Read Out Driver (ROD) R&D for ATLAS LAr Calorimeter Upgrade

Joseph Mead
Brookhaven National Lab

Common ATLAS CMS
Electronics Workshop
March 19-21, 2007

ROD R&D outline

- A possible design scenerio of the ATLAS LAr FEB is to digitize and transmit every channel for every bunch crossing.
  - Receive very large data rates from this new FEB
  - Buffer data digitally until Level 1 accept is received.
  - Calculate energy and timing of calorimeter signals from discrete time samples using FPGA technology
  - Since all information would be available at ROD, possibility to explore implementing LVL 1 sums digitally in ROD's.
    - Transmit digital sums to LVL1
    - LVL1 accept signals connect to ROD's instead of FEB's
    - Would require discussions with LVL1 groups
  - Explore different system level architecture: AdvancedTCA
    - High Availability: Redundancy and shelf management features built into it's specification.

ROD data rates

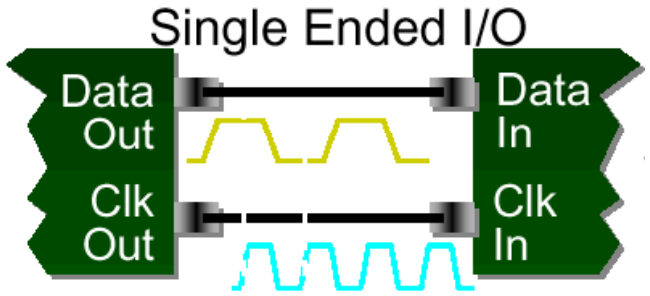
- Assuming a new FEB architecture which would digitize and transmit all samples from 128 channels at 40Mhz.

- 128channels x 40Mhz x 16bits = 81.92 Gbits/sec.
  - 98 Gbits/sec with 8b/10b encoding overhead
  - Entire LAr : 1524 FEB * 98 Gbits/sec = 150 Tbits/sec

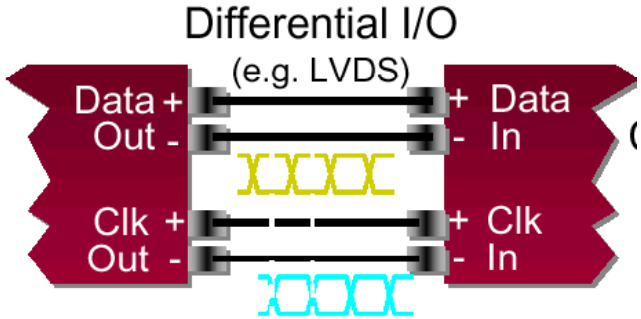
- Must be split among multiple fibers. Assume using an industry standard parallel fiber connector MPO/MTP, which has 12 fibers and is about 15mm wide.

- If 12 fibers per FEB were used, the data rate on each fiber would be 6.825Gbits/sec (8.2Gbit/sec with 8b/10b coding)

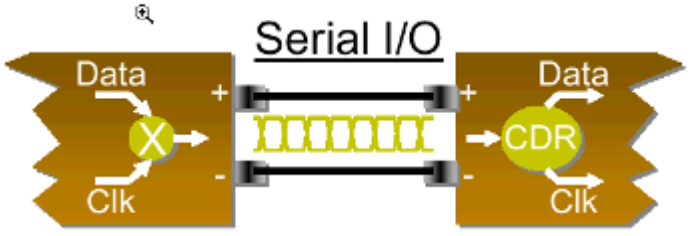
Interconnect Technology Overview



- Multi-Drop Busses
  - Noise limited above ~200Mhz
  - VME, PCI



- Switched fabrics (still parallel)
  - Source synchronous clk
  - clock skew limited above ~1Ghz
  - Rapid I/O, HyperTransport



- Serial Switched fabrics
  - Eliminates traditional noise and clock skew issues. 10Gbps
  - PCI-Express, Infiniband
  - SERDES (SERializer-DESerializer)

Parallel Fiber Optic Link Receiver

Reflex Photonics: 40 Gb/s SNAP 12 Parallel
Fiber Optic Receivers

- 12 independent parallel optical channels
- Mechanical size : 49mm x 17mm x 11mm
- Channel data rate of up to 3.5 Gb/s
    - 42 Gb/s per module
    - 10 Gb/s per channel in development
- Low power comsumption < 1W per module
    - No heat sink required
- Drop in compatible with SNAP 12 MSA connector
- Both 62.5um and 50um multi mode ribbon fibers supported
    - 100m range with 62.5 um
    - 200m range with 50 um
- Individual channel fault monitoring
- $425 ea (single piece pricing)

SERDES options

- Commercial receiver
  - Broadcom 1 channel 10 Gbps input : 16- 622Mbps outputs
  - Vitesse VSC7123: 4 channel 1.36Gbps
  - Mindspeed: 8 channel 2.5Gbps
- Custom receiver
  - OptoElectronics Working Group
  - GBT
- FPGA based receiver
  - Xilinx : Virtex 4 RocketIO
  - Altera : Stratix II GX
    - 622-Mbps to 6.5Gbps transceiver rates
    - Lots of programmability for compliance with wide range of standards and protocols
      - PCI Express, OC-192, 10Gb Ethernet, Serial RapidIO
    - Devices available with 4, 8, 12, 16, or 20 high-speed serial transceiver channels providing up to 255 Gbps of serial bandwidth

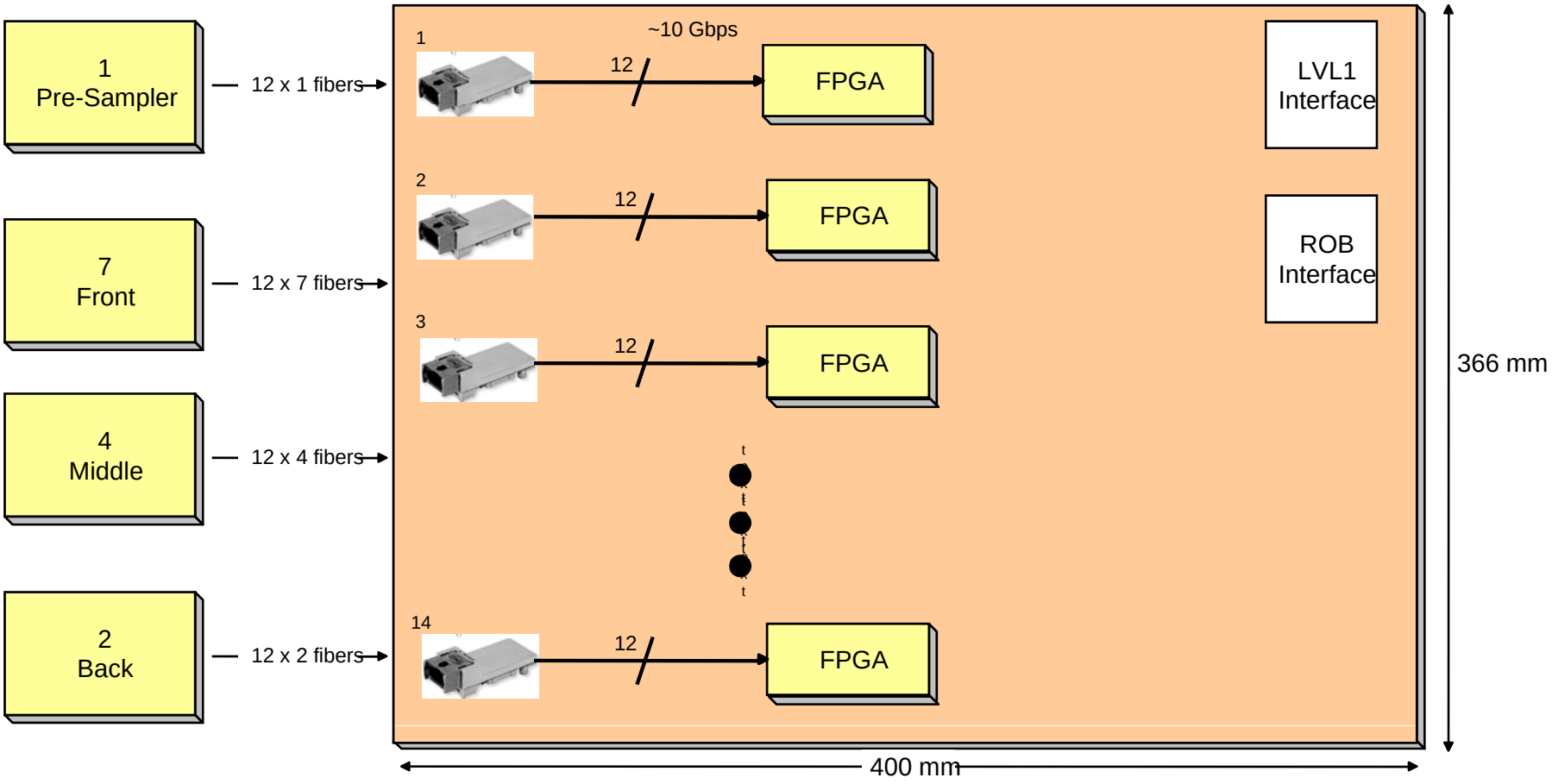
ROD using FPGA SERDES

(14 FEB's per ROD)

Input data rate = 1.1Tbps
168 fibers

FEB (1524 modules)

ROD (108 modules)



Xilinx and Altera High Speed Transceivers

Virtex-4 FX FPGA Device/Package Combinations

Feature/Product	XC 4VFX12	XC 4VFX20	XC 4VFX40	XC 4VFX60	XC 4VFX100	XC 4VFX140
EasyPath Cost Reduction Solutions	--	XCE 4VFX20	XCE 4VFX40	XCE 4VFX60	XCE 4VFX100	XCE 4VFX140
Logic Cells	12,312	19,224	41,904	56,880	94,896	142,128
Block RAM/FIFO w/ECC (18 kbits each)	36	68	144	232	376	552
Total Block RAM (kbits)	648	1224	2,592	4,176	6,768	9,936
Digital Clock Managers (DCM)	4	4	8	12	12	20
Phase-matched Clock Dividers (PMCD)	0	0	4	8	8	8
Max Differential I/O Pairs	160	160	224	288	384	448
XtremeDSP™ Slices	32	32	48	128	160	192
PowerPC Processor Blocks	1	1	2	2	2	2
10/100/1000 Ethernet MAC Blocks	2	2	4	4	4	4
RocketIO Serial Transceivers	0	8	12	16	20	24
Configuration Memory Bits	5,017,088	7,641,088	15,836,464	22,262,016	35,122,240	50,900,352
Max SelectIO	320	320	448	576	768	896

Virtex 4 RocketIO Transceivers
170mW @ 6.5Gbps

Stratix II GX Transceivers
225mW @ 6.375Gbps

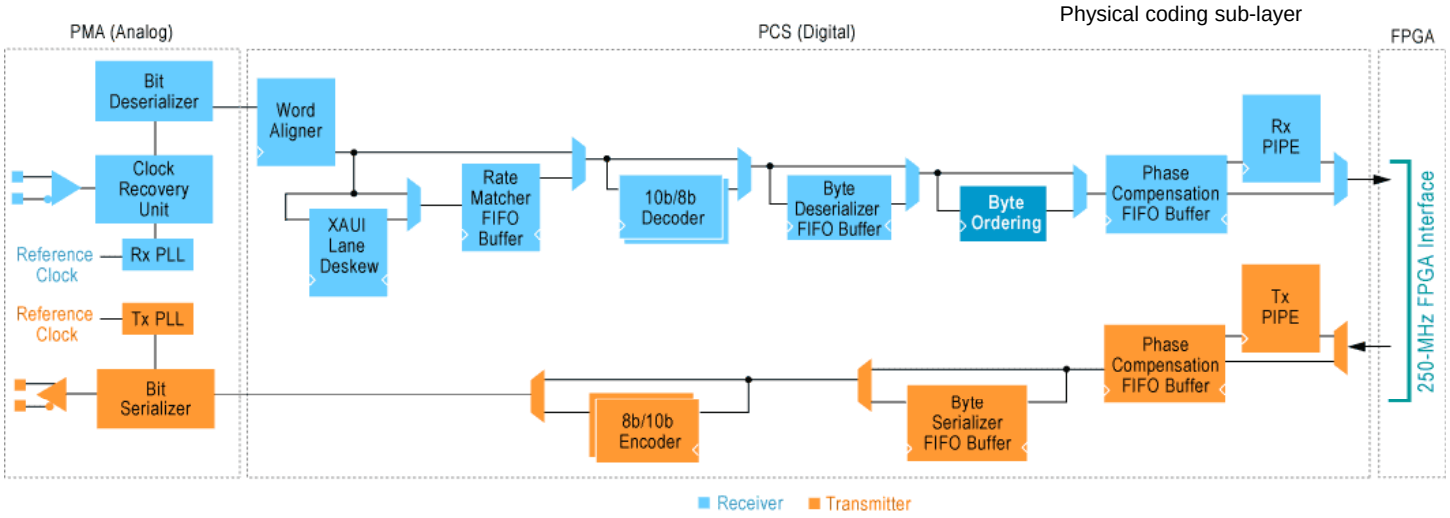
Table 2. Stratix II GX Device Features

Device	Transceiver Channels	Equivalent Logic Elements	Total Memory Bits	18-Bitx18-Bit Multipliers (1)	PLLs (2)	Availability (3)	
						General Engineering Samples	Production Devices
EP2SGX30C	4	33,880	1,369,728	64	4	-	Now
EP2SGX60C	4	60,440	2,544,192	144	4	-	Now
EP2SGX30D	8	33,880	1,369,728	64	4	-	Now
EP2SGX60D	8	60,440	2,544,192	144	4	-	Now
EP2SGX60E	12	60,440	2,544,192	144	8	-	Now
EP2SGX90E	12	90,960	4,520,448	192	8	Now	Now
EP2SGX90F	16	90,960	4,520,448	192	8	Now	Now
EP2SGX130G	20	132,540	6,747,840	252	8	Now	Now

$720
100pcs

Stratix II GX SERDES

Physical Medium Attachment



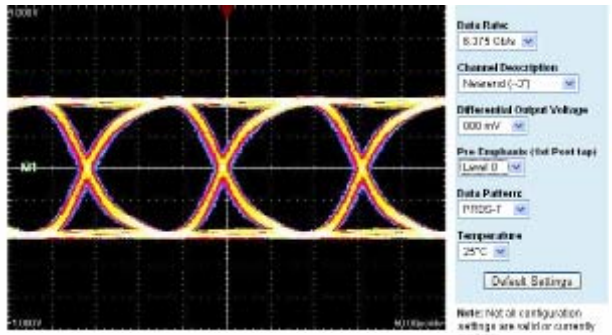
Stratix II GX Eye Diagram Viewer

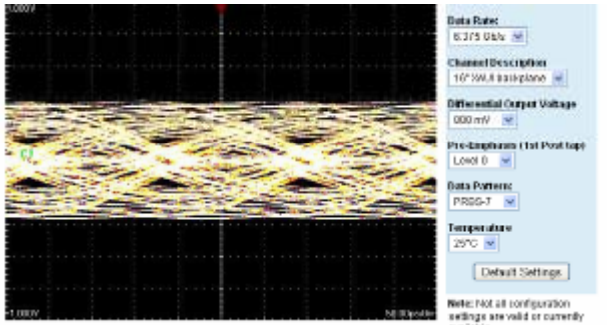
Programmable on-the-fly Pre-Emphasis (transmitter) and Equalization (receiver) provide effective compensation for channel degradation in low-cost FR-4 PCB's

Near End ~3"

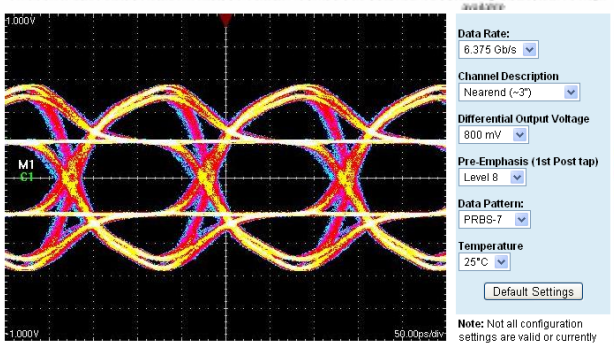
Far End ~16" (through 2 connectors)

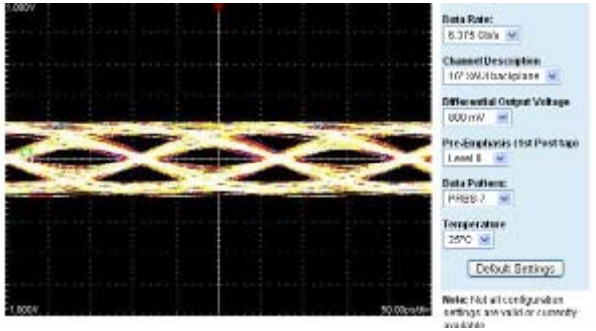
No Pre-Emphasis



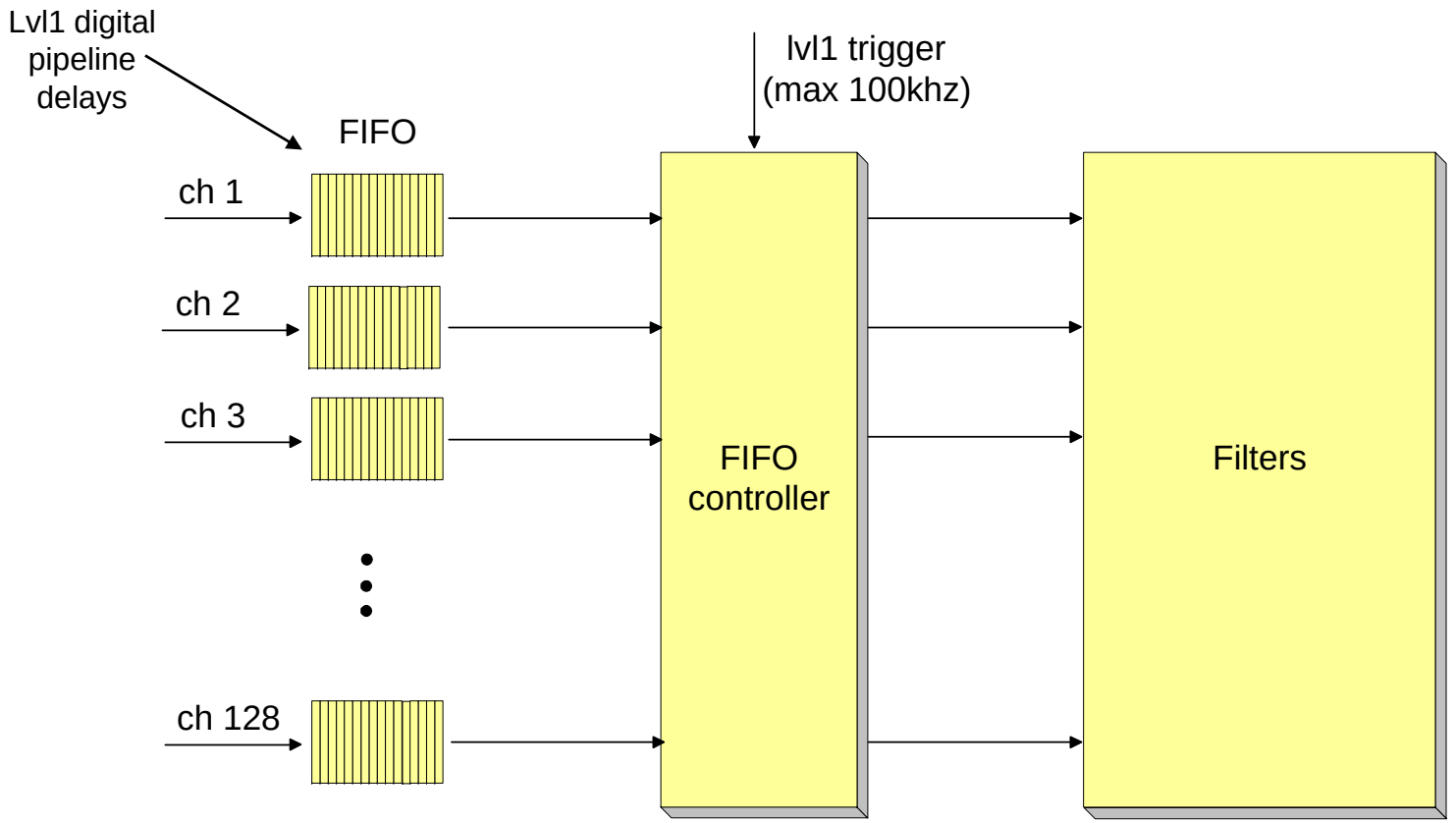


Level8 Pre-Emphasis





FPGA storage requirements



Length of FIFO's = 40Mhz * 2.5us = 100 * 16bits = 1.6kbits/channel

Total memory = 128 channels * 1.6kbits/channel = 204kbits

Modest Size FPGA has > 1Mbit memory

Energy & Timing Calculations

- Energy and Timing calculations performed at LVL1 rate of ~100khz
- Calculate energy for these data using optimal filtering weights:

$$E = \sum a_i \ (S_i - PED) \qquad i=1,..,5 \ \ PED=pedestal$$

- If E > threshold, calculate timing and pulse quality factor

$$\tau = \sum b_i \ (S_i - PED)$$
$$\chi^2 = \sum (S_i - PED - Eg_i)^2$$

- Histogram E, τ, χ^2

- Look at using FPGA solution which allows a more parallel implementation over traditional DSP's.
- Use dedicated FPGA DSP blocks (Multiply-Accumulate blocks)

GMACS

Billions of Multiply and Accumulate operations per second

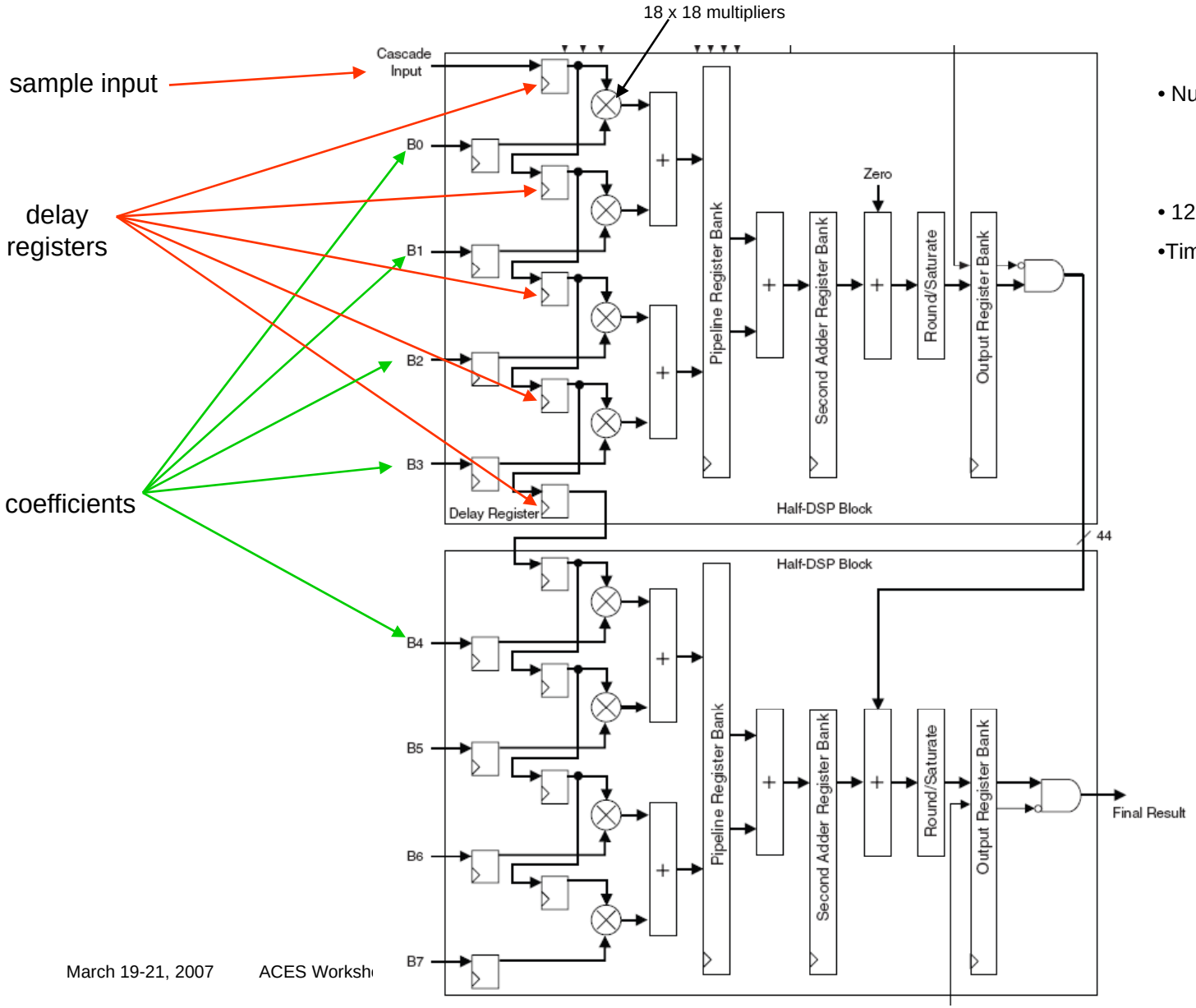
Table 1. Comparison of a Stratix III FPGA to Leading Performance DSP Processors

Device Type	Stratix III EP3SL70	TI C64x
Clock Frequency	300 MHz	1 GHz
Number of Devices	1	10
Total GMACS	86	80
Total Power	~3 watts	~10–15 watts
Total Board Area	~1.225 mm^2	~ 5.000 mm^2
Total Cost	$400 (1 x $400 @ 1 K units)	$3000 (10 x $300 @ 1 K units)

Energy & Timing Calculations

Altera Stratix III DSP Block

$E = \sum a_i \ (S_i - PED)$



- Number of DSP Blocks
  - EP3SL50 27
  - EP3SE110 112
- 128 channels processed by FPGA
- Time multiplex channels
  - 40 Mhz sample rate
  - 550 Mhz DSP block

ROD Level 1 Sums

- Currently, sums are analog and performed in the FEB crate.
- Each trigger tower requires 60 analog summations (EM barrel)
  - 4 channels Pre-sampler
  - 32 channels Front
  - 16 channels Middle
  - 8 channels Back
- Benefit to doing digital summation in ROD?
  - could provide more flexibility?
    - refined granularity?
  - However, would require discussions with LVL1 groups
- Latency considerations: Increase in LVL1 delay probably can't be avoided because of extra fiber optic lengths

Level 1 trigger latency

TDR Latency

Calorimeter trigger	**56.1**	
Signal processing and cables up to input of trigger preprocessor	20.6	Including worst-case 60 m cables to USA15
Preprocessor (e/γ, τ/h)	15.0	
Preprocessor (jet,E_T)	17.0	
Electron/gamma finding	14.0	
Tau/hadron finding	14.0	
Jet finding	18.0	
Missing E_T	18.5	
Total scalar E_T	18.5	
CTP	**4.0**	
TTC	**3.1**	Includes cable from CTP
Fibres to FE electronics	**16.0**	Using worst-case 80 m fibres
TTC receiver (in FE electronics)	**3.0**	
TOTAL	**82.2**	

Cable delay would increase to ~100m + 8.5

Digital Summation + 4 ?

System Level Considerations

- System Monitoring
  - histogram and sample data downloads
  - need reasonable throughput
- Configuration Download
  - loading FPGA code
  - loading filter coefficients
- Calibration
  - Special modes of operation to handle calibration, coefficient loading, etc.
- System Health
  - monitoring of crate voltages, temperature, fans, etc.
- Scalability
  - Fast low-latency communication between ROD's if level 1 sums are computed digitally and if sums span more than a single ROD

Advanced Telecom Computing Architecture (ATCA)

- Developed by The PCI Industrial Computer Manufacturer's Group (PICMG) An industry consortium that has standardized many popular standards such as ISA and PCI technologies for industrial backplane applications.

- PICMG 3.0 (ATCA)
  - High Availability, redundancy built into everything: power supplies, fabrics, etc. Boards are designed for hot swap.
  - Multi-gigabit serial transport (no parallel busses)
    - choice of protocols: GigE, Infiniband, PCIe, Serial Rapid I/O
  - Shelf Manager provides intelligent diagnostics, watches over basic health of system
  - Large form factor and power budget
    - 8U cards
    - 200W per slot, 3kW per chassis
    - 16 boards per crate
  - Relatively new platform, Potentially big market (Telecom). Some projections of $20 Billion.

Advanced Telecom Computing Architecture

Integrated
Shelf Manager

User Defined
Connector
Area

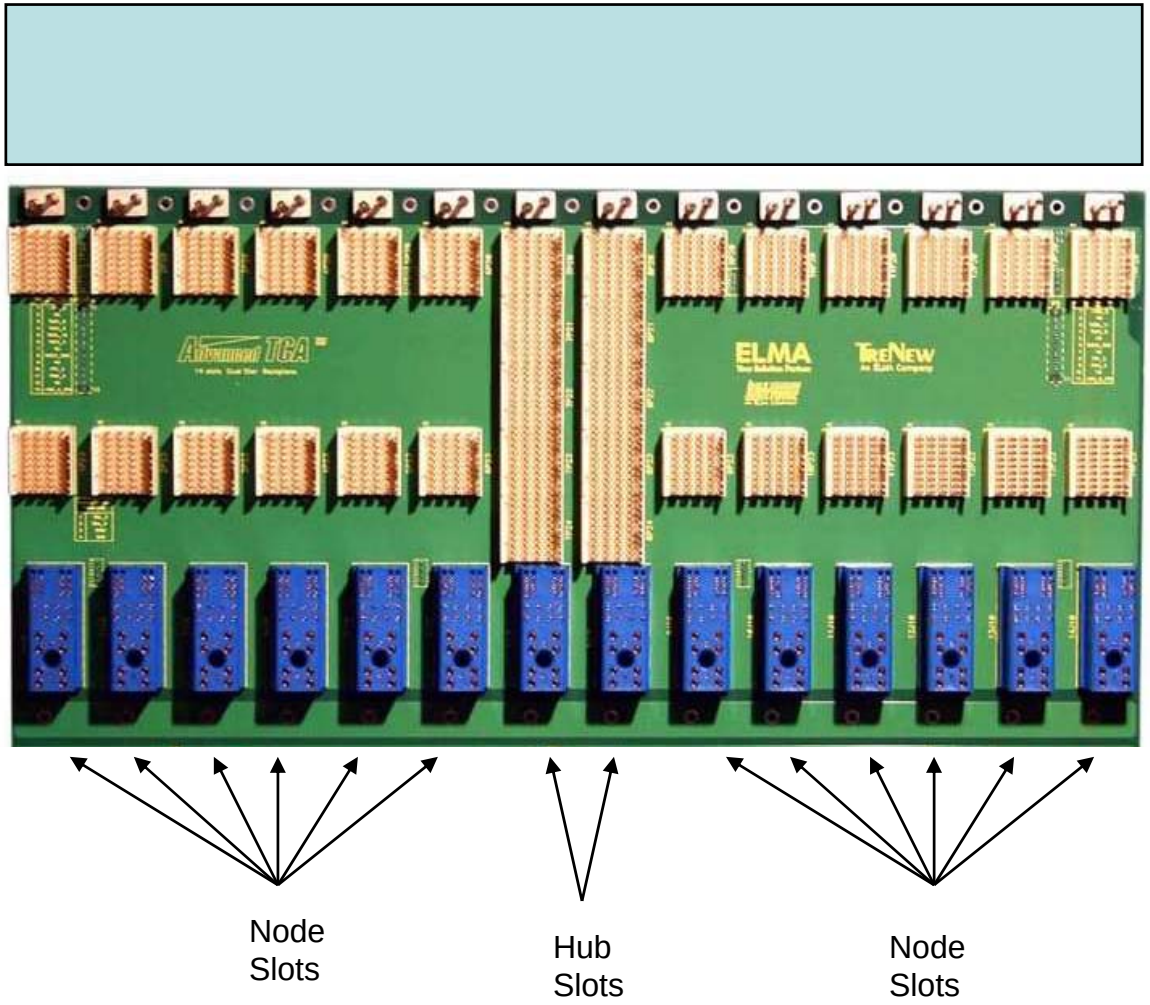
Front Module
Size
8U x 280mm

Transition
Module size
8U x 70mm

Air Intake Area

Redundant
Power Entry
Modules, -48v

ATCA backplane - Dual Star



Zone 3 Used Defined I/O area, Connection to Rear Transition Modules

Zone 2 Data Transport Interface. Base Interface: 10/100/1000 Ethernet

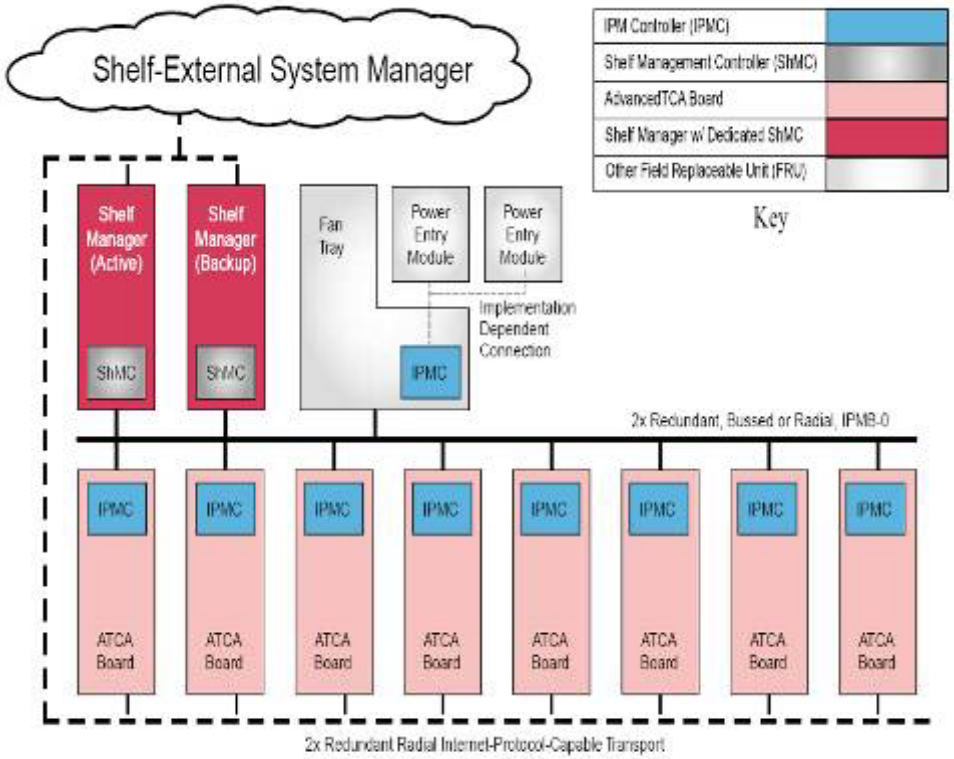
Fabric Interface: Ethernet, Infiniband, PCI Express, StarFabric

Zone 1 Power – Dual Redundant -48v

Shelf Management

ATCA Shelf Management

- Purpose of Shelf Management System is to monitor, control and assure proper operation of components.
- Based on Intelligent Platform Management Interface (IPMI)
  - Used in Computer Server environment
  - Standardized management architecture for component and chassis level elements
    - Simple physical interface I2C bus
    - Monitors board health, such as voltages, temp.
    - Controls system level fan speeds, power supplies.
  - Each ATCA board must provide an IPMC controller requires intelligence



PICMG 3.0 Shortform Specification

Plan of work for FY07

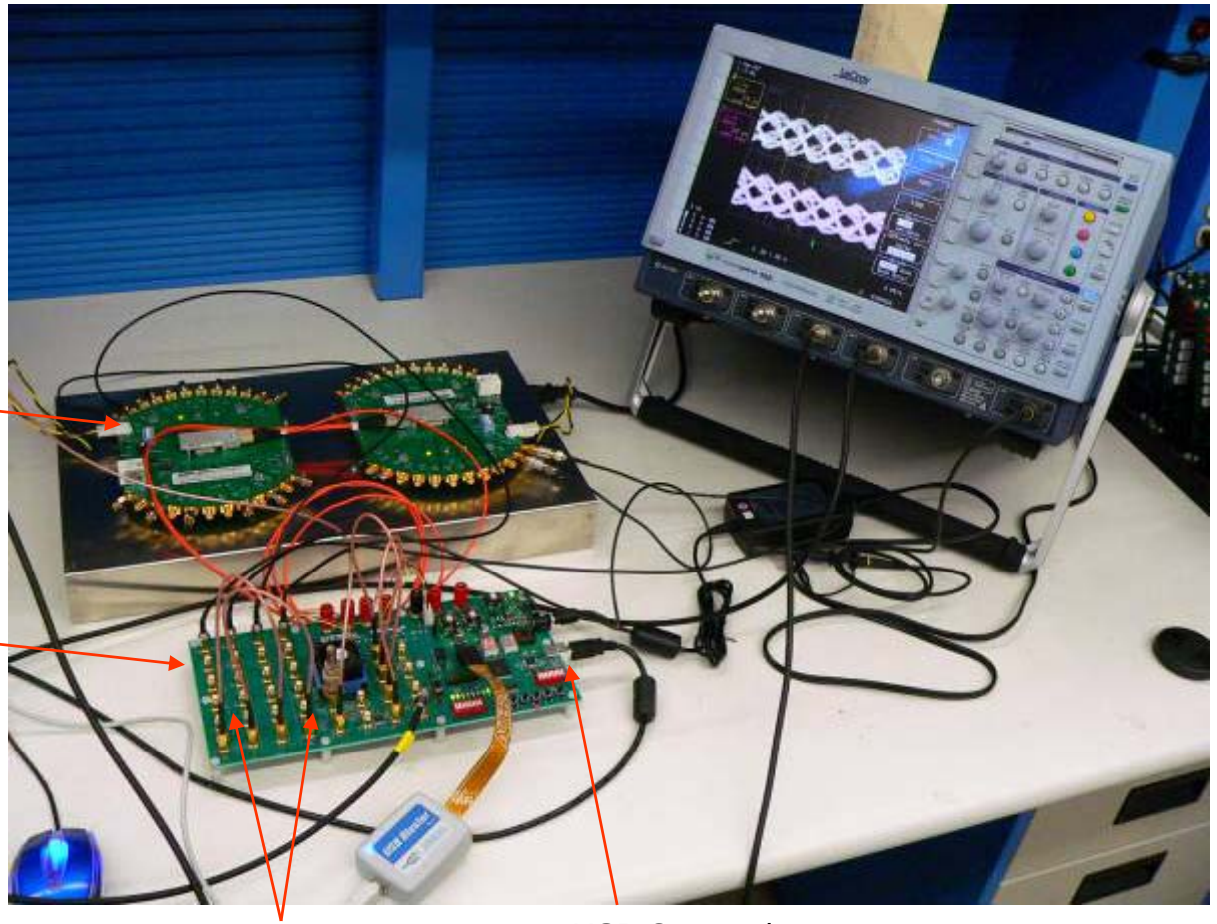
- Investigate multi-gigabit link with Xilinx and/or Altera FPGA SERDES
  - Use evaluation board to quickly learn features
- Investigate further the use of ATCA crates.
  - Start experimenting with ATCA crate and Xilinx ATCA Reference Board
  - purchase ATCA Hub Board with Integrated Shelf Manager
- Design a Sub-ROD module.
  - Process data from a single FEB.
    - 128 ADC Channels @ 40Mhz, 16bit
    - 1 Parallel Fiber with 12 SERDES on FPGA
  - Implement Energy and Timing calculations using FPGA.

Investigate multi-gigabit link with FPGA SERDES

University of Arizona: Ken Johns, Joel Steinberg

Reflex Photonics 12 channel transmitter and receiver evaluation board

Altera Stratix II GX Transceiver Signal Integrity Development Kit



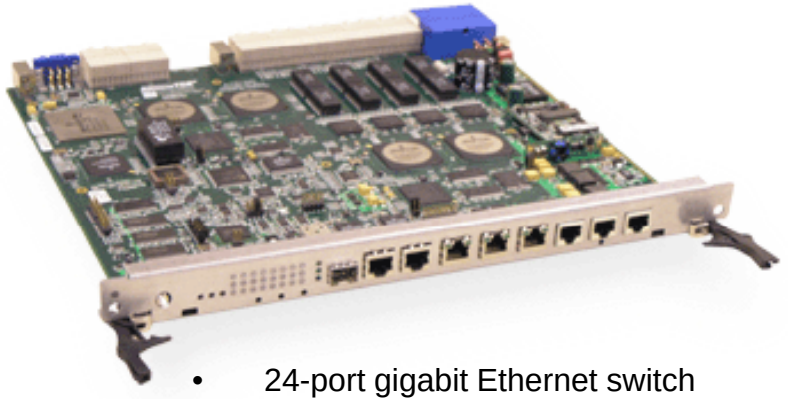
SMA Transmit and Receive Connectors

USB Connection to computer for changing SERDES configurations

More Investigation of ATCA

- Reference development board developed to demonstrate the use of the ATCA PICMG standard for high-speed networking and commnunications
- PICMG 3.0 Compliant
- Provides 15 channels of 2.5Gbps serial links to full mesh fabric backplane.
- Headers for application specific daughterboards
- Management firmware runs on the Virtex-II Pro's PowerPC processor running an embedded Linux OS

- 14 slot ATCA Crate
- Learn Shelf Management Features
- Use base interface to send configuration info to ROD
- Use fabric interface to send ROD data out of the shelf.



- 24-port gigabit Ethernet switch
- 1 port to each of the 14 node slots
- Communicates to each ROD board which has Ethernet TCP/IP interface
- Built in Shelf Management Features

Develop Sub-ROD module

- Virtex-4 ML403 Embedded platform ($495)
- Embedded PPC 405
  - Linux as OS, having OS permits easier intelligent controller development (interface to shelf manager), proven TCP/IP stack, etc.
- 10/100/1000 Ethernet Port
  - Start using Ethernet as this is used as the base interface on ATCA
- 64 MB DDR SDRAM, 64Mb Flash, IIC EEPROM
- 64 bit User Expansion Connector .

- Custom Daughterboard
  - Parallel Fiber Optic Link
  - FPGA with 12 on-chip SERDES
  - Interface with ML 403

Participating Institutions (EOI)

- U.S.A.
  - Brookhaven National Laboratory
    - Hucheng Chen, Joe Mead, Francesco Lanni
  - University of Arizona
    - Ken Jones, Joel Steinberg
  - Stony Brook University
    - Dean Schamberger
- France
  - LAPP, Annecy
    - Jacques Colas, Guy Perrot
- Italy
  - INFN, Milan
    - Mauro Citterio

Conclusions

- If new FEB design transmits all data to ROD without zero suppression
  - Huge data rates (~100Gbps / FEB)
  - Technology would need to exist to handle all those bits.
    - parallel fiber interfaces currently at ~40 Gbps. Need to reach 100Gbps
    - FPGA SERDES currently running at ~6.5 Gbps. Need to reach 10 Gbps
- LVL1
  - Possible to implement LVL1 sums digitally in ROD's
    - possible benefits would be more flexibility
    - feedback with LVL1 groups would be required
- Energy & Timing calculations
  - Technology already exists
  - All FPGA approach would reduce size and power needed
- System Level
  - Investigate ATCA crate
    - High Availability features: redundancy, shelf management protocols