The Selective Read-Out Processor for the CMS Electromagnetic Calorimeter

Nuno Almeida, Philippe Busson, Jean-Louis Faure, Olivier Gachelin, Philippe Gras,
Irakli Mandjavidze, Michel Mur, Joao Varela

Abstract—This paper describes the selective read-out processor (SRP) proposed for the electromagnetic calorimeter (ECAL) of the CMS experiment at LHC (CERN). The aim is to reduce raw ECAL data to a level acceptable by the CMS DAQ system. For each positive level 1 trigger, the SRP is guided by trigger primitive generation electronic to identify ECAL regions with energy deposition satisfying certain programmable criteria. It then directs the ECAL read-out electronic to apply predefined zero suppression levels to the crystal data, depending whether the crystals fall within these regions or not. About 200 of high speed 1.6 Gbit/s I/O channels, asynchronous operation at up to 100 kHz level 1 trigger rate, a stringent real-time requirement of 5 µs latency and flexibility in choice of selection algorithms are the main challenges of the SRP application. The architecture adopted for the SRP is based on modern parallel optic pluggable modules and high density FPGA devices with embedded processors and multi-gigabit transceivers. Implementation studies to validate proposed solutions are presented. The performance of envisaged selection algorithms is investigated with the CMS detector simulation software. The robustness of optical communication channels is estimated via direct measurements and calculations. The feasibility to perform data reduction operations within the allocated timing budget is verified by running a representative SRP firmware on a development board with a Xilinx Virtex2Pro FPGA device.

I. INTRODUCTION

WITH its some 80,000 lead tungstate (PbWO$_4$) crystals the CMS electromagnetic calorimeter (ECAL) is designed to play an essential role in exploiting the physics potential of the LHC [1]. It is divided into four partitions: left and right half-barrels and endcaps (Fig. 1). Each half-barrel contains 18 super modules (SM) of 1700 crystals and each endcap contains two Dee-shaped sections of 3908 crystals.

The ECAL is one of the main CMS trigger detectors. The

synchronous 40 MHz pipelined level 1 trigger acts on calorimeter trigger primitives – essentially transverse energies deposited in trigger towers (TT). There are 4032 TTs: 1224 per half-barrel and 792 per endcap.
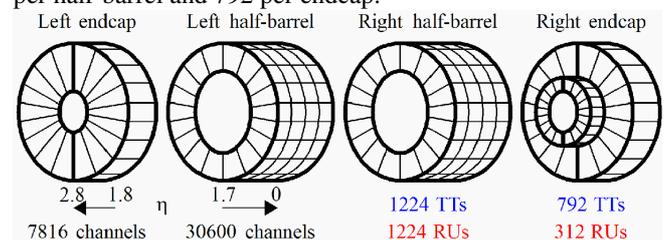

Fig. 1.   Four partitions of the CMS electromagnetic calorimeter

Barrel TTs are formed from regular 5x5 square arrangements of crystals, which in fact matches the read-out segmentation of the detector (Fig. 2). Endcap TTs have less regular shape. For each bunch crossing computation of the TT energies are initiated in the ECAL front-ends and finalized in the Trigger Concentrator Cards (TCC) within the off-detector electronics (ODE) housed in VME crates and racks in the CMS underground service cavern [2]. The TCCs transmit the computed primitives to the level 1 trigger system. There are 108 TCCs: 18 per half-barrel and 36 per endcap.
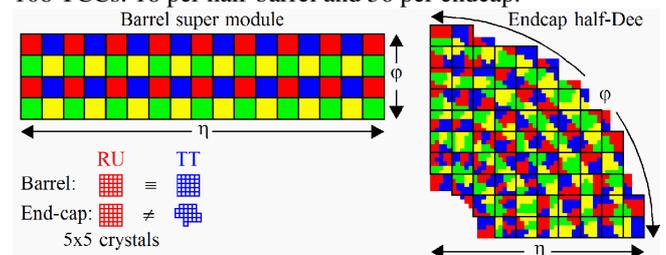

Fig. 2.   Read-out units and trigger towers in half-barrels and endcaps

Organization of the detector read-out follows mechanical arrangement of crystals in square 5x5 read-out units (RU). There are 3072 RUs: 1224 per half-barrel and 312 per endcap. For each positive level 1 trigger, digitized and formatted data of all 25 crystals in a RU is sent by front-end electronics to the corresponding ODE Data Concentrator Card (DCC). The DCC is responsible for reception of level 1 accepted ECAL data from up to 68 RUs, for checking of their integrity, for merging them into an event fragment and for transmission of the latter to the CMS high level triggers (HLT) and data acquisition (DAQ) system. There are 54 DCCs: 18 per half-barrel and 9 per endcap.

Due to the slow nature of calorimeter signals, 10 consecutive time samples per channel need to be integrated for

accurate estimation of the energy deposited in a crystal. The totality of calorimeter data received by DCCs for level 1 accepted events amounts ~1.5 Mbyte including fixed ECAL specific protocol overhead. This by far exceeds the 100 Kbyte part of the total CMS event size allocated to the calorimeter. In addition, the CMS event building system requires uniform distribution of ECAL event fragment sizes around an average value of 2 Kbyte. Under the full read-out hypothesis, when all crystal data are systematically transferred to the HLT and DAQ system, the DCC event fragments are of the order of 40 Kbyte. Also, in this case, total bandwidth of ECAL data (up to 150 Gbyte/s at 100 kHz level 1 trigger rate) surpasses the 100 Gbyte/s event builder throughput. Therefore, calorimeter data to be passed to the higher event selection levels have to be reduced by a factor of almost 20. Studies in [3] and [4] have shown that a crystal-by-crystal zero suppression (ZS) technique cannot be used as it would significantly degrade energy resolution and introduce important non-linearity in the energy scale. To achieve the necessary data reduction factor without affecting physics performance of the ECAL, a selective read-out approach has been proposed.

II. SELECTIVE READ-OUT PROCESSOR

Selective read-out implies determining for level 1 accepted events ECAL areas of interests with energy depositions that satisfy certain programmable criteria. For crystals that form the identified zones of interests all energy samples are kept for further levels of event selection. To achieve the necessary data reduction factor, the rest of the crystals are optionally read out, if their energy is above a certain zero suppression threshold.

The envisaged selective read-out algorithm [4], [5] defines ECAL areas of interests evaluating TT transverse energies. The areas are larger than individual TTs and knowledge of energies in groups of neighboring TTs is needed. Also, the areas of interests may span across the calorimeter regions served by TCCs and DCCs. A dedicated hardware device, called Selective Read-out Processor (SRP), is the part of the ECAL off-detector electronics, which on event-by-event basis collects from TCCs a complete map of energy depositions in TTs, executes the selection algorithm and instructs DCCs on read-out actions to be taken per individual RUs (Fig. 3).

A. Functional Description

For each level 1 trigger accept signal, the SRP receives from 108 TCCs classification flags for all 4032 calorimeter TTs. These flags, 3 bits per TT, indicate the threshold level that has been passed by the energy deposited within a given TT. The SRP scans the calorimeter in η and φ directions and classifies every TT as "suppressed" (energy is below a low threshold), "single" (energy is in-between the low and a high thresholds) and "central" (energy is above the high threshold). In the presence of a central TT, 8 or 24 nearest trigger towers in square 3x3 or 5x5 regions are classified as "neighbors".

The SRP then produces selective read-out flags for all 3072 RUs. These flags, 3 bits per RU, indicate the type of TTs that are formed by the crystals of a given RU. The SRP delivers the

selective read-out (SR) flags to 54 DCCs. The selective-read out flags arrive to DCCs before the corresponding crystal data comes out of a fixed pipeline delay, so that the DCCs can perform on-line data reduction. For crystals that form "central", "neighbor" or "single" trigger towers all 10 energy samples are kept. Crystals contributing to "suppressed" TTs are ignored or are read out with a high ZS threshold.
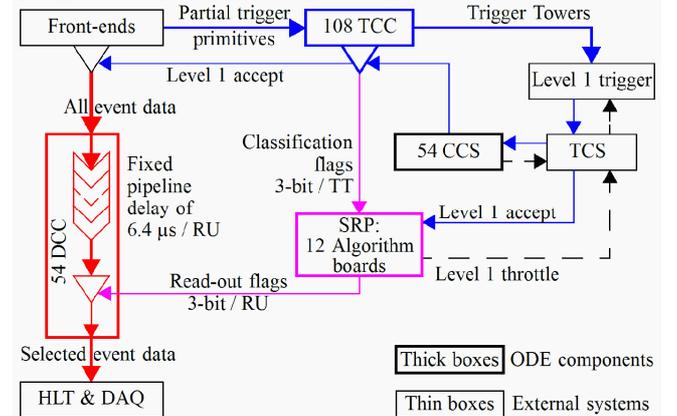

Fig. 3. SRP within the ECAL Off-Detector Electronics

The data reduction efficiency of this kind of selective read-out algorithms has been studied in [5]. The CMS event reconstruction software has been used to generate a representative set of jet events with transverse momentum in the 50-100 GeV range. Calorimeter level 1 trigger at great extent is generated by single electrons originating from such jets. High luminosity conditions were reproduced superimposing 17 minimum bias events, on average.
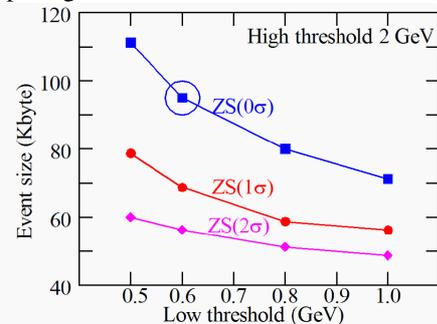

Fig. 4. ECAL average data volumes obtained with the selective read-out

Fig. 4 shows ECAL average data volumes obtained with a variant of the generic SR algorithm: RUs that form Suppressed TTs are ignored, while crystal data in RUs that participate to Single, Neighbor and Center TTs are further reduced by applying ZS with a given threshold. The feasibility to fit an average ECAL event to the 100 Kbyte quota with a uniform distribution of event fragment sizes has been shown. In these conditions, neither noticeable degradation in energy resolution, nor perceptible non-linearity in energy scale have been observed [4]. Other types of SR algorithms are under study.

The full precision ECAL data, retained by SR algorithms, is always complemented with coarse grain data of energy deposited in the calorimeter. For events accepted by level 1 trigger, TCCs send computed trigger primitives to DCCs, which are included in event fragments. These data contribute to measurements of jet energies and missing transverse energy.

B. SRP Architecture

The SRP is an asynchronous device that operates at level 1 trigger frequency of up to 100 kHz. An absolute maximum of its timing budget to produce and deliver selective read-out flags equals to 6.4 µs. This is fixed by the 256-deep input pipelines in DCCs, where the calorimeter front-end data is delayed while the SRP performs trigger tower classification and derives selective read-out flags (Fig. 3).

According to the adopted architecture, the SRP is housed in a single 6U VME64x crate. It is composed of 12 conceptually identical Algorithm Boards (AB). The crate will also contain a crate controller – any standard VME-PC interface adopted by the CMS collaboration for slow control and monitoring.

For accepted events, the ABs examine in parallel energy depositions in distinct calorimeter areas. Mapping of TCCs, DCC and ABs is shown on Fig. 5. In the half-barrels, TCC and DCC mapping follows their mechanical segmentation in super modules: a TCC-DCC pair serves one SM. Three ABs cover a half-barrel, each receiving TT classification flags from 6 TCCs and delivering SR flags to the corresponding 6 DCCs. In the endcaps, four TCCs and a DCC serve approximately 40° φ sectors. Three ABs cover entire endcap, each serving 3 such sectors. An endcap AB receives TT classification flags from 12 TCCs and delivers SR flags to 3 DCCs.
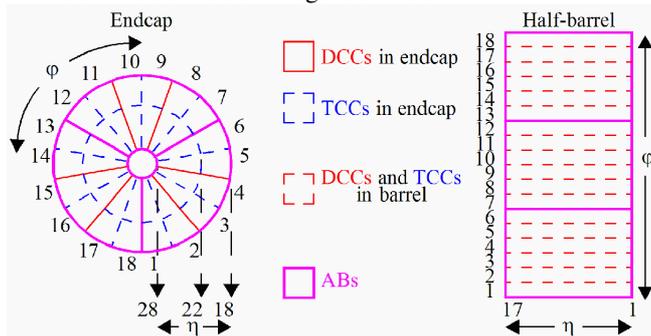

Fig. 5.   Mapping of TCCs, DCCs and ABs in half-barrels and endcaps

Thus, each TCC sends TT classification flags to only one AB and each DCC receives SR flags from only one AB. Each AB receives TT classification flags from at most 12 TCCs (endcaps) and sends SR flags to at most 6 DCCs (half-barrels). In total there are 108 TCC to SRP and 54 SRP to DCC unidirectional connections. In addition, to be able to classify trigger towers on their edges, adjacent ABs exchange classification flags of their frontier TTs. An endcap AB sends (receives) TT data to (from) its 2 fellow endcap ABs and 3 neighboring barrel ABs. A barrel AB exchanges TT data with 8 adjacent ABs: 5 in the half-barrels and 3 in an endcap. All these sum up to 39 bidirectional inter-AB connections.

C. Algorithm Board

The considerable number of SRP I/O channels (~200), the asynchronous operation at up to 100 kHz level 1 trigger accept rate, the stringent maximum latency of 6.4 µs and needed flexibility to change to some extend selection algorithms require bringing together the latest advances of modern FPGA devices and optical communication technologies.

The Algorithm Boards, currently under development, are single-slot 6U VME64x compliant cards (Fig. 6). They contain a high integration 2vp70 device from the Xilinx Virtex2Pro FPGA family, which implements the core functionality of ABs. Several interfaces with components of the ECAL OD electronics and external systems allow for configuration, initialization and monitoring of the cards and ensure the necessary connectivity for their normal operation. These are the communication channels with the TCCs and DCCs, interfaces with the CMS trigger control system (TCS) [6], a VME slave interface, a configuration and test interface and some auxiliary interfaces such as RS232 console and Ethernet link for the PowerPC processor embedded in the FPGA.
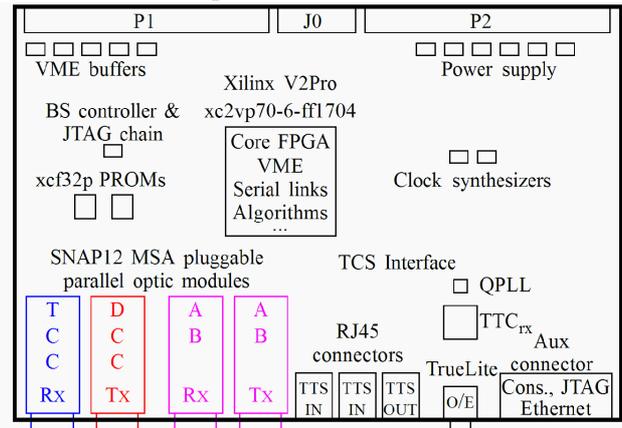

Fig. 6.   6U VME64x compliant algorithm board

1) Communication channels: The FPGA includes 20 RocketIO multi-gigabit transceiver (MGT) cores operating at up to 3.125 Gbit/s transmission rates [7]. The bi-directional RocketIO cores perform data framing, CRC calculation, data serialization/de-serialization and 8b/10b encoding/decoding of serial data streams. Up to 12 RocketIO MGTs serve for unidirectional communications with TCCs and DCCs. Up to 8 RocketIOs are used for communications with adjacent ABs.

The RocketIO serial inputs and outputs are connected to two pairs of pluggable 12-channel parallel optical transmitter and receiver modules. These modules are standard SNAP12 multi source agreement (MSA) devices [8]. One pair of the transmitter and receiver modules establishes optical communication links with the TCCs and DCCs (Fig. 7.a). These cards deploy a standard small form factor pluggable (SFP) optical transceiver [9] respectively to send classification flags to and to receive selective read-out flags from the SRP. A commercial passive optical distribution module combines up to 12 individual TCC (or DCC) fiber optic cables to a standard 12-fiber MTP assembly that is connected to a parallel optical receiver (or transmitter) module. Another pair of the parallel optical modules is used for connections with the adjacent ABs (Fig. 7.b). A commercial passive optical cross-connect ensures all-to-all connectivity. All communication channels run at 1.6 Gbit/s speed.

The choice made for the communication links has many considerable advantages. No constraints are imposed on distances between the ODE components. The high, 1.6 Gbit/s

link speed reduces the contribution of data movement in the total SRP latency. Parallel optical technology offers a solution for a compact design. The pluggable nature of the parallel optical modules and SFP transceivers gives a wide choice of manufacturers (pin compatible MSA devices), simplifies maintenance, and if necessary, allows for increase of link speeds to up to 2.5 Gbit/s without modifying PCBs. The use of a commercial passive optical cross-connect for the AB communications simplifies the design, as there is no need to develop a custom (electrical) backplane.
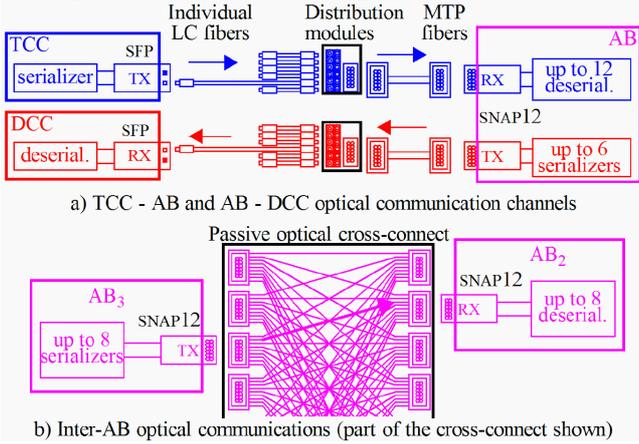

a) TCC - AB and AB - DCC optical communication channels

b) Inter-AB optical communications (part of the cross-connect shown)

Fig. 7. Organization of the SRP communication channels

2) TCS interface: A well proven design including an optical receiver from TrueLight Corp., a TTCrx ASIC and a QPLL chip both developed at CERN, is implemented on ABs. It is used to receive an optical signal from the time, trigger and control (TTC) system and to extract a clean 40 MHz LHC clock with very low jitter, level 1 trigger and various broadcast and individually-addressed control signals such as, for example, start/stop of run, bunch crossing zero, etc.

The SRP, as a part of the CMS Trigger/DAQ system, needs to signal overflow warning, busy state and other conditions to the trigger throttling system (TTS). As the TTS merging tree requires only one signal per SRP partition, each AB is able to receive TTS signals from two fellow ABs in a partition and to multiplex them with its own TTS signal prior to sending it out.

3) VME interface: A simple VME64x compliant slave interface supporting plug and play configuration and A32/D32 transactions is directly implemented in the core FPGA. External buffers are used to adapt 5V VME signals to its 3.3V IO ports. The interface is used to parameterize the AB firmware according to the run conditions and to monitor its functionality while running.

4) Configuration and test interface: The five maintenance bus lines of the VME64x backplane are connected to a boundary scan bridge supporting up to 3 local scan chains that can be accessed individually or combined serially. One local chain is dedicated for boundary scan tests of the core FPGA and the TTCrx ASIC. Another local chain is used for in-circuit programming through a JTAG port of the core FPGA and associated PROMs. The two xcf32p PROMs from Xilinx have 32 Mbit capacities each and support firmware compression with multiple design revision capability. Apart from the VME

backplane, the on-board JTAG chain can be directly accessed via dedicated pins on the front panel auxiliary connector.

5) Auxiliary interfaces: One of the two PowerPC processor cores, embedded in the Virtex2Pro FPGA device, is used to monitor functionality of the AB and to communicate via VME with the ECAL local DAQ and the CMS run control systems. Currently it is envisaged to run a standalone program on the processor, however, the use of a Linux operating system is not excluded. For this purposes a RS232 console and Ethernet circuitry is deployed on the board.

Hardware design for barrel and endcap ABs is identical. The distinction is made trough the corresponding versions of firmware. In addition, the proposed AB architecture eliminates the need for development of dedicated testing hardware. By modifying the FPGA firmware, an Algorithm Board can be transformed into a tester device for the ABs. The tester firmware can generate output data of up to 12 TCCs and send them to an AB card under test via a parallel optic transmitter, which in the normal mode of operation is dedicated to transmission of selective read-out flags. Similarly, the tester firmware can receive selective read-out flags destined to of up to 6 DCCs and verify their integrity and correctness.

Apart from reducing the efforts needed to keep SRP operational during the long life cycle of the CMS experiment (only one type of boards to maintain, use of pluggable MSA optical devices), the design permits performance increase and functional modifications, if necessary in future. A spare bit in TT classification and SR flags allows a wider choice of selective read-out algorithms. The data exchanged among the ABs is enough for algorithms with very large 9x9 TT sliding windows. Finally, the core FPGA can be replaced by a more powerful xc2vp100 device without PCB modifications.

III. IMPLEMENTATION STUDIES

The studies have been conducted using commercial development kits from Memec Inc. with Xilinx xc2vp7 and xc2vp30 FPGAs, a pair of evaluation kits for HFBR-779/7789 parallel optic modules and a pair of evaluation kits for HFBR-520ALP optical SFP transceivers, all from Agilent. The HFBR-779/7789 parallel optic pair operates in 1-2.5 Gbit/s per channel range. The HFBR-520ALP transceiver is designed for Fiber Channel speeds of 1.0625 and 2.125 Gbit/s.

A. Validation of the SRP Communication Channels

A common protocol for all SRP communications has been proposed and validated. At the lower protocol layer the 8b/10b coding scheme ensures enough level transitions for proper clock recovery from received serial data. Elementary data units are 16-bit words. In absence of application data the link is kept synchronized by continuous transmission of idle words 0xBC50 (K28.5D16.2) containing a synchronization comma character. Application data packets are delimited by start of frame 0xF7FB (K23.7K27.7) and end of frame 0xFDF7 (K29.7K23.7) control words. The latter is preceded by a 32-bit cyclic redundancy check (CRC) sequence for detection of eventual transmission errors in application data packets.

A reusable communication channel firmware has been developed in VHDL, which is independent from upper layers of the protocol stack and user applications. It instantiates a communication entity based on Xilinx RocketIO MGT that handles link level synchronization, data framing and CRC calculation. The communication entity keeps track on link activity and maintains detailed statistics.

In the test setup, up to 12 bidirectional communication channels have been instantiated and activated. For protocol validation purposes a simple tester firmware has been developed: each communication channel has been coupled with a pair of data producer and data consumer entities. Data producers repeatedly send data packets of a given size followed by a programmable inactivity period. A packet contains a header with the communication channel ID and the packet number and a succession of 16-bit data words caring predefined patterns. The tester firmware deploys an embedded PowerPC processor that runs a standalone C application to initialize, control and monitor the hardware.

Interoperability between SFP transceivers and parallel optic modules has been checked for 1.6 and 2.5 Gbit/s link speeds in both configurations shown on Fig. 7.a. Operation of parallel optic modules has been demonstrated with all 12 channels running at 2.5 Gbit/s speed (Fig. 7.b). Bit error ratio (BER) of the communication links has been estimated by long term runs and by jitter and eye opening measurements on a specialized LeCroy serial digital analyzer. Fig. 8 illustrates a bath tub diagram. It is derived from eye opening measurements on a differential LVDS output of the parallel optic Rx module receiving data from a SFP transceiver at 1.6 Gbit/s. At a BER of 10^{-12}, eye opening at the receiver output is estimated to be 0.78 unit interval (UI) that is much bigger than the required 0.35 UI needed for proper operation of the RocketIO receiver.
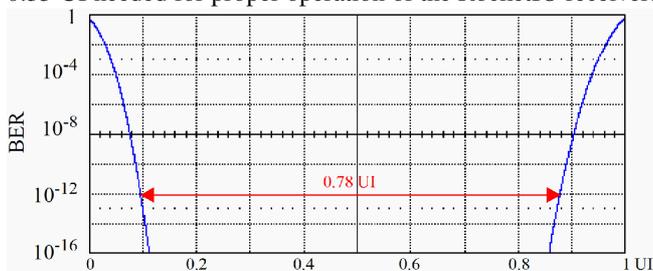

