IEEE Real Time Conference 2012 Poster PS2-18

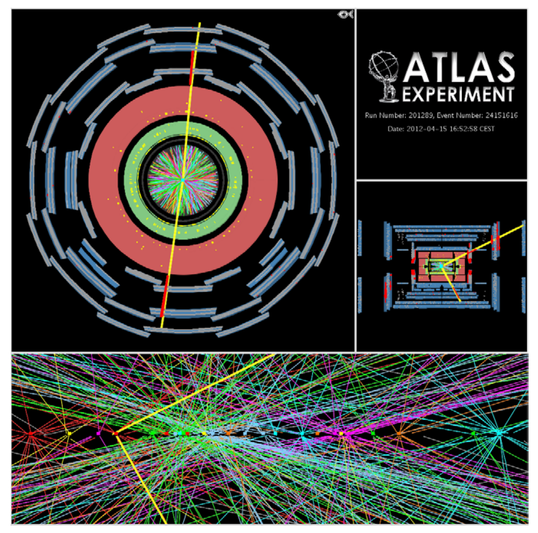
Data Formatter System for the ATLAS Fast TracKer

Jamieson Olsen[1], Ho Ling Li[2], Ted Liu[1], Yasuyuki Okumura[1,2], Bjoern Penning[1]

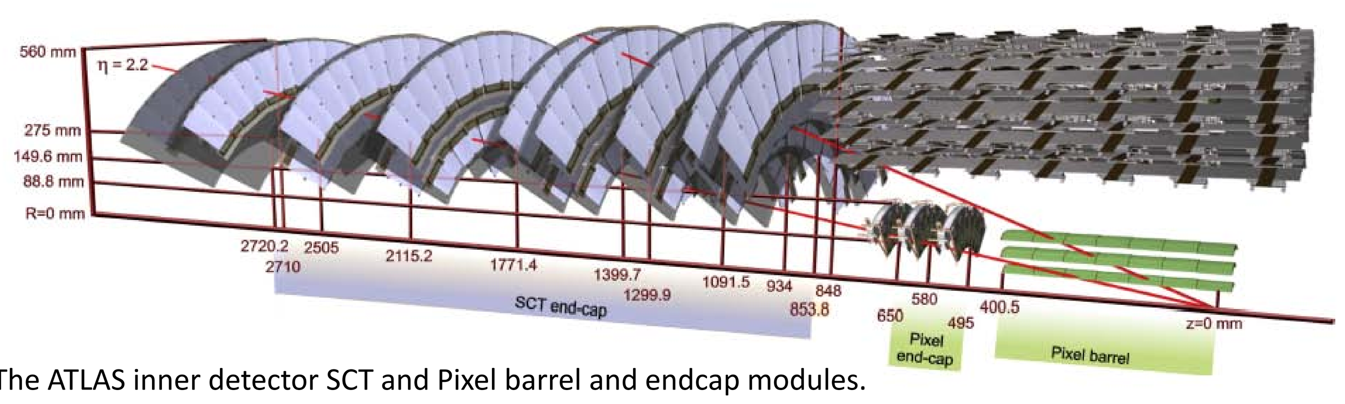
[1] Fermi National Accelerator Laboratory, United States
[2] University of Chicago, United States

The Tracking Challenge

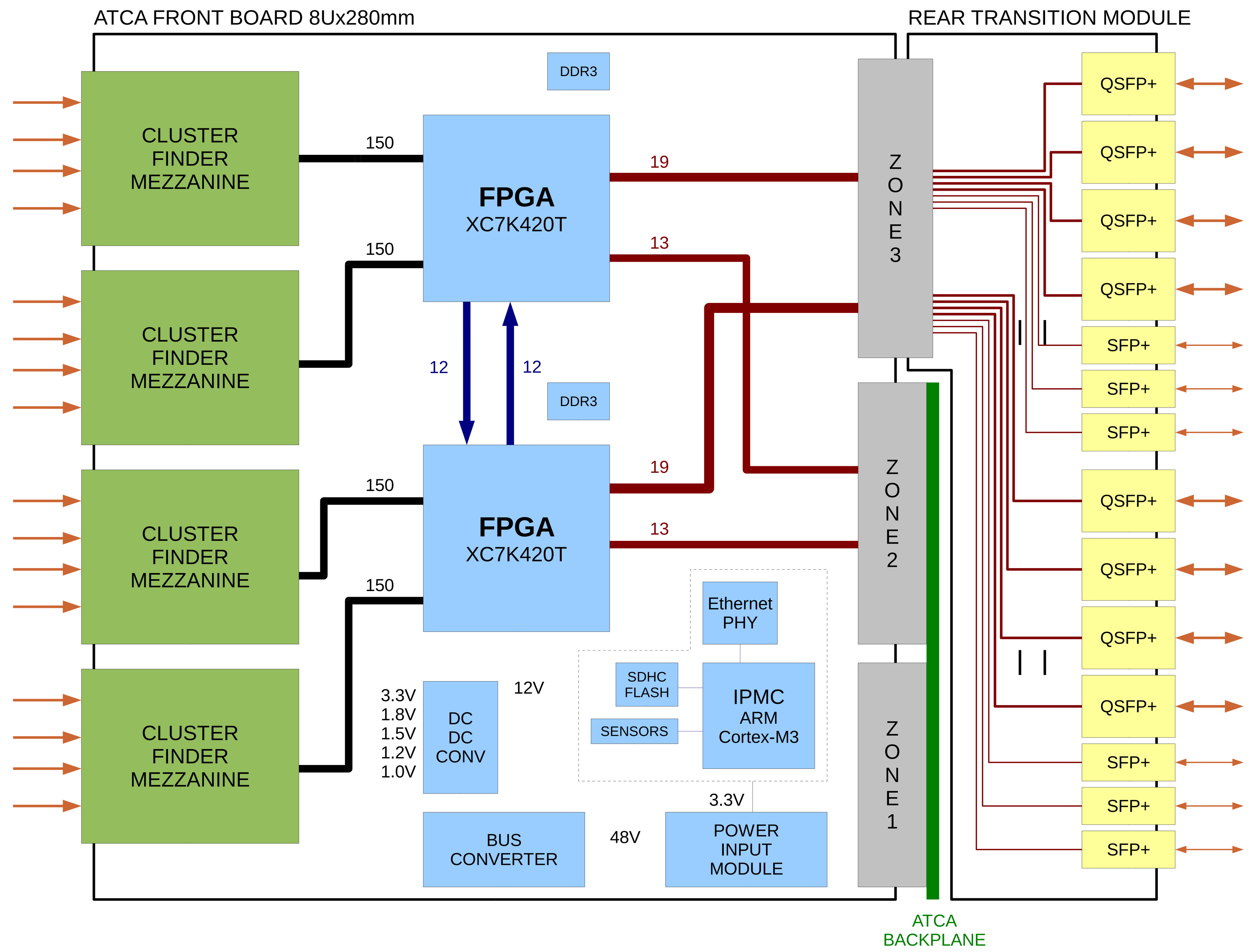
Crossings in the LHC occur at the nominal rate of 40MHz with a design luminosity of 1×10^{34} and approximately 25 overlapping proton-proton collisions per crossing. The ATLAS detector trigger system must reject a vast majority of these events, and only 200 events per second can be stored for later analysis. Instantaneous luminosity is expected to increase to $3{\times}10^{34}$ with an average of 75 collisions per crossing. Under these conditions the existing ATLAS trigger is strained and the need for a tracking trigger is clear. The **Fast TracKer (FTK)** involves adding a hardware-based level-2 track trigger to the ATLAS DAQ system.



A Z → mu mu event with high pileup.

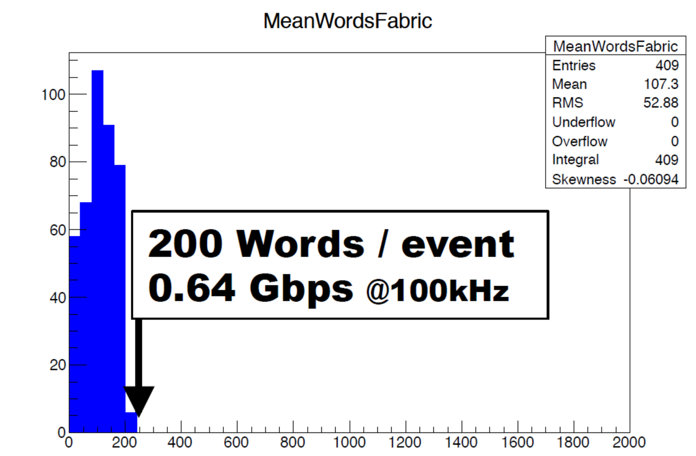


The ATLAS inner detector SCT and Pixel barrel and endcap modules.



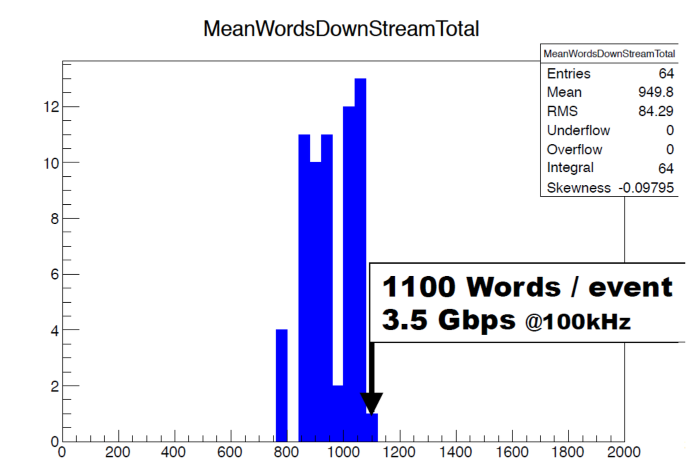
Data Volume Analysis

A software model of the Data Formatter system is used to analyze data flow using real event data (raw SLINK records from the RODs). Using this model is it possible to accurately measure data bandwidth on every link in the system. The following histograms were generated using recent events with pileup μ=10 @ 7 TeV and 100kHz trigger rate.



Backplane Fabric Links

The busiest backplane fabric link has a date rate of 0.64Gbps. A conservative estimate with μ=80 @ 14TeV shows this worst case link would be 5.9Gbps. A major advantage of the full mesh backplane is the ability to balance traffic by redirecting data from busy links to alternate, under-utilized paths between FPGAs.



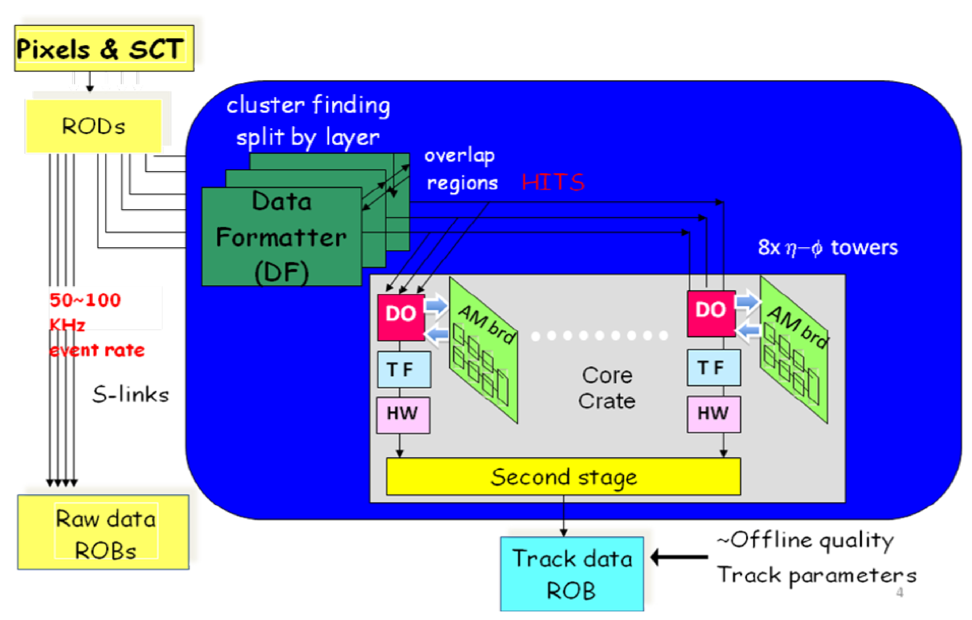
Downstream Links

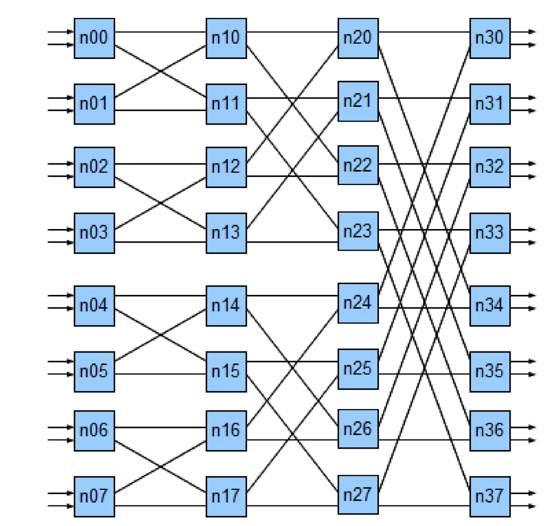
The worst case downstream link is 3.5Gbps with μ=10 @ 7 TeV. With μ=80 @ 14TeV this worst case link is 32Gbps, which is within the capacity of a single QSFP+ transceiver.

FTK Architecture

The FTK system will find and fit tracks using the inner detector silicon layers for every event that passes the level-1 trigger. It receives the Pixel and SCT data at full speed from an duplicate output added to the ROD optical transmitter mezzanine cards. The FTK system is a scalable, highly parallel processor which uses an associative memory approach to quickly find track candidates in coarse resolution roads. Roads which match the selection criteria are then analyzed using full resolution silicon hits and the track parameters are reported to the level-2 trigger.

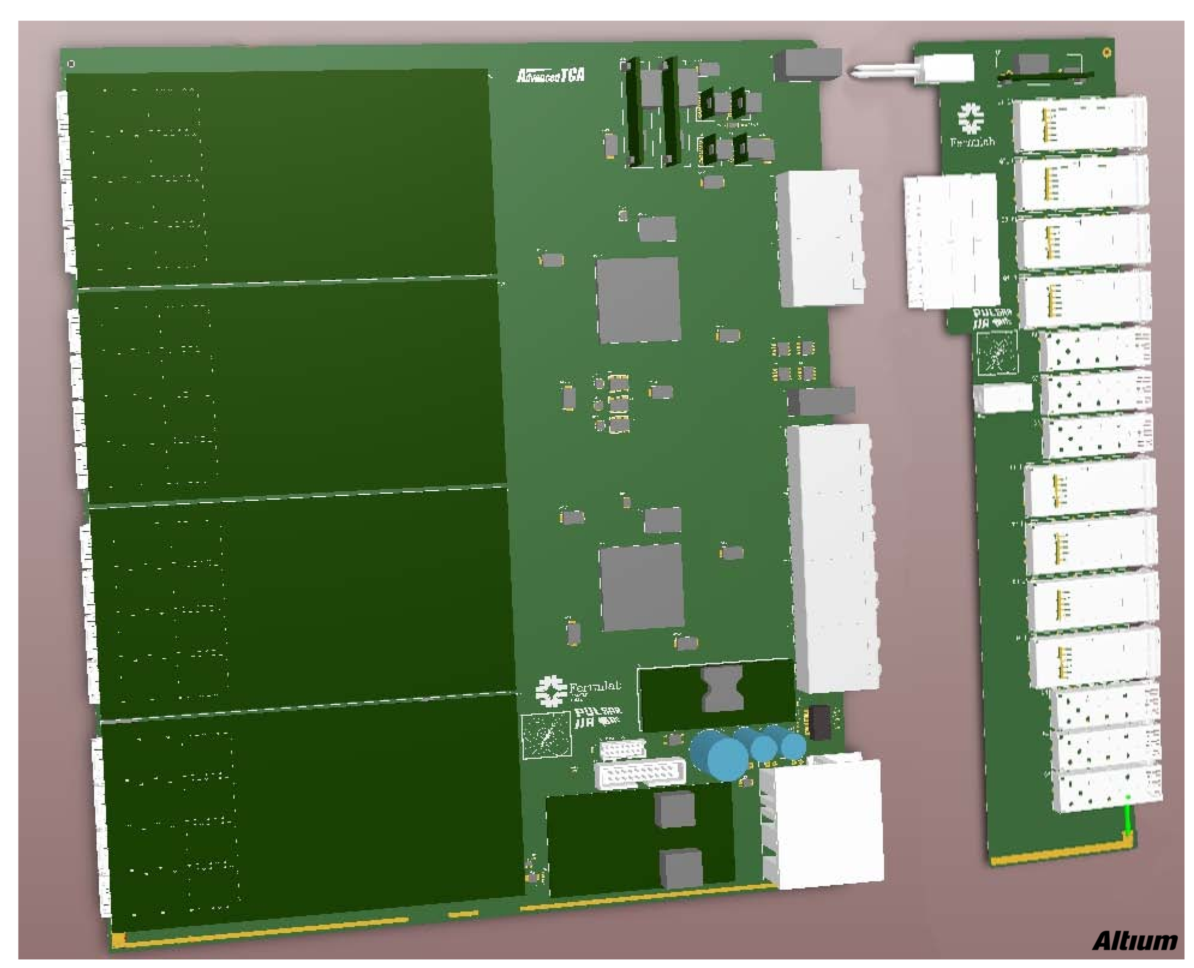
The FTK system includes a Data Formatter to remap the ATLAS inner detector geometry to match the 64 FTK η-ϕ towers. The Data Formatter system also performs pixel clustering and data sharing in overlap regions.

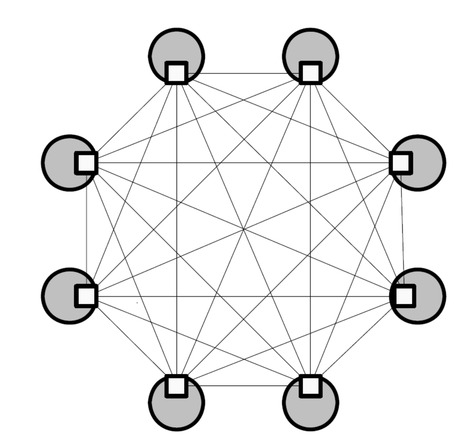




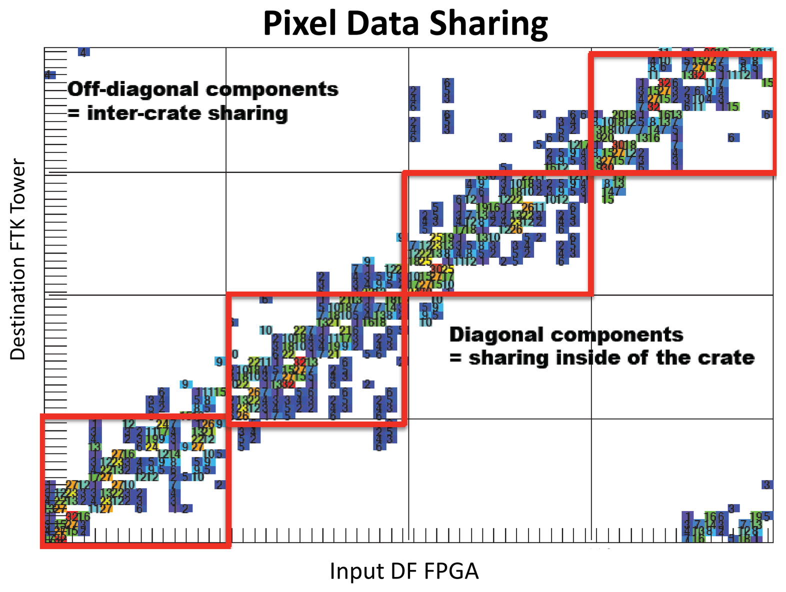
Data Routing Firmware

Each FPGA receives data from two mezzanine cards, the local bus, up to 13 backplane fabric links and inter-shelf fiber links. Incoming data packets may be sent downstream or retransmitted to another FPGA over another link. The performance of an internally buffered banyan network switch (right) is currently under investigation.

