Available on CMS information server **CMS CR-2006/092**

CMS Conference Report

15 February 2006

Flexible custom designs for CMS DAQ

Dominique Gigi

CERN

URL: http://cms-frl.home.cern.ch/cms-frl/default.htm

Abstract

The CMS central DAQ system is built using commercial hardware (PCs and networking equipment), except for two components: the Front-end Readout Link (FRL) and the Fast Merger Module (FMM).
The FRL interfaces the sub-detector specific front-end electronics to the central DAQ system in a uniform way. The FRL is a compact-PCI module with an additional PCI 64bit connector to host a Network Interface Card (NIC). On the sub-detector side, the data are written to the link using a FIFO-like protocol (SLINK64). The link uses the Low Voltage Differential Signal (LVDS) technology to transfer data with a throughput of up to 400 MBytes/s.
The FMM modules collect status signals from the front-end electronics of the sub-detectors, merge and monitor them and provide the resulting signals with low latency to the first level trigger electronics. In particular, the throttling signals allow the trigger to avoid buffer overflows and data corruption in the front-end electronics when the data produced in the front-end exceeds the capacity of the DAQ system.
Both cards are compact-PCI cards with a 6U form factor. They are implemented with FPGAs. The main FPGA implements the processing logic of the card and the interfaces to the variety of busses on the card. Another FPGA contains a custom compact-PCI interface for configuration, control and monitoring. The chosen technology provides flexibility to implement new features if required.

Presented at IPRD06, Siena Italia October 2006

1. INTRODUCTION

The CMS DAQ system [1] has been designed to collect event fragments from approximately 650 data sources at a first level trigger rate of 100 kHz. Fragments are assembled into entire events by the Event Builder and transferred to the Filter farm where higher-level trigger decisions are made by software. Accepted events are permanently stored for further physics analysis. With a mean event size of 1 MByte the system has to sustain a data throughput of 100 GByte/s. This can be achieved with the parallel architecture shown in Fig. 1.

The various sub-detector readout systems store data continuously in 40 MHz pipelined buffers. Upon arrival of a synchronous L1 trigger (3 μs latency), the corresponding data are extracted from the front-end buffers and pushed into the DAQ system by the Front-End Drivers (FEDs). The data

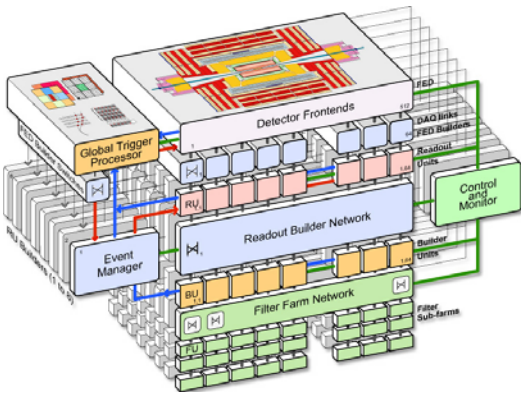

Figure 1. CMS DAQ block diagram

from the FEDs are read into the FRLs that are able to merge the data of two FEDs. The total number of FRLs is 468. The FRL interfaces with the first stage of the Event Builder based on Myrinet [2] network technology.

The design of the FED is sub-detector specific. However, a common interface from the FED to the DAQ system has been implemented. The hardware of this interface is based on S-Link64[3]. The FED encapsulates the data received from the front-end electronics in a common data structure by adding a header and a trailer that mark the beginning and the end of an event fragment. Their contents include event information partly used in the event building process, such as the event number derived from the Trigger Timing and Control (TTC) signals. The payload of the event fragments is only inspected in the filter farm.

The FEDs can request to throttle the Level-1 trigger through the synchronous Trigger Throttling System (sTTS). The sTTS must have fast reaction since the buffer capacity in the FEDs is limited. Trigger throttling signals from around 650 individual FEDs have to be merged in order to request trigger throttling for one of 32 partitions controlled by the Level-1 trigger. This task is performed by custom-built compact-PCI Fast Merging Modules (FMMs).

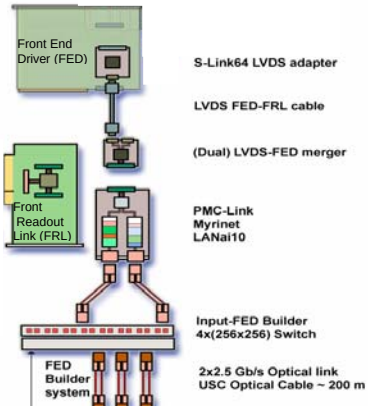

Figure 2. Diagram of one element of the sub-detector readout interface.

2. CMC AND FRONT-END READOUT LINK

The physical implementation of one element of the sub-detector readout interface is shown in Figs. 2 and 3, whose components are described in the following subsections. In total there are 468 of such elements

2.1. The S-link64 CMC

The S-Link64 Sender card is a Common Mezzanine Card (CMC).It is plugged into the sub-detector FED and receives data from the FED via an S-Link64 port. The maximum supported frequency at the S-Link64 connector is 100 MHz (64-bit data width). The card is able to buffer up to 1.6 KByte of data before generating back-pressure to the FED.

The CMC has an LVDS converter to interface with a maximum 10 m long copper cable. This cable [4] comprises 18 twisted pairs, individually shielded. The default data transfer rate over the LVDS cable is 400 MByte/s (64 bits @ 50 MHz clock), twice the sustained design speed of the DAQ system.

The CMC transmitter converts the S-Link64 signals to LVDS format, and vice-versa for a few signals transmitted in the opposite direction (back-pressure and some commands).

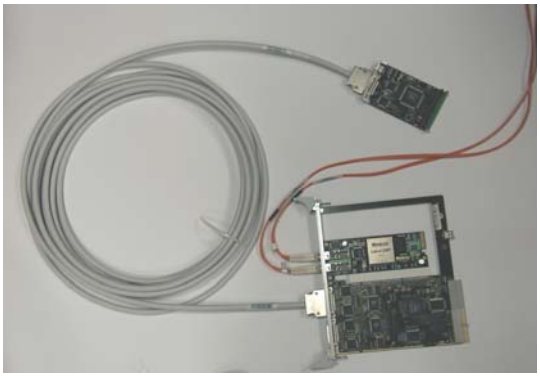

Figure 3. Picture of the S-link64 CMC (top-right), LVDS cable, FRL with Myrinet NIC.

An FPGA is used between the S-Link64 connectors and the LVDS transmitter to convert the transmitter frequency of the data, and perform a CRC calculation. It can also generate a test pattern like that specified in the S-Link64 specification [3]. This allows testing of the media and the link without any transfer of data coming from the FED.

2.2. The Front-end readout link card

The FRL is a 6U, single width, Compact-PCI card. The block diagram is shown in Fig. 4.
It has three different interfaces:

- an input interface which handles up to two LVDS cables;
- an output interface to the Event Builder implemented as a 64bit@66MHz PCI connector for a NIC;
- a configuration and control interface which is a PCI bus interface connected to the Compact-PCI backplane.

The FRL receives, buffers and optionally merges the event fragments, checks the payload CRC, and pushes the data to the NIC in fixed size blocks. The input buffer memories have a 64 KByte size. An add-on board that contains only the LVDS converter and the 64KB buffer can extend the number of input links to four.

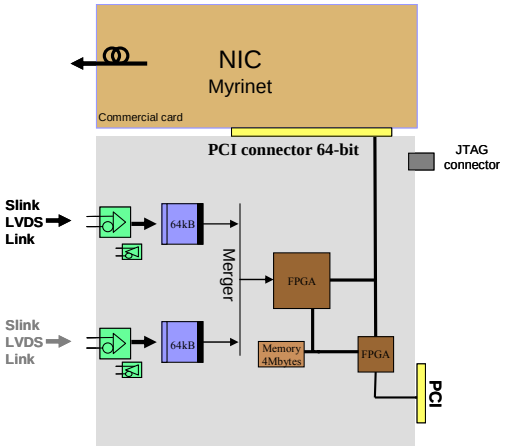

Figure 4. FRL block diagram

The FRL card also provides monitoring features, like the ability to spy on a fraction of events via the Compact-PCI bus, and to accumulate histograms of fragment size distributions.

