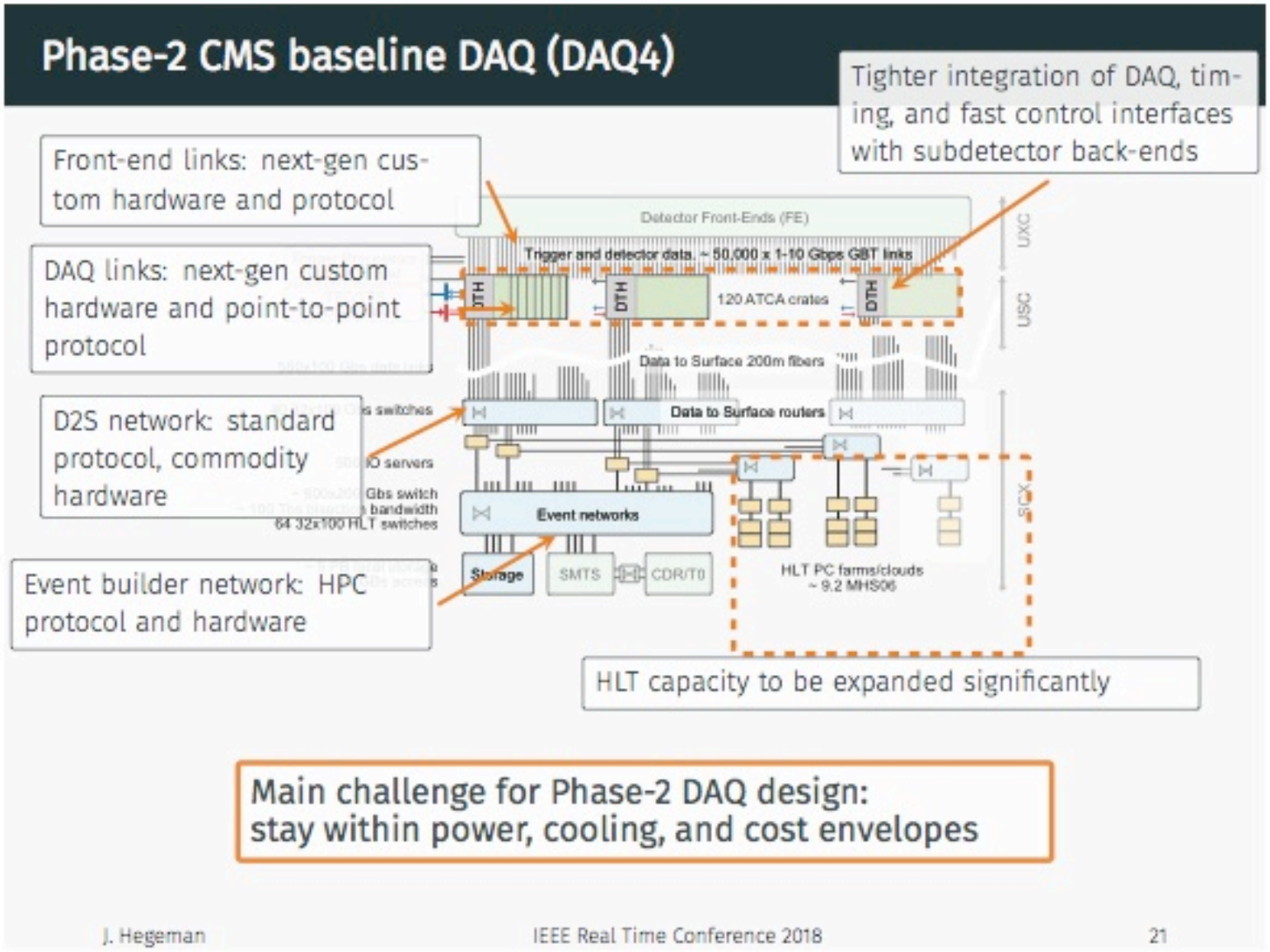
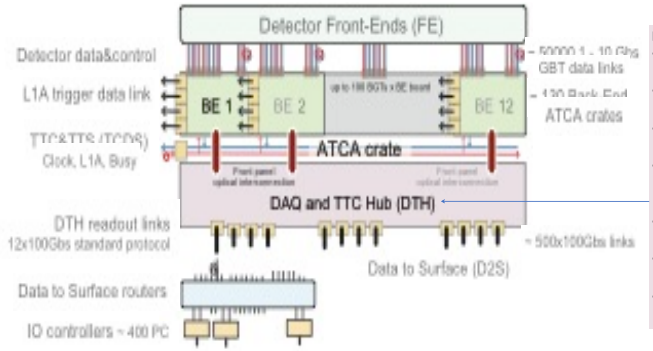
#545 Design and development of the DAQ and Timing Hub for CMS Phase-2



The DAQ and Timing Hub for CMS Phase-2

21st IEEE Real Time Conference, Colonial Williamsburg, VA USA, 9-15 June 2018
CMS DAQ group, presented by J.Hegeman and F.Meijers CERN EP/CMD

ABSTRACT: The CMS Detector will undergo a major upgrade for the Phase-2 of the LHC physics program, which will start around 2026. The detector will be read out at a rate of 750 kHz by some 50k high-speed optical links, for an average event size of 7.5 MB. In the baseline architecture for the Phase-2 DAQ [1], the optical links from detector front-ends are aggregated in detector-dependent ATCA based back-end boards. A DAQ and Timing Hub (DTH) aggregates data streams from multiple back-end boards over point-to-point links. The DTH combines these streams to feed high speed commercial 100 Gb/s optical links, forming the data-to-surface (D2S) network. It provides buffering for time decoupling and transmission using a reliable high-level protocol, such as TCP/IP. The D2S links carry the data to surface, connecting the DTH output via standard network to I/O servers for the event building. The DTH is also distributing trigger accept and timing signals, as well as trigger control codes for calibration and synchronisation to the back-end electronics, from where they are redistributed to the front-ends. This poster presents the system level functionality and performance requirements of the DTH. The DTH will have a modular design, based on Timing unit and a DAQ unit (with 0.4 Tb/s bandwidth). The first step of the DTH development roadmap (P1 for a DTH-400) is currently under design.



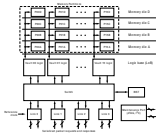
DAQ and Timing Hub functions:

- Standard interface from sub-detector back-ends to central-DAQ
- Data aggregation from back-end node boards in ATCA crate via front panel mid-board optics
- Convert from custom DAQ protocol to 100 Gb/s TCP/IP (switch+NIC)
- Total maximum throughput 0.4 Tb/s for DTH-400 (1 DAQ unit) and 0.8 Tb/s for DAQ-800 (2 DAQ units)
- Distribution of TTC signals and collection of back-end/buffer status (TTS) over backplane
- Emulation and data generation for testing purposes
- Monitoring of DAQ and TTC/TTS functionality as well as performance

DAQ FPGA (Xilinx KU15P) Gb/s for HMC
Serdes usage:
- 16x25 Gb/s (or 24x16Gb/s) for input from BE via FireFly
- 4x(4x25) Gb/s for TCP/IP out
- 32x16
- 1x16 Gb/s for TCDS-FPGA communication

TCDS FPGA (Xilinx KU15P)
Serdes usage:
- 7xGTX + 6 x GTY to backplane
- 1xGTX = 1xGTY to frontpanle
- 1x16 Gb/s for DAQ-FPGA communication
- 4x25 GB/s to firefly

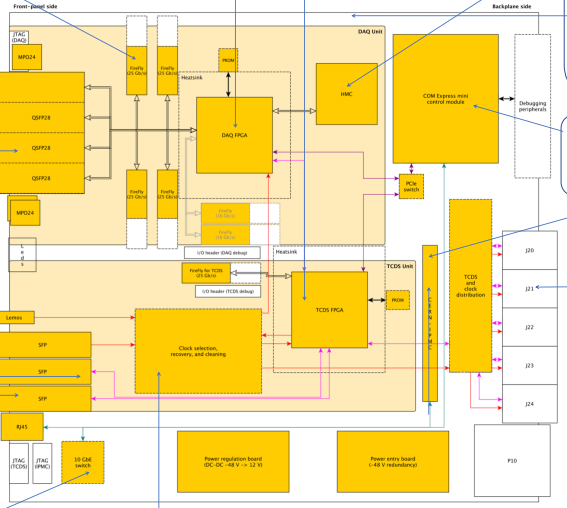
Fast serial memory **HMC** (Hybrid Memory Cube) for output buffering:
- 2 Gbyte
- 240 Gb/s bandwidth and 4x8 16Gb/s links



Samtec **FireFly**

With Y-cable to MPO24 front panel 3x2x4 duplex 16/25 Gbps
- Communication with BE boards via **p2p custom protocol** (slink)

DTH prototype P1 diagram (1xDAQ unit + 1xTiming unit) for ATCA Hub-slot



4x QFSP28 cage for 100 Gb/s Ethernet transceiver
- Reduced **TCP/IP** implemented in **FPGA** logic [4], using **HMC**
- On PC, using **Linux TCP/IP** stack, uses 1 x86 core per 50 Gb/s

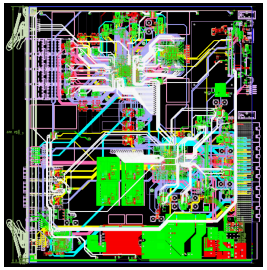
Com-Express on-board PC running linux for configuration, control, and monitoring.
Access to devices via **PCIe**.

IPMC from CERN [3] for low-level control

10 Gb/s over **backplane** to/from back-end boards **serial TTC/TTS** protocol

SFP+ for LHC **clock**

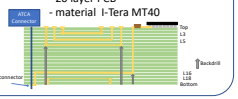
SFP+ for **10 Gb/s serial TTC/TTS protocol** from/to distribution tree



Ethernet and switch for connection to all components, for configuration, control, monitoring

Extensive **clock cleaning** and monitoring infrastructure

ATCA board layout
- 20 layer PCB
- material I-Tera MT40



(1) The Phase-2 upgrade of the CMS DAQ Interim Technical Design Report , CERN LHCC-2017-014.
(2) TCDS (Trigger and Control distribution System)
(3) IPMC (MModule developed by CERN ESE group)
(4) 10Gbps TCP/IP streams from the FPGA for the CMS DAQ Eventbuilder Network (TWEPP 2013 Perugia Italy)

Design Choices

- ATCA
  - infrastructure needed on board (IPMC, etc.)
  - processor and software for configuration, control and monitor
    - zync
    - ComExpress PC on the board
    - IPbus over ethernet
    - pci-express over backplane
  - FPGA serial links for input / output
    - backplane
    - mid-board optics (FireFly, etc)

Design Choices

- Loss-less concentration / event-building
  - Ethernet: custom protocol over UDP, converged ethernet, TCP/IP ?
  - RDMA or others in FPGA
- Memory resources with large size and bandwidth
  - Block- RAM
  - recently appearing fast serial memory (Hybrid Memory Cube)
- Timing and Trigger control