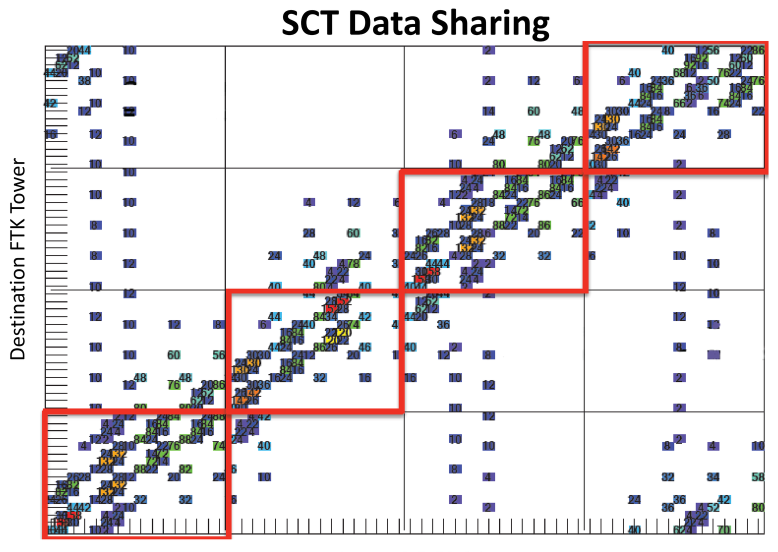




Pixel Data Sharing



SCT Data Sharing



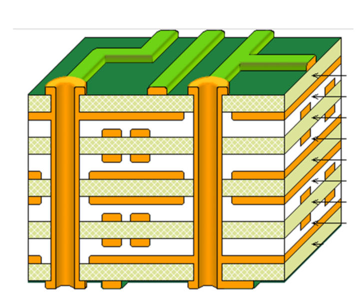
222 optical fibers (SLINK) are connected to 32 Data Formatter boards and the inter-board and inter-shelf data transfers are shown above for Pixel and SCT detector types. In some cases a single input link supplies module data for many FTK towers – and this data must be transferred over the backplane fabric or between shelves over dedicated fiber links.

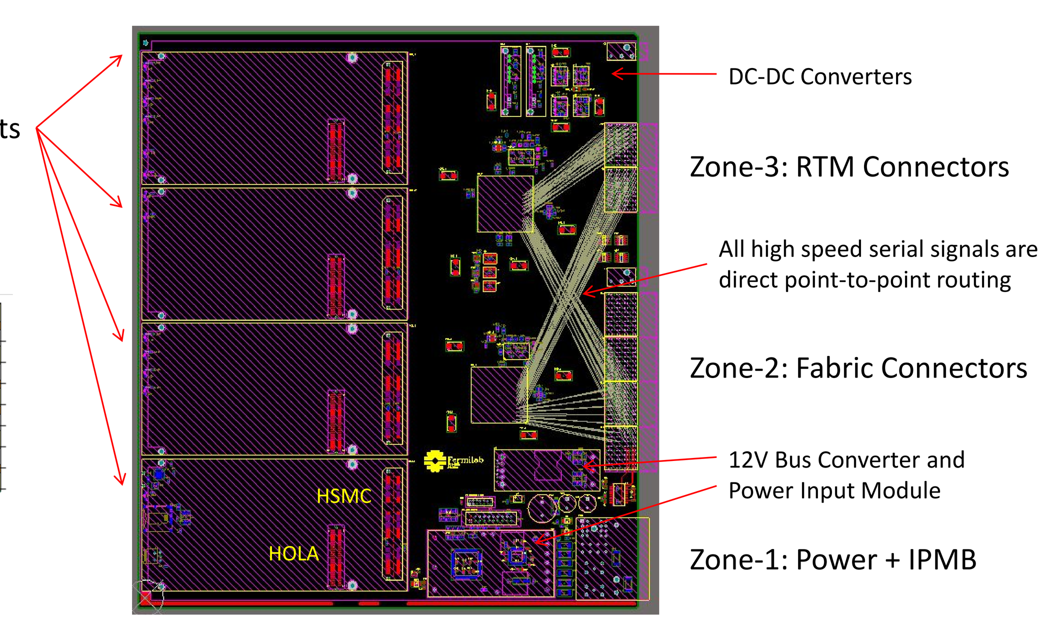
Data Formatter Architecture

The Data Formatter hardware design is determined by input and output requirements; while also maintaining the flexibility needed to accommodate future expansion and allowing for changes in the number of input links and inner detector module assignments. Early in the design process it became apparent that data sharing between FPGAs was irregular and dependent upon the mapping between inner detector modules and input links. Thus it was critical to select a backplane topology that allows any FPGA to communicate directly with any other FPGA within a crate. Likewise inter-crate data transfers must also be flexible and scalable. The ATCA full mesh backplane provides direct high speed serial links between all slots and is a natural fit for the Data Formatter system.

Data Formatter Features

- Two large FPGAs per board, one FTK η-ϕ tower per FPGA.
- 32 DF Boards, 8 boards per ATCA shelf.
- Four mezzanine cards per board, HSMC and CERN SLINK/LDC (HOLA) connectors are supported.
- Each FPGA has 19 bidirectional SERDES connections to the RTM; each connection supports data rates up to 10Gbps.
- Each FPGA has an external memory chip 256MB DDR3-800 with 16-bit data path (up to 1.6GB/sec).
- All Top FPGAs in the shelf are directly connected through the backplane fabric (up to 10Gbps per port).
- All Bottom FPGAs in the shelf are directly connected through the backplane fabric (up to 10Gbps per port).
- Top and Bottom FPGAs are connected with an LVDS local bus which supports data rates up to 20Gbps bidirectional.
- ARM Cortex M3 microcontroller is used for ATCA Shelf Management (IPMB bus) and slow controls (100BASE-T Ethernet).
- Firmware images may be downloaded to the SDHC flash memory via HTTP or FTP over the Ethernet interface.
- Inter-shelf data sharing occurs over RTM transceivers.


10 layer controlled-impedance PCB



Key Performance Specifications

FPGA	Mezzanine Card / RTM Optics	ATCA Fabric
Xilinx Kintex 7 420k Logic Cells 32 SERDES (up to 12Gbps) 4MB Block RAM LVDS local bus up to 20Gbps 256MB DDR3-800	**Mezzanine Card** HSMC Connector 40 LVDS signals Data rates up to 64Gbps **RTM Optics** 8 x QSFP+ (up to 40Gbps) 6 x SFP+ (up to 10Gbps) Up to 380Gbps	14 Slot Full Mesh 13 ports / FPGA All fabric channels support data rates up to 40Gbps

ATCA Hardware

- Designed by the telecommunications industry with emphasis on high performance, redundancy and high availability.
- 14 slot full mesh *Fabric Interface* backplanes are rated for up to 40Gbps direct bidirectional communication between every slot, with no switching or blocking.
- A dual star Gigabit Ethernet *Base Interface* is also provided on the backplane.
- Dual redundant shelf manager boards monitor fan speeds, temperatures, etc.
- Fans and boards are hot swappable.
- Redundant hot swappable 48VDC power supplies.
- Forced air cooling supports up to 300W per slot.
- ATCA front boards are 8U x 280mm and the shelf is 12U in height.
- Rear transition module (RTM) boards are 8U x 70mm.
- Many switches, routers, and single board computers are available.

Current Status

- The prototype board is in layout, first boards are expected later this year.
  - The prototype has two XC7K325T devices (16 GTX SERDES transceivers).
- Software simulations of the Data Formatter system are ongoing.
- Firmware development is proceeding:
  - Packet routing algorithms.
  - Spy buffers, diagnostics, and slow control interface.
  - DDR3 Interface.
  - Clock distribution and GTX Transceivers.
- ARM Cortex M software development using the KEIL ARM-MDK tools.
- Full Mesh ATCA Shelf and power supplies are on hand.