The bridge between the compact-PCI bus and the internal PCI bus is implemented using an FPGA (Altera [5] EP20K100EFC324-1). All other functions are performed by the main FPGA (Altera EP1S10F672C6).

The custom firmware of the NIC is programmed to communicate with the main FPGA, whereby the main FPGA pushes data blocks via Direct Memory Access (DMA) into the onboard-NIC SRAM memory, together with a descriptor for each block.

The FRL cards are placed into crates with a Compact PCI backplane, up to 16 cards per crate. Each crate is connected to a PC via a PCI bridge interface. The PC serves for configuration, control and monitoring. There are in total 50 FRL crates.

3. FAST MERGING MODULE

The FMM is designed to merge either one group of up to 20 TTS signals or two groups of up to 10 TTS signals. The TTS signals are transmitted from the FEDs to the FMMs on four shielded twisted pair Ethernet cables with RJ45 connectors. Each pair in the cable transmits one bit (using LVDS). The four bits encode one of the following TTS states (in order

of increasing priority): READY, WARNING, BUSY, OUT-OF-SYNC, ERROR, DISCONNECTED.
The merged state at the FMM output is determined by counting the number of active inputs in any of the above states, applying a threshold (usually zero) and selecting the state of highest priority among the states which satisfy the threshold condition.

The FMM is a 6U, double width, compact-PCI card (see Fig. 5). It is structured similar to the FRL: an Altera FPGA (EP20K100EFC324-1) acts as a bridge between the external compact-PCI bus an an internal bus. The main processing is done in a Xilinx [6] Virtex-II Pro FPGA which is connected to the internal bus. Besides performing the merging of TTS states as outlined above, the main FPGA also performs several monitoring functions.

All transitions at the FMM input are detected and stored to a synchronous static RAM memory (ZBT) controlled by the main FPGA. For each transition, the states of all inputs after the transition and a 40-bit time-stamp with 25ns resolution are recorded. The memory can hold the history of up to 128k transitions and is used in a ring buffer mode. In this mode of operation the history memory is polled by software running on the PC that controls the compact PCI crate. A DMA mode of operation with buffer loaning is foreseen but not yet fully implemented. It will allow to transfer the transition history entries directly from the FMM to the PC memory.

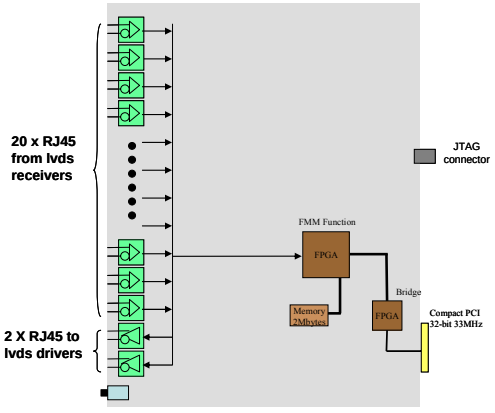

Figure 5. FMM block diagram

The main FPGA further monitors the dead times of the TTS input and output signals, separately for the WARNING (low trigger rate) and BUSY (no triggers) states.

4. CONCLUSIONS

Two custom modules have been designed for the CMS DAQ. These are compact-PCI boards based on FPGAs. The full production of 800 CMCs , 650 FRLs, and 80 FMMs has been completed and 800 LVDS cables have been purchased. Multiple test stages were realized to validate the production in terms of functionality and reliability. The test results are recorded in a database, for traceability. The equipment has been installed in the CMS underground cavern and commissioning will start in January 2007. A small part of the production is operational in setups for detector tests.

REFERENCES

1. CERN/LHCC 2002-26, CMS TDR, Data Acquisition and High Level Trigger
2. Myrinet company http://www.myricom.com
3. A. Racz, R. McLaren, E. van der Bij, The S-Link64 bit extension specification: S-Link64, available at http://hsi.web.cern.ch/HSI/s-link
4. Cable manufactured by 3M company (Multi-conductor round cable v98)
5. ALTERA component http://www.altera.com
6. Xilinx component: http://www.xilinx.com