

A Reconfigurable Cluster Element (RCE) DAQ Test Stand for the Atlas Pixel Detector Upgrade

Rainer Bartoldus, Mike Huffer, Martin Kocian,
Emanuel Strauss, Su Dong, Matthias Wittgen (SLAC)
Erik Devetak, David Puldon, Dmitri Tsybychev (Stony Brook)

September 22

TWEPP 2010, Aachen

Introduction

- SLAC R&D project for a generic DAQ/trigger technology based on generic building blocks
 - ATCA packaging standard
 - Reconfigurable Cluster Element (RCE) processing boards
 - Cluster Interconnect Module (CIM)
 - Rear transition modules for custom user I/O
- Aimed at accommodating the common needs of a wide range of DAQ applications with hardware used by
 - LSST (Large Synoptic Survey Telescope)
 - LCLS (Linac Coherent Light Source)
 - