Trigger Hardware Development - The CMS GCT Muon and Quiet Bits System

Modular Trigger Processing Architecture

Matt Stettler, Magnus Hansen, Greg Iles
CERN
Costas Foudas
Imperial College
John Jones
Princeton University

9/2007

GCT Muon and Quiet Bits System

- Secondary Function of the GCT
  - Receive data from 18 RCT crates
  - Reformat and pass on to GMT
- Details
  - 18 fiber inputs
    - 1.6Gbps from GCT source cards
      - Located in RCT racks
    - Standard 8b/10b encoding
    - Each fiber represents 40 degrees in phi, ½ barrel in eta
  - 24 fiber outputs
    - 1.44Gbps to GMT
    - DC coupled copper cables
      - Designed to comply with previous version of GCT
    - Data organized in 120 degree phi, full barrel eta

Design Approach

- Current GCT modules ill suited for task
  - Architecture suited for electron and jet processing
    - Special purpose modules mimic detector organization
  - Dataflow matched to RCT data output
    - Reformatting data to GT requirements awkward
    - System is designed to concentrate, rather than reformat data
      - Bottlenecks exist for pass through
- Lessons learned on current GCT
  - Dataflow defined in hardware
    - Large parallel buses
    - Logically inflexible
  - Large circuit boards required
    - Expensive to develop in both time and money
    - Tend to higher overall risk (technical/schedule/funding)

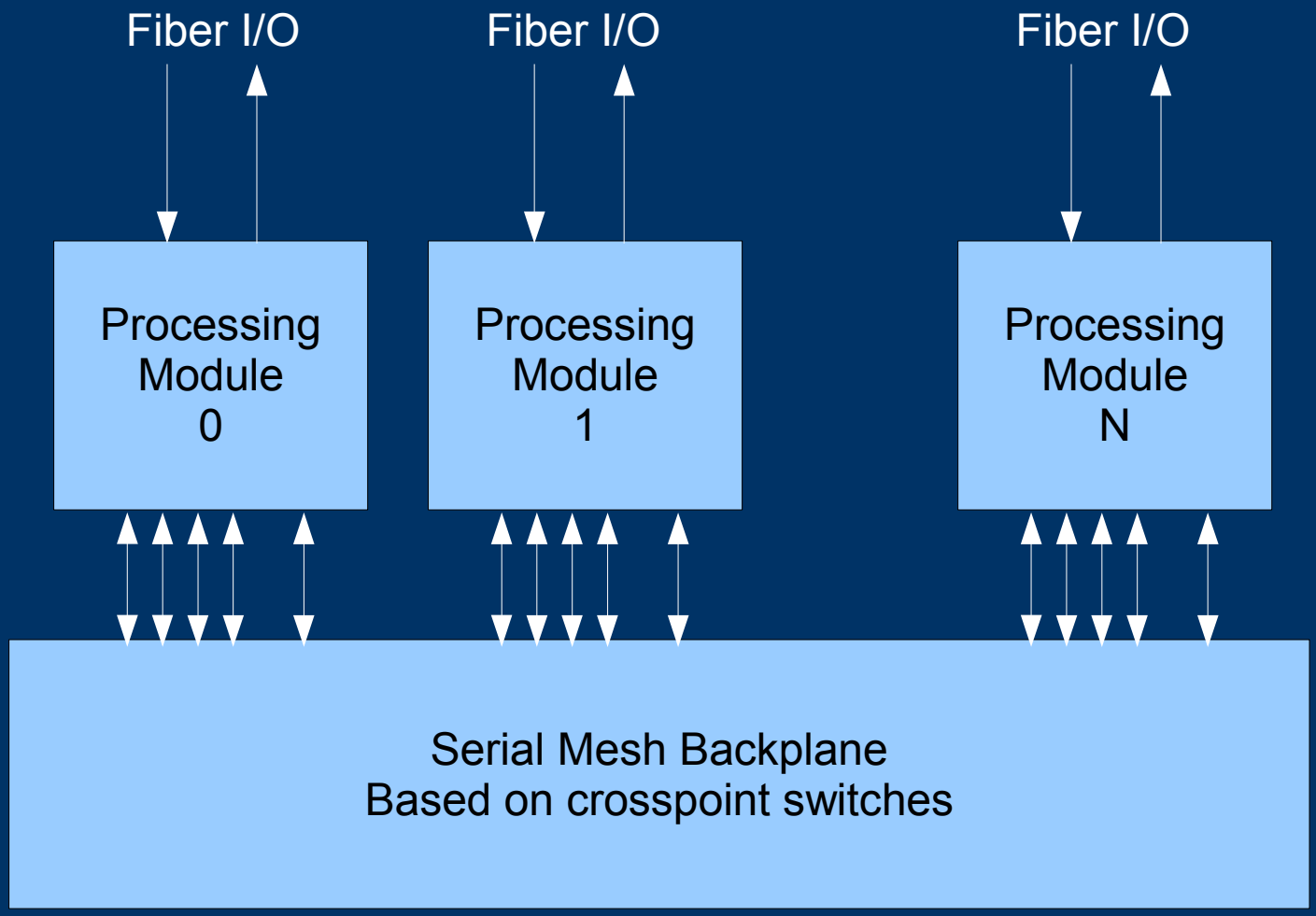
Desired Attributes for new Systems

- Modular
  - Reasonably fine grained
    - Smaller circuit boards
    - Easier and less expensive to develop
  - Should scale well
- Well defined internal interfaces
  - Allows modules to be developed independently
    - Perhaps shared across projects
  - At least electrically compatible
- Flexible in both logic and interconnect
  - Retain flexibility of FPGAs
  - Add complementary data routing flexibility
    - Modify data flow without altering hardware
    - Perhaps give the ability of dynamic reconfiguration

Relevant Current Technology

- Large FPGAs with built in SERDES links
  - Very high density logic
  - SERDES I/O allows physical concentration of data
    - 16-24 3+Gbps links available on large devices
    - Normal I/O pins support 1+Gbps links
- LVDS/CML crosspoint switches
  - Up to 144x144 non blocking matrices
  - Asynchronous, Protocol agnostic
    - Multi rate switching supported intrinsically
  - Currently available at 4+Gbps, 10Gbps announced
- Serial Memory becoming available
  - FBDIMM modules designed for server market
    - Support very large arrays (>100GB)
    - Electrically compatible with LVDS/CML

Overall architecture of CMS HCAL Muon and Quiet bit distribution

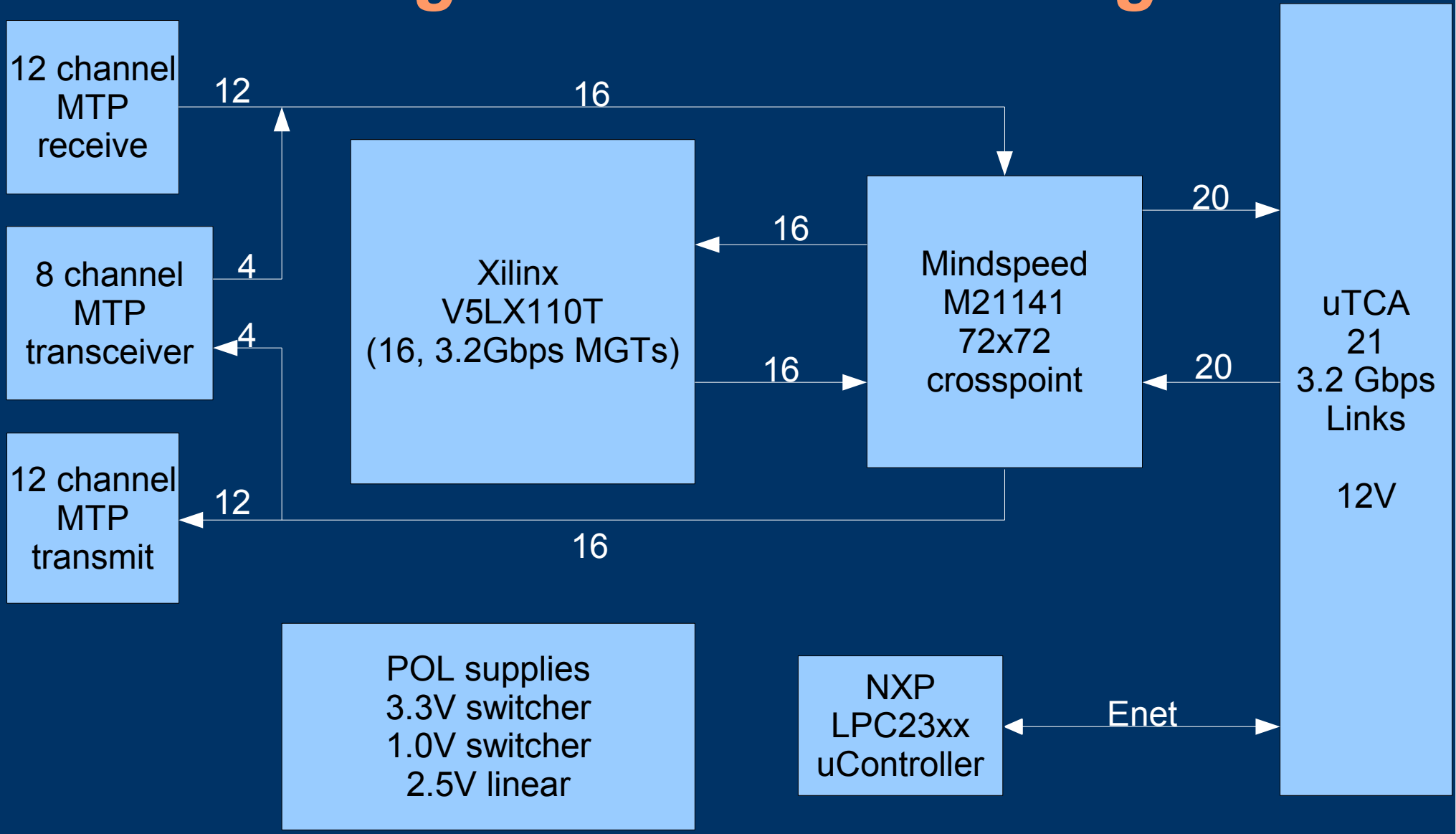


Implementation

- Planned for CMS Global Calorimeter Trigger HCAL muon & quiet bit distribution
  - This system is in layout now
- Micro TCA based, 12 slot chassis
  - 20 high speed links per slot
- Processing module
  - Based on Xilinx V5LX110T
  - Mindspeed M21141 72x72 crosspoint
  - MTP Fiber interface (SNAP-12)
- Backplane
  - Commercial unit possible, but limiting
  - Custom unit based on Mindspeed M21161 144x144
    - Programmable topology
    - 500 Gbps maximum bandwidth

Micro TCA Crate/Backplane

Processing module Block diagram



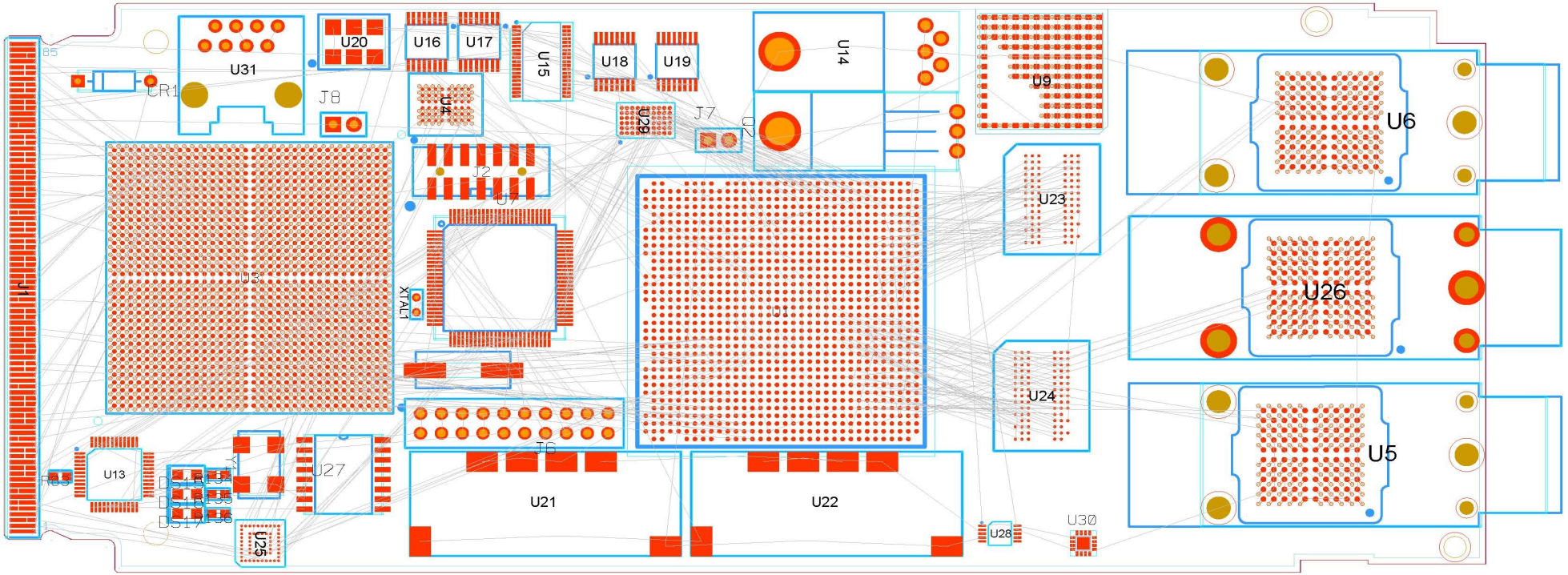
Key Features of Processing Module

- Xilinx Virtex 5LX110T FPGA
  - 16 low power MGTs, Superior logic density/speed
  - Standard I/O 1Gbps capable
- Data I/O direct from fiber
  - 16 channels full duplex @ 3.2 Gbps
  - Fiber utilized for inter-crate data links
  - Same fiber modules used on previous system
- Crosspoint routes to FPGA, backplane, and fiber
  - 1:N data replication supported at wire speed
  - FPGA output data sent to backplane or fiber
  - Data sharing between systems intrinsically supported
- Slow control via standard Ethernet
  - Based on NXP 23xx uController
  - Standard Ethernet protocols

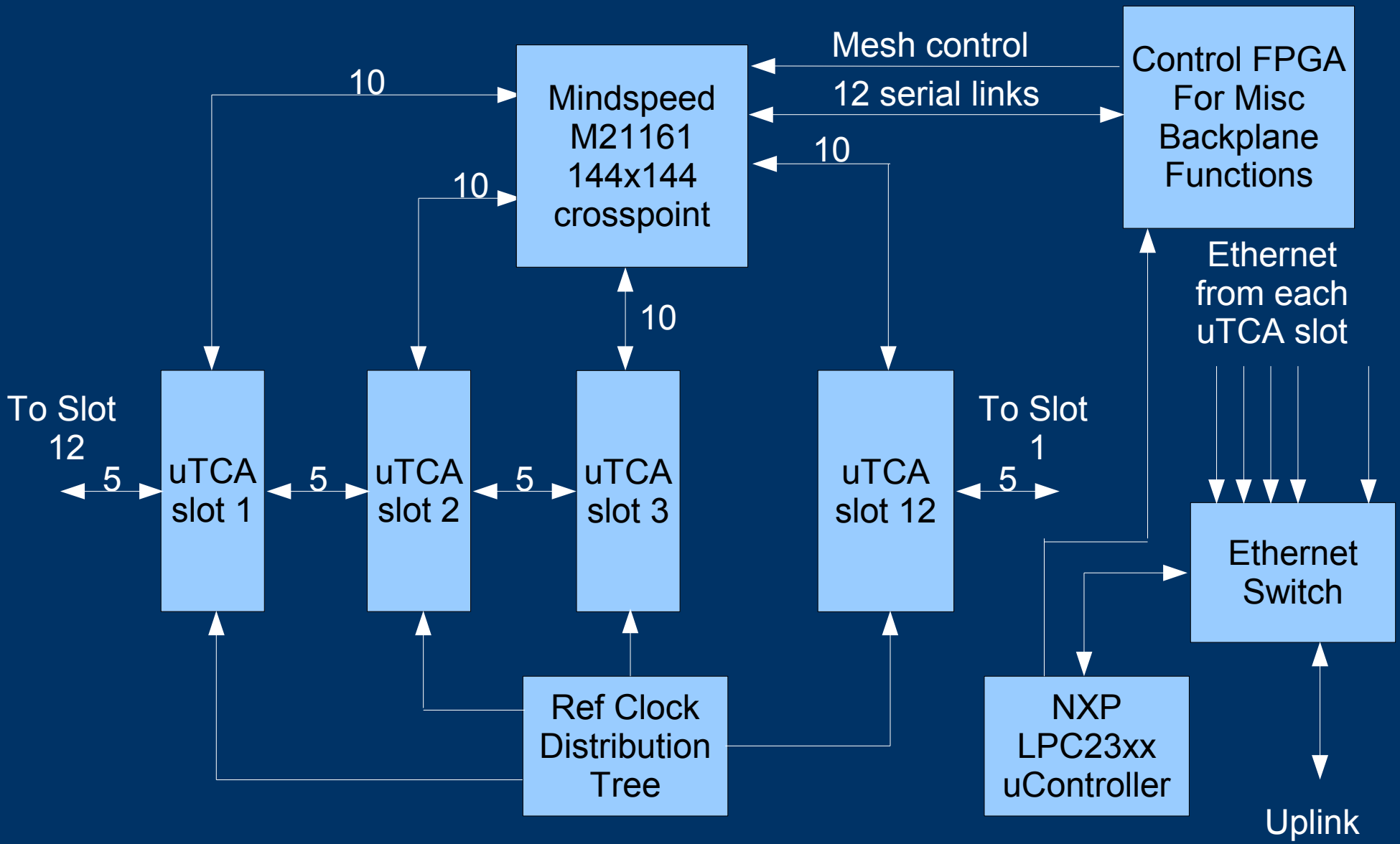
Additional Features

- Module development is co-funded by Los Alamos
  - Intended for signal processing research
- Additional RAM required for this purpose
  - 512MB DDR2 SDRAM added to module
    - Two banks of 128Mx16 @ >500MHz
    - Enables sophisticated SoC possibilities
- Initially unused Xilinx V5 features
  - PCIExpress endpoint
    - Can be connected directly to fiber I/O
    - Commercial PCIExpress fiber extenders available
      - Could connect to DAQ PCs directly for next generation readout
  - GigE MAC
    - Can be interfaced directly to fiber
    - Obvious DAQ and control potential

Preliminary Layout



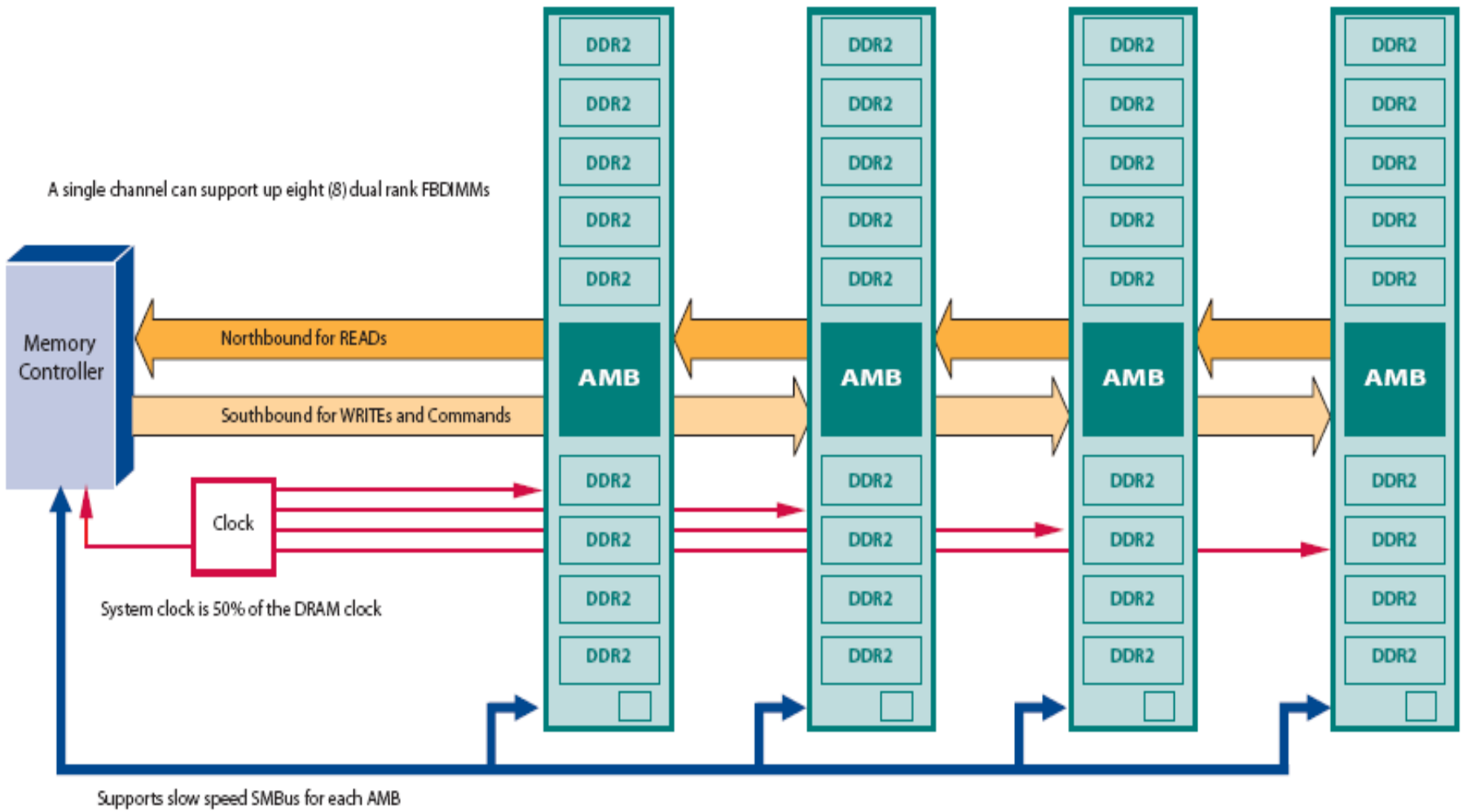
Custom Backplane block diagram



Key Features of Custom Backplane

- Active backplane based on crosspoint switches
  - Mindspeed M21161
  - Active backplane eliminates congestion
    - uTCA Hub slots cannot support full mesh connections
- Programmable topology
  - Mesh, star, etc based on crosspoint configuration
- 20 links/slot for high speed data communication
  - M21161 supports 4.1Gbps data rate
    - System will be limited to 3.2 Gbps by FPGA/MTP Optics
- 1 link/slot for 100Mb Ethernet control interface
  - Star topology interface, switch IC on backplane
  - Backplane configuration via Ethernet also
    - Same uController as processing module

FBDIMM Architecture
(related technology)



Key Features of FBDIMM

- Serial memory is designed for large blocks
  - Packet based logical model
  - Optimized for block transfers
- Utilize Advanced Memory Buffers (AMB)
  - Control local banks of DDR2 devices
  - Create daisy chain of banks
- Interface to host is via CML links
  - Currently working at 3.2 Gbps, eventually up to 9.6
  - Utilize 12-14 full duplex links
- Up to 8 AMBs can be daisy chained
  - Achieves 192GB with current technology
- Can be combined with serial mesh
  - Very large distributed memory arrays possible

Serial Link Latency

- Historically the limiting factor
  - Deserialization has dominated latency
  - This has improved with new FPGA families
- Current performance (Virtex5)
  - 10-12 byte clocks for receive
    - This is 2.5-3 40MHz clocks in our implementation
  - 5-7 byte clocks for transmit
    - 1-2 40MHz clocks
- Still significant
  - Biases FPGA choice to fewer large devices
  - Number of “hops” needs careful attention
- Crosspoints do not contribute to latency
  - Nearly wire speed (ps delay) - insignificant
  - No serialization/deserialization