The AMC13XG: a new generation clock/timing/DAQ module for CMS MicroTCA

E. Hazen,1 A. Heister, C. Hill, J. Rohlf, S.X. Wu and D. Zou

Boston University, Boston, MA, U.S.A.

E-mail: hazen@bu.edu

ABSTRACT: The AMC13 provides clock, timing and DAQ service for many subdetectors and central systems in the upgraded CMS detector. This year we have developed an upgraded module, the AMC13XG, which supports 10 gigabit optical fiber and backplane interfaces. Many of these modules are now being installed in the CMS experiment during the current LHC shutdown. We describe the implementation using Xilinx Kintex-7TM FPGAs, commissioning, production testing and integration in the CMS HCAL and other subsystems.

KEYWORDS: Trigger concepts and systems (hardware and software); Data acquisition circuits; Modular electronics; Data acquisition concepts

1Corresponding author.
1 Introduction

The LHC (Large Hadron Collider) at CERN is currently shut down while many elements of the machine and detectors undergo upgrades, planning to restart in 2015 at 14 TeV. The CMS detector is performing a series of updates [1] to make the experiment more efficient and to upgrade the detector to cope with higher luminosity. The Level-1 Trigger [2], pixel detector [3] and hadron calorimeter [4] will all be upgraded, along with central services such a clock, timing and controls.

The CMS detector has chosen to widely adopt the µTCA [7] standard for upgrade of off-detector electronics. This telecommunications standard provides high-speed backplane signaling which provides convenient paths for clock, timing and DAQ (data acquisition). In the current VMEbus [6] systems, these signals are implemented using custom backplanes, front panel cables and fibers and other ad-hoc solutions.

2 Upgrades to CMS central services

The CMS trigger and data acquisition is almost entirely digital. Detector signals are digitized close to the sensor, split, and sent on a low-latency path to the Level 1 Trigger, which is implemented in hardware. The digitized data are also stored in a 3.2 μs pipeline buffer. When an L1A (Level 1 Accept) is received from the Level 1 Trigger, the appropriate data are copied from the front end pipelines to readout buffers for transmission to CMS CDAQ (Central Data Acquisition).
The CMS Level 1 trigger [2] integrates inputs from the electromagnetic calorimeter (ECAL), hadronic calorimeter (HCAL) and muon subsystems to evaluate trigger conditions each bunch crossing (BX) period. A series of upgrades is planned [5], resulting in a complete replacement of the level 1 trigger in 2016. The upgraded trigger will be built entirely using µTCA hardware, and an AMC13 [9] module will be used in each crate.

Triggers are distributed and managed by the trigger control system. Currently this comprises three independent systems. The Trigger Timing and Control (TTC) system distributes L1A signals and synchronization commands. The Trigger Throttling System (TTS) collects front-end status information and propagates those up to the central Trigger Control System (TCS). The TCS allows or vetoes Level-1 triggers from the Global Trigger (GT) based on the TTS state and on the trigger rules.

These three systems will be combined in the new TCDS (Trigger Distribution and Control System) [10] (see figure 1). On the left is shown the current system in which timing and control information is transmitted to each front-end over a simplex optical fiber carrying 160 MHz biphase mark encoded data (the TTC fibers). The front end state (busy, ready, overflow warning, etc.) are returned to the TTS system on standard four-pair network cables using the LVDS standard. On the right is shown the upgraded system. For new front-ends (using AMC13 modules) a duplex fiber carries timing information to the sub-detectors using the same TTC protocol, and returns status information over the paired fiber (also using a variant of the TTC protocol). The pink boxes (Global Trigger, TCDS and new front-ends) represent µTCA crates, each containing an AMC13.

The CDAQ will be completely replaced during the LHC long shutdown 1 (LS1) in 2013–2014 [11]. The details of this are beyond the scope of this document, but the changes to the front-end interface will be described.

The current front-ends send data to CDAQ over S-Link64 [12], using standard mezzanine transmitter cards and a copper cable carrying serial LVDS data. The upgraded front-ends will send data to CDAQ over an optical fiber using a new protocol S-Link Express [11]. The new protocol and implementation are designed to provide similar key features to the old:
Figure 2. MicroTCA backplane ports use for CMS.

- Detailed implementation of both link ends by CDAQ (hardware for S-Link, firmware for S-Link Express)
- Link initialization, monitoring and testing under control from receiving end
- Easy integration of standard module into front-ends (S-Link LSC or AMC13)

The S-Link Express transmitter firmware has been implemented in the AMC13 module (v1 only as of this writing) and tested extensively with a prototype receiver card (the MOL). We anticipate shortly to integrate the newest firmware in the AMC13XG.

Currently, it is foreseen that most front-end subsystems will use $\mu$TCA crates with AMC13XG as their interface to CDAQ. However, some systems (notably the tracker) have bandwidth requirements which far exceed that of the AMC13XG and they may implement the S-Link Express transmitter directly in the proposed $\mu$TCA replacement for the front-end driver (FED) modules.

3 MicroTCA in CMS

3.1 Backplane signal use

The $\mu$TCA standard essentially defines a backplane which can accommodate up to 12 Advanced Mezzanine Card (AMC) [8] modules along with up to two MicroTCA Carrier Hub (MCH) modules. The AMC standard defines up to 21 “ports” on a backplane connector, each of which can in principle transfer data at up to 10 Gb/s. Each set of point-to-point connections routed from one or both MCH to the AMC ports is assigned a “fabric” name in the $\mu$TCA standard.

On a typical $\mu$TCA backplane, only a few of the 21 ports are actually connected. A specific backplane topology (“dual-star”) has been chosen for use in CMS, and this topology is illustrated schematically in figure 2.

In the dual-star topology, two MCH sites (MCH1 and MCH2) are provided for central hubs. In a telecommunications application, these provide redundant backup, with each routed with an independent set of point-to-point links to the various ports on the 12 AMC sites in the crate. For CMS, we use the MCH2 site to house an AMC13 module, which takes advantage of this connectivity to provide central services to the AMC modules. The MCH1 site is occupied by a commercial MCH
module which provides management and Ethernet services to all modules in the crate (including the AMC13). Fabric (port) use is summarized below:

- **Fabric A** (port 0) is used for Gigabit Ethernet (GbE) communication with all modules provided by a GbE switch on the commercial MCH. A cross-over GbE connection provides service to the AMC13 in MCH2 site as well.

- **Fabric A** (port 1) collects DAQ data from the AMC modules to the AMC13 for transmission via S-Link Express to CDAQ. Fabric A operates at 5.0 Gb/s.

- **Fabric B** (port 3) is used to transmit timing and control signals using the TTC protocol from AMC13 to the AMC cards. The return channel from AMC to AMC13 is currently unused.

- **CLK1** (port FCLKA) carries the LHC machine clock (40.079 MHz) from the AMC13 to the AMC modules.

- **Fabrics D-G** are routed from each AMC to each MCH but are unused for central CMS services. They are available for use in e.g. trigger processor crates if appropriate interconnections are provided in the MCH sites.

3.2 The AMC13 — hardware and firmware

3.2.1 Hardware implementation

The \( \mu \)TCA standard provides for up to four PC boards (“tongues”) in an MCH, each with a 170 pin backplane edge-card connector. The standard AMC13 has tongue 1 and tongue 2 boards which engage the backplane connectors, and a tongue 3 board which does not. The board stack of the AMC13 is shown in figure 3.

An overall block diagram of the AMC13 is shown in figure 4. Tongue 1 contains four front-panel mounted SFP sites for optical transceivers (one for TTC, 3 for DAQ), the initial portion of the clock chain, a Xilinx Kintex™ FPGA, two 256 MB DDR3 memories, and several voltage regulators. Tongue 2 houses the Module Management Controller (MMC) and flash memory (described in section 3.2.5) along with the second part of the clock chain. Tongue 3 serves mainly to house
Figure 4. AMC13XG block diagram.

front-panel connectors for JTAG and the MMC console, though additional signals are routed to this board for special applications.