Fig. 8.   Bit error measurement at 1.6 Gbit/s speed (1UI = 625ps)

In addition to measurements, the jitter budget has been calculated for all types of communication channels using maximum values of contributed jitter characteristics known for all active devices and passive optical components including splitters and 50 m long fibers. In all cases induced total and deterministic jitters were less than respectively 0.65 UI and 0.41 UI jitter tolerances of RocketIO receivers.

Optical power budgets have been calculated as well. The TCC-AB and AB-DCC channels feature greatest attenuation due to 4 connector pairs in their respective optical paths. Nevertheless, in the worst case, there is at least 3.5dB margin between the cumulated attenuation and the input sensitivity of the optical receivers.

B. System level studies

A simplified, but sufficiently representative model of the AB has been developed in VHDL. It was complemented by a TCC emulator generating TT classification flags and a DCC emulator receiving the selective read-out flags. The simplified AB model exchanges TT classification flags with only two (fake) adjacent ABs and serves only 1/6th of the assigned barrel calorimeter area (*i.e.* a super module).

The firmware has been tested on an xc2vp7 development kit running at 125 MHz clock frequency. The latency between the level 1 trigger accept signals arriving to the TCC emulator and the corresponding selective read-out flags delivered to the DCC emulator has been measured. It amounts to 252 clock cycles, corresponding to ~2 µs and is far below the ultimate timing budget of 6.4 µs. Because of the intrinsic parallelism of the AB firmware - in reception of TT classification flags from TCCs, in exchange of boundary TT data with adjacent ABs, in production and in delivery of SR flags to DCCs - the overall SRP latency for the final version of the AB firmware will not increase significantly, even though sophistication of the selective read-out algorithm is expected.

IV. CONCLUSION

The on-line raw data reduction scheme adopted for the CMS electromagnetic calorimeter has been presented and the architecture approved for the selective read-out processor has been detailed. Implementation studies have shown validity of the proposed principles and now SRP development is underway. It is planned to commission the SRP in early 2006. The design is especially well adapted to the long life cycle of the CMS experiment. It requires reduced maintenance efforts and allows for performance increase and functional modifications, if needed in future.

REFERENCES

[1] The CMS Collaboration, "The Electromagnetic Calorimeter Project – Technical Design Report," CERN/LHCC 97-33, 15 December 1997.

[2] R. Alemany et al., "CMS ECAL Off-Detector Electronics", *In Proc. 11th Int. Conf. Calorimetry In HEP,* Perugia, Italy, March 2004.

[3] Ph. Busson, P. Paganini, "ECAL Zero Suppression Algorithms at High Luminosity", CMS Internal Note CMS IN 2002-06, 11 February 2002.

[4] S. A. Rutherford, "Study of the Effects of Data Reduction Algorithms on Physics Reconstruction in the CMS ECAL", CMS Note 2003-001, 17 January 2003.

[5] N. Almeida, J. Varela, "ECAL Data Volumes & Selective Readout", CMS Internal Note, CMS IN 2002-09, 18 February 2002.

[6] J. Varela, "CMS L1 Trigger Control System", CMS Note 2002-033, 13 September 2002.

[7] Xilinx Corp., "RocketIO Transceiver User Guide", UG024 (v2.3.2), June 24, 2004

[8] Emcore Corp, PicoLight Inc., "12-channel Pluggable Optical Module MSA", Specifications, Rev 1.1, May 15, 2002

[9] SFF Committee, "SFF-8074i Specification for SFP (Small Form factor Pluggable) Transceiver", Rev 1.0 May 12, 2001