components of the aggregation and routing implementation. The high utilization of the FIFO components is justified due to the buffering necessity of the DACQ system.
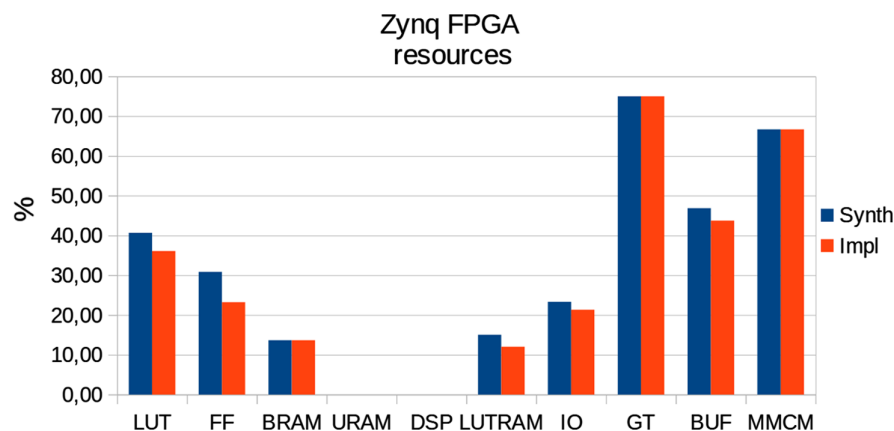
Fig. 7 Zynq FPGA resource utilization report. This design fits very well in the current platform and there are available resources to implement advanced features if needed.

physical/interconnection requirements. To solve these inconveniences, we have implemented a traffic generator system using one of the Kintex Ultrascale in the XDACQ board. This system imitates the behavior of the FEE modules sending packets bursts from each interface at the same time. In addition, it requires the utilization of a crossed SAMTEC cable to establish communication between the different Kintex Ultrascale FPGA devices. The internal architecture of the generator uses the AXIS traffic generator module[25] configurable from an AXI4 Slave and custom IP core, which calculates the checksum of the data in a 8-bit word and the number of packets and bytes generated. In order to generate high-bandwidth data using the minimum resources in the FPGA device, only one AXIS traffic generator IP core is used but data are replicated in each 1 GbE interface using an AXIS broadcaster. For each packet generated, 17 are sent by the FPGA device. This setup is shown in Fig. 8 that includes a diagram and a picture of the board interconnected via the SAMTEC cable. In the PC side, data are received through an Endace DAG 10X2-S network controller[26] and measured with nload tool.[27]

The proposed system has been evaluated in different scenarios to ensure that it fulfills the CHEC requirements. The first test case evaluates the system behavior when several 1 GbE interfaces are activated to transmit using the full Gigabit capacity, as shown in Fig. 9. In this case, the independent variable is the number of interfaces transmitting at the same time and the test has been performed for different packet sizes: 1500, 4500, and 9000 bytes. The results demonstrate that the DACQ system is able to cope with the 92.9% of the 10G total capacity.
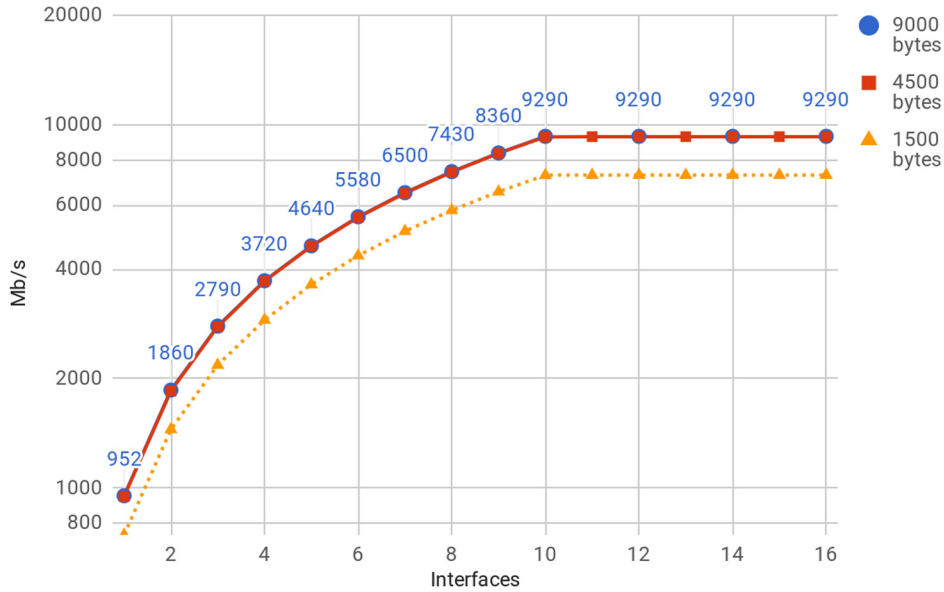
The second test scenario measures the data bandwidth limits using the 16 1 GbE links at the same time. Under this context, the independent variable is the data bandwidth per interface and its results are summarized in Fig. 10. They evidence a perfect fit between the system performance and the theoretical one. Moreover, the output interface is limited for the 10G bandwidth and it is not possible to obtain a higher performance for a long time period. If this corner condition is not considered, the overall bandwidth of the system can be reduced dramatically. For this reason, the aggregation system implements a FIFO control mechanism that allows system bandwidth to remain constant when this condition is exceeded, as shown in Figs. 9 and 10.

In addition to the performance tests, we have measured the routing system latency and the propagation delay of its main component: the RTU module.

The RTU module has been tested together with the aggregation mechanism in high bandwidth conditions to evaluate the control path latency and the isolation between aggregation and routing paths. Figure 11 shows that the RTU module does not introduce any penalty in the system due to its deterministic latency. Furthermore, the RTU module does not use any memory element to hinder the packet traffic, therefore it does not limit the maximum amount of bandwidth for the routing path. The latency of the RTU module does not depend by the number of active 1 GbE links, as can be seen in Fig. 12. The results demonstrate that the aggregation and routing paths are properly isolated and the control flow is not affected for the high bandwidth activity in the network.

In addition to the laboratory tests previously described, the XDACQ board has been already successfully integrated in the CHEC prototype (Fig. 13); prior to this, several integration tests were performed at the Deutsches Elektronen-Synchrotron (DESY) Institute[28] in Zeuthen, Germany, and at the Max Planck Institute for Nuclear Physics (MPIK)[29] in Heidelberg, Germany. The tests went from the verification of the basic functionality of the board such as packet routing and SPI communication, to high-level stress tests, such as the repeated

Fig. 8 XDACQ board test setup. A Kintex Ultrascale FPGA device is configured to generate several packets through the SAMTEC connection. On the other side, the second Kintex Ultrascale FPGA device receives and pass them through the switching connector to reach the 10G port.

Fig. 9 Bandwidth experiment for aggregation system. System bandwidth with different packet sizes and several interfaces transmitting at the same time. It shows that the system is able to cope with the maximum 10G bandwidth (10 interfaces at the same time).
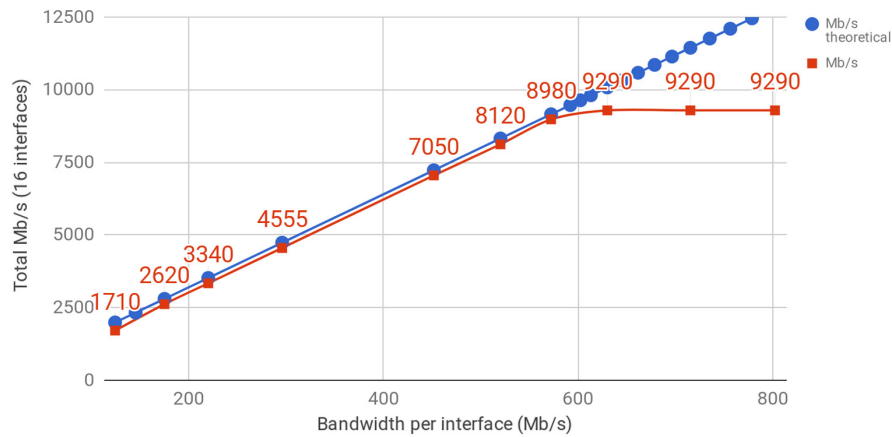

Fig. 10 Comparison of system performance with theoretical one. For bandwidth larger than 600 Mbps per interface, the system is able to use the 92.9% of the 10G link capacity.
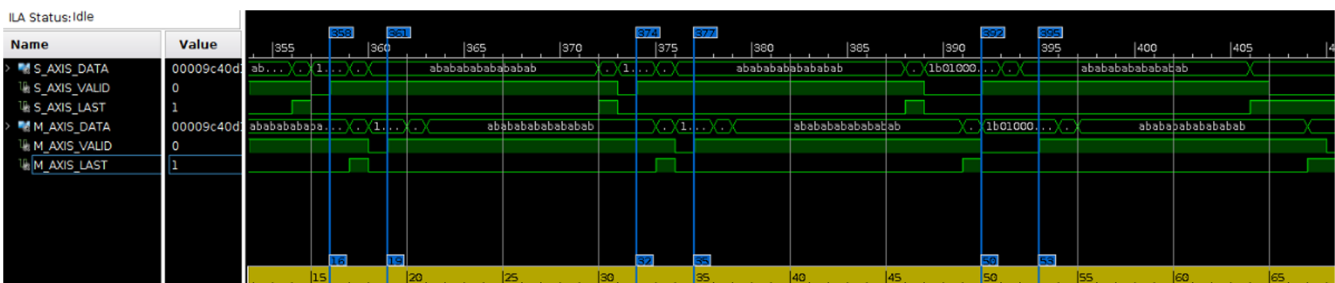
Fig. 11 RTU latency test. This picture shows the real behavior of the RTU IP core. It has been obtained by means of Vivado logic analyzer software, which is able to introduce logic probes in the FPGA design. Once introduced probes in the RTU module, the routing system is enabled and some burst of three packets are sent from PC to the 10G port. Regarding the picture, each pair of markers in the picture represents the time between the packet ingress to and the packet egress. The RTU module transfers a packet to its output interface when the packet MAC address is used to determine the final destination for a specific packet. Therefore, the picture demonstrate that the RTU module always presents a deterministic and fixed latency of three cycles.
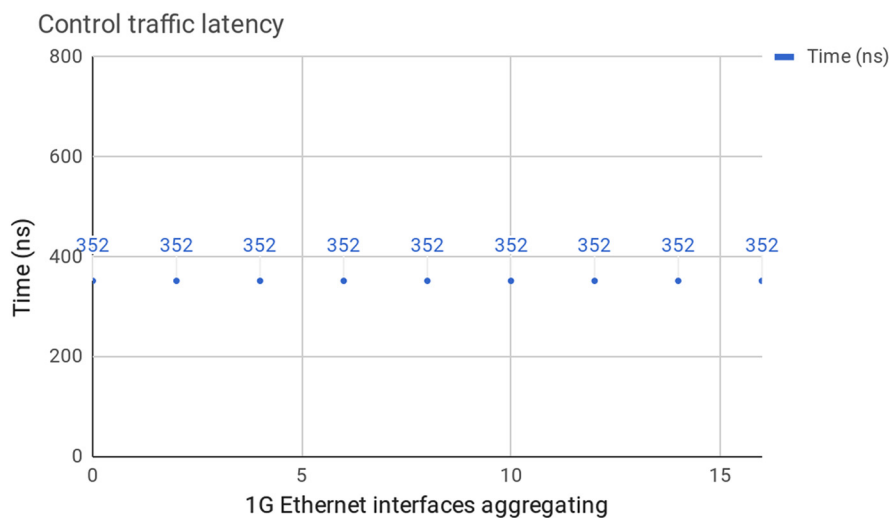
Fig. 12 Latency experiment for the routing system. Latency test for the control flow through the RTU module while the aggregation mechanism is active. The figure illustrates the latency behavior for the control packets that is not affected by the aggregation logic in the system. It demonstrates that the data packet flow and the control one are properly isolated.
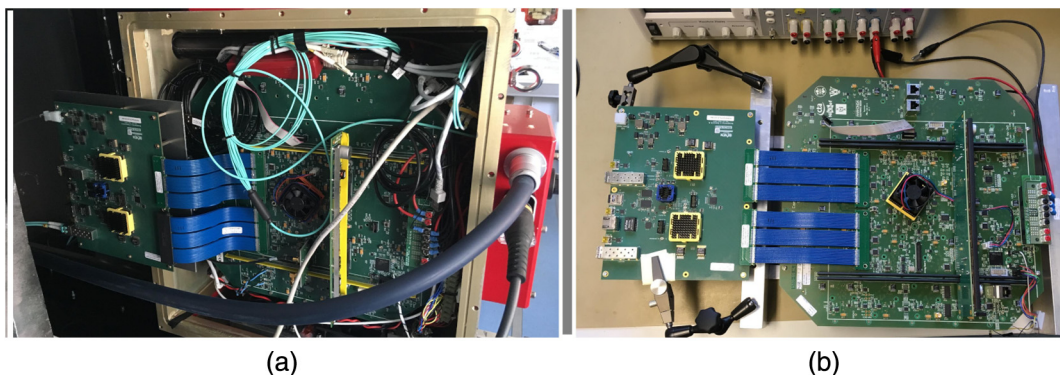
Fig. 13 DESY camera test setup. (a) Integration and (b) validation of the XDACQ board on the CHEC camera.

simulation of observation runs, with frequencies up to 2.5 times (1500 Hz) the required one (total bandwidth up to 6.3 Gbps), involving the exchange of hundreds of thousands of control packets and the collection of several TBs worth of data, with no errors.

5 Conclusion

In this contribution, we have shown how an asymmetric network can be used as a cost-effective and flexible solution for the DACQ systems. A high bandwidth upstream traffic is transmitted using a static routing scheme while a flexible, fully programmable, low bandwidth management traffic can be added to this network topology. As a consequence, the network elements and topology used for DACQ system can be easily deployed. As challenging target example, we have worked on the CTA project and particularly for the XDACQ platform. It includes this hardware architecture combining two Kintex Ultrascale FPGA devices and a SoC communicated by means of a high speed bus based on Aurora 8b/10b technology. The Kintex Ultrascale FPGA device contains many memory elements and logic block that enable the high bandwidth link aggregation and routing mechanisms. The Zynq SoC is in charge of control and complex software tasks, such as routing table maintenance, FPGA programming, and diagnostics among others. Moreover, the use of Linux OS allows one to present a friendly and standard interfaces and toolset to the users. Some examples of these kind of applications are SSH sessions and FTP service. However, additional applications can be easily installed due to the existence of Linux OS.

The proposed DACQ system has been tested and measured to get the maximum bandwidth that can be used and, therefore, calculate the system performance. The results described in the previous section argue that the system is able to work properly up to 9.29 Gbps for the aggregation components. Moreover, it shows that the RTU module presents a deterministic latency avoiding any penalty due to the operation of this component. The aggregation and routing paths are properly isolated and the control path is not affected by the high bandwidth network conditions in the aggregation side.

The results presented here allow us to conclude that the proposed solution is able to reach the demanded bandwidth fulfilling the CTA requirements with a small resources consumption and with a simple and predictive network routing architecture.

6 Future Work

Finally, we propose the following future work lines as the most promising and remarkable ones:

- Develop traffic control mechanisms to avoid the packet loss when a higher bandwidth more than 10 Gbps is required. This is useful to provide alarm signals to the software for monitoring purposes.

- Improve the Aurora 8b/10b channel between Zynq SoC and Kintex Ultrascale FPGA devices to allow a full duplex communication. It would enable a complete communication between the camera server and ARM processor in the Zynq SoC. It would be interesting to establish SSH sessions through the 10G port.

- Extend the current aggregation model to build an asymmetric system with static routing and fully programmable aggregation mechanisms.

- Update the current design to deal with higher bandwidth interfaces, such as 25 GbE ones.

References

1. G. Jereczek et al., "A lossless network for data acquisition," *IEEE Trans. Nucl. Sci.* **64**, 1238–1247 (2017).
2. G. Jereczek, "Software switching for high throughput data acquisition networks," PhD Thesis (2017).
3. Q. Du, G. Gong, and W. Pan, "A packet-based precise timing and synchronous DAQ network for the LHAASO project," *Nucl. Instrum. Methods Phys. Res. Sect. A* **732**, 488–492 (2013).
4. S. K. A. Consortium, "Square kilometer array website," https://www.skatelescope.org/ (August 2018).
5. T. B. Gibbon et al., "Fiber-to-the-telescope: MeerKAT, the South African precursor to square kilometre telescope array," *J. Astron. Telesc. Instrum. Syst.* **1**, 028001 (2015).
6. C. Zorraquino et al., "Asymmetric data acquisition system for an endoscopic PET-US detector," in *19th IEEE-NPSS Real Time Conf.*, pp. 1–3 (2014).
7. N. Shukla et al., "Design of computer network topologies: a vroom inspired psychoclonal algorithm," *Appl. Math. Modell.* **37**(3), 888–902 (2013).
8. T. Uchida et al., "New communication network protocol for a data acquisition system," *IEEE Trans. Nucl. Sci.* **53**, 286–292 (2006).
9. B. Karlow et al., "Tradespace investigation of strategic design factors for large space telescopes," *J. Astron. Telesc. Instrum. Syst.* **1**, 027003 (2015).
10. J. Song et al., "Energy efficiency evaluation of tree-topology 10 gigabit Ethernet passive optical network and ring-topology time- and wavelength-division-multiplexed passive optical network," *Opt. Eng.* **54**, 090502 (2015).
11. A. Recnik et al., "An efficient real-time data pipeline for the CHIME pathfinder radio telescope x-engine," in *2015 IEEE 26th Int. Conf. Appl.-specific Syst., Archit. Process.*, IEEE (2015).
12. P. C. Jain, "Recent trends in next generation terabit Ethernet and gigabit wireless local area network," in *Int. Conf. Signal Processing and Communication (ICSC)*, pp. 106–110 (2016).
13. W. Steiner, S. S. Craciunas, and R. S. Oliver, "Traffic planning for time-sensitive communication," *IEEE Commun. Stand. Mag.* **2**, 42–47 (2018).
14. H. J. Chao and B. Liu, *High Performance Switches and Routers*, John Wiley and Sons, Inc., Hoboken, New Jersey (2006).
15. C. T. A. Consortium, "Cherenkov telescope array website," http://www.cta-observatory.org (August 2018).
16. B. Acharya et al., "Introducing the CTA concept," *Astropart. Phys.* **43**, 3–18 (2013).Seeing the High-Energy Universe with the Cherenkov Telescope Array—The Science Explored with the CTA.
17. C. Balázs et al., "Sensitivity of the Cherenkov telescope array to the detection of a dark matter signal in comparison to direct detection and collider experiments," *Phys. Rev. D* **96**, 083002 (2017).
18. T. Hassan et al., "Monte Carlo performance studies for the site selection of the Cherenkov telescope array," *Astropart. Phys.* **93**, 76–85 (2017).
19. Y. Terada et al., "Time assignment system and its performance aboard the Hitomi satellite," *J. Astron. Telesc. Instrum. Syst.* **4**, 011206 (2017).
20. M. Jiménez-López et al., "A fully programmable white-rabbit node for the SKA telescope PPS distribution system," *IEEE Trans. Instrum. Meas.* PP(99), 1–10 (2018).
21. J. Lapington et al., "The GCT camera for the Cherenkov telescope array," *Nucl. Instrum. Methods Phys. Res. Sect. A* **876**, 1–4 (2017). The 9th international workshop on Ring Imaging Cherenkov Detectors (RICH2016).
22. D. Gascon et al., "Reconfigurable ASIC for a low level trigger system in Cherenkov telescope cameras," *J. Instrum.* **11**, P11017 (2016).
23. R. Rajda et al., "DigiCam—fully digital compact read-out and trigger electronics for the SST-1m telescope proposed for the Cherenkov telescope array," in *Proc. 34th Int. Cosmic Ray Conf. (ICRC2015)*, The Hague, The Netherlands, arXiv:1508.06082v1 (2015).
24. D. Hoffmann et al., "Prototyping a 10 gigabit-Ethernet event-builder for the CTA camera server," *J. Phys. Conf. Ser.* **396**, 012024 (2012).
25. Xilinx, "Xilinx AXI traffic generator," https://www.xilinx.com/products/intellectual-property/axi_tg.html (August 2018).
26. Endace, "Endace dag 10x2-s datasheet," https://www.endace.com/introducing-dag.pdf (August 2018).
27. L. Man Pages, "Nload(1)—Linux man page," https://linux.die.net/man/1/nload (August 2018).
28. DESY, "Deutsches elektronen-synchrotron," http://www.desy.de (August 2018).
29. MPIK, "Max Planck Institute for Nuclear Physics," https://www.mpi-hd.mpg.de/mpi/en/start (November 2018).

Miguel Jiménez-López received his Msc degree in computer science from the University of Granada, Spain, in 2013. He is finishing a PhD degree in computer science in the Department of Computer and Technology of University of Granada, Spain. His main research interests are high accurate synchronization technologies, especially in high data bandwidth systems. From April to July 2017, he was actively collaborating in the CTA project due to an international research stay at NIKHEF, Amsterdam, The Netherlands.

Jorge Manuel Machado-Cano received his BSc degree in computer science from University of Granada, Spain, in 2017. He is working as FPGA engineer at Seven Solutions. His main interests are related with high bandwidth system capabilities and data processing on FPGA based embedded systems.

Manuel Rodríguez-Álvarez received his BSc degree in electronics in 1986 and his PhD degree in physics in 2002 both from the University of Granada, Spain. He is currently an associate professor at the Department of Computer Architecture and Technology of University of Granada. His research interests include the dissemination of precise timing over optical fiber networks, and he collaborates with research facilities as SKA working on subnanosecond time transfer solutions based on White Rabbit.

Maurice Stephan received his diploma in physics in 2009 and a PhD in science in 2014 both from RWTH Aachen University. His research interest focuses on instrumentation for imaging applications and data processing. Starting in 2015, he was involved with the development of the GCT cameras at the University of Amsterdam and NIKHEF, Amsterdam. In 2018, he joined the German Aerospace Center (DLR), where he now develops instruments and methods for the protection of maritime infrastructures.

Gianluca Giavitto received his PhD degree in physics from the Universitat Autonoma de Barcelona in 2013. His main research interests are VHE gamma-ray emission from pulsars and development of cameras for imaging atmospheric Cherenkov telescopes. His work led to the detection by ground-based instruments of VHE gamma-ray pulsations from the Crab and Vela pulsars. He has also collaborated on MAGIC and H.E.S.S. experiments. He is currently working on the development of the CHEC camera at DESY, Germany.

David Berge received his master in physics in 2002 from the University of Berlin and a PhD in science in 2006 from Max-Planck-Institute for Nuclear Physics. He is leading the gamma-ray group at the DESY site in Zeuthen. His research is focused on cosmic particle accelerators and the search for dark matter. In 2017, he accepted an offer for a joint professorship for particle and astroparticle physics at DESY in Zeuthen and the University of Berlin.

Javier Díaz received his MS degree in electronics engineering in 2002 and a PhD in electronics in 2006 both from the University of Granada. His main interests are related with high performance image processing architectures, safety-critical systems, highly accurate time synchronization and frequency distribution techniques. Currently, he works as a university professor and collaborates with research facilities as CERN, IFMIF-EVEDA, CTA, or SKA working on subnanosecond time transfer solutions based on White Rabbit technology.