3.2.2 Clock and timing

The AMC13 timing and clock distribution is shown in figure 4. The encoded TTC signal is received on an SFP optical receiver on the front panel of the tongue 1 board. An Analog Devices ADN2814 [14] separates the 160 MHz carrier clock frequency from the 80 Mb/s TTC data stream (A and B channels at 40 Mb/s each). The recovered clock is fanned out, distributed to the Kintex™ 7 FPGA on tongue 1, and divided by 4 through a Micrel SY89872U [15] Programmable Clock Divider/Fanout. The resulting 40 MHz is passed to tongue 2 and fanned out through a distribution tree to the backplane. The TTC data stream is transmitted from the Kintex-7™ on tongue 1 to the Spartan-6™ on tongue 2.
The µTCA standard specifies M-LVDS (TIA/EIA-899 [16]) level signaling for backplane clocks. Drivers which comply with this standard such as the DS91M125 [17] are relatively low-performance devices compared to standard LVDS drivers, and suffer from a wide range of permitted propagation delays. To compensate for this, a spare output of each DS91M125 is routed to the Spartan-6™. The Spartan-6™ phase aligns each group of three TTC data outputs with it’s corresponding clock group and transmits the data over the backplane.

The intention is that the TTC data may be recovered on the AMC card using a simple double data rate (DDR) receiver without any special timing considerations.

3.2.3 Data acquisition and event building

The AMC13 acquires event data fragments from each AMC card through a backplane link operating at 5.0 Gb/s on µTCA Fabric A as shown in figure 5. Firmware in the AMC accepts data through a FIFO-like interface, synchronized to any convenient local clock. An “almost full” signal provides flow control. A framing signal marks the start of each event fragment. Additionally, buffer status information may be transmitted through the link to modulate the rate of triggers distributed by the TCDS system described in section 2. The backplane link protocol provides CRC checking with buffering and re-transmission of corrupted blocks.

TTC signals are decoded in the Kintex-7 FPGA to recover L1A and other timing signals. The L1A are queued in a FIFO. Each L1A is combined with buffered event fragments from each AMC input to produce an output event in the CMS common data format. The event builder output may feed the S-Link Express transmitter to CDAQ, and may additionally be captured in a 512 MB DDR3 SDRAM on the AMC13 for monitoring and local DAQ on an external computer.

3.2.4 Monitoring

The AMC13 firmware implements a very large array of counters, which tally the number of words, events and errors of various types at each stage in the processing pipeline. This data is made available via the IPBus interface for software monitoring.
A prescaled subset of events may be captured in the SDRAM buffer for readout separately from the CDAQ S-Link Express output. These events may be selected by simple 1-of-n prescaling, events where the L1A number matches a pattern, or events with a specific type of error (e.g. L1A number out of sequence, CRC error). In addition to the error event, a “window” of surrounding events is captured to aid in diagnosis.

3.2.5 Infrastructure

Primary communication with the AMC13 is through GbE, switched by the MCH. The AMC13 has two GbE endpoints, one on each FPGA, implemented using IPBus firmware [19] developed for CMS. IPBus provides a convenient way to access registers on an FPGA through an Ethernet interface. These registers control and monitor operation of the AMC13. Each IPBus instance must have its own IP address, which may be set by a variety of means: (1) Default setting based on hardware serial number; (2) Set by value stored in MMC EEPROM; (3) set by RARP protocol from a software daemon and (4) set through a direct IPMB command.

The required Module Management Controller (MMC) is implemented on tongue 2 in an Atmel AT32UC3A1 micro-controller using a custom-written application in C [18]. The MMC is responsible for power management, monitoring of sensor inputs (voltage and temperature) and handling of hot-swap functions. Additional features provided by the CMS MMC include the ability to set the IP addresses for the GbE. The MMC includes a small EEPROM which may be used to store non-volatile parameters.

The MMC communicates with the MCH in the µTCA crate over the Intelligent Platform Management Bus (IPMB) which implements the Intelligent Platform Management Interface (IPMI) protocol [20]. The MCH provides a means to communicate with the AMC13 and other AMC cards through their MMC using a LAN-to-IPMI bridge function. This provides a back-up means of communicating with the module under fault conditions (e.g. when the module’s IP address is unknown).

A 128 MB serial peripheral interface (SPI) flash memory provides storage for FPGA configurations. This memory is divided into four regions: (1) a header describing the flash layout; (2) a fall-back configuration for the Spartan-6™; (3) the operating configuration for the Spartan-6™ and (4) the configuration for the Kintex-7™. The Spartan-6™ FPGA automatically loads its configuration from the flash on power-up (reverting to the fall-back if the primary fails) and provides an interface for flash programming via IPBus.

3.3 AMC13 software

AMC13 software support consists currently of several ongoing projects, which are listed in table 1.

3.4 AMC13 integration and testing

The AMC13 hardware design was validated by running an extensive series of tests to exercise each key component. DDR3 memory was tested by an industry-standard read/write test using pseudo-random data. The high speed links between chips, backplane and fiber optic modules were tested with loop-back through twice the longest anticipated connection. As an example, figure 6 illustrates an “eye pattern” test [22] performed on the backplane links at 5.0 Gb/s and on a 30 m optical fiber at 10.0 Gb/s. The vertical scale is the electrical threshold on the link receiver. The horizontal scale
Table 1. AMC13 software projects.

| Project          | Status | Contact | Notes                                               |
|------------------|--------|---------|-----------------------------------------------------|
| Base Classes     | Working| E. Hazen| Basic C++ classes for module identification initialization, control and monitoring |
| AMC13Tool        | Working| D. Zou  | Command-line tool for interactive and scripted operation |
| XDAQ (HCAL)      | Working| A. Heister| XDAQ [21] application for data taking in HCAL      |
| AMC13 GUI        | Prototype| A. Heister| GUI for initialization, control and monitoring by non-experts |

Figure 6. High speed link loop-back testing.

is the sampling time, where the full scale corresponds to one bit period. The dark blue region corresponds to the area where zero errors were detected. Each test was typically run for several days, with no errors observed.

Functional testing of the AMC13 has consisted mainly of preparing for and carrying out a full data-taking run in parallel with the existing VME system in HCAL. In late 2012 a µTCA crate was installed in the read-out chain of the CMS experiment in parallel with the HCAL VME readout. Data from runs with proton-proton and proton-lead collisions were collected, and the compared bit-for-bit between the VME and µTCA files. All data matched perfectly.

4 Summary and future plans

The AMC13XG provides clock, control and DAQ functions for use in µTCA systems in CMS and other experiments such as Muon g-2 [23]. We are currently (late 2013) preparing a production run of 50 modules, with an additional 50 expected to be produced in 2014. The initial production and test site will be at Boston University where the module was developed, with an additional site planned for set up at CERN in 2014.

We hope that this module will be the first of a series of common hardware projects for the LHC developed by collaborating institutes.
Acknowledgments

Most of this work was funded by a grant from the United States Department of Energy.

The entire CMS $\mu$TCA scheme was developed over beers with Greg Iles and Magnus Hansen, without whom this all would not be possible.

References

[1] CMS collaboration, *Technical Proposal for the Upgrade of the CMS Detector*, CERN-LHCC-2011-006, CERN, Geneva Switzerland (2011).

[2] CMS collaboration, *CMS Technical Design Report for the Level-1 Trigger*, CERN-LHCC-2013-011, CERN, Geneva Switzerland (2013).

[3] CMS collaboration, *CMS Technical Design Report for the Pixel Detector*, CERN-LHCC-2012-016, CERN, Geneva Switzerland (2012).

[4] CMS collaboration, *CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter*, CERN-LHCC-2012-015, CERN, Geneva Switzerland (2011).

[5] CMS collaboration, *TriDAS project: Technical Design Report, Volume 1: The Trigger Systems*, CERN-LHCC-2000-038, CERN, Geneva Switzerland (2000).

[6] ANSI/VITA, *ANSI/VITA 1-1994, VME64 Standard*, available from www.vita.com/home/Specification/Specifications.html.

[7] *Micro Telecommunications Computing Architecture Base Specification R1.0*, 6 July 2006, available from http://www.picmg.org.

[8] *AdvancedMC Mezzanine Module Specification Rev 2.0*, 15 November 2006, available from http://www.picmg.org.

[9] AMC13 technical specifications, available from www.amc13.info.

[10] J. Troska et al., *TCDS Overview*, Trigger Electrical Design Review, 13 November 2013, Geneva Switzerland.

indicocern.ch/getFile.py/access?contribId=2&resId=1&materialId=slides&confId=279026.

[11] A.G. Holzner, *The new CMS DAQ system for LHC operation after 2014 (DAQ2)*, forthcoming in the proceedings for the 20th International Conference on Computing in High Energy and Nuclear Physics (CHEP1023), Amsterdam The Netherlands (2013). Presentation link:
indicocern.ch/contributionDisplay.py?contribId=139&sessionId=2&confId=214784.

[12] S-LINK is a CERN specification for a FIFO-like data-link. See hsi.web.cern.ch/HSI/s-link/.

[13] G. Iles, M. Hansen, T. Gorski and E. Hazen, *CMS MicroTCA Crate Specification and AMC Card Requirements*, CMS Internal Document, cms1.bu.edu/hazen/cms_info/amc13/docs/cms_utca_crate_1_0.pdf.

[14] *Continuous Rate 10 Mbit/s to 675 Mbit/s Clock and Data Recovery IC with Integrated Limiting Amp*, Analog Devices ADN2814 Data Sheet, www.analog.com/static/imported-files/data_sheets/ADN2814.pdf.

[15] *2.5V, 2GHz Any Diff. IN-to-LVDS Programmable Clock Divider/Fanout Buffer with Internal Termination*, Micrel SY89872U Data Sheet, www.micrel.com/index.php/en/products/clock-timing/clock-data-distribution/clock-dividers/article/12-sy89872u.html.
[16] Electrical Characteristics of Multipoint Low Level Differential Signaling (MLVDS) Interface Circuits for Multipoint Data Interchange, March 2002, available from www.tiaonline.org, or Introduction to M-LVDS (TIA/EIA-899), Rev. January 2013, Texas Instruments, www.ti.com/litv/pdf/slla108a.

[17] 125 MHz 1:4 M-LVDS Repeater with LVDS Input, Texas Instruments DS91M125 Data Sheet, www.ti.com/product/DS91M125.

[18] T. Gorsky, MMC Reference Design Quick Specification, University of Wisconsin, Madison U.S.A., 26 April 2011. www.hep.wisc.edu/eced/SLHC/MMC/AMC13_MMC_v1p2_download/UWHEP_MMC_Ref_QuickSpec_Rev_1.pdf.

[19] IPBus firmware/software suite available from svnweb.cern.ch/trac/cactus/wiki.

[20] Intel Corporation, Intelligent Platform Management Interface Specifications, www.intel.com/content/www/us/en/servers/ipmi/ipmi-specifications.html.

[21] The CMS XDAQ Project, https://svnweb.cern.ch/trac/cmsos.

[22] Eye Pattern, see http://en.wikipedia.org/wiki/Eye_pattern.

[23] R.M. Carey et al., The New (g-2) Experiment: A Proposal to Measure the Muon Anomalous Magnetic Moment to ±0.14 ppm Precision, FERMILAB-PROPOSAL-0989 (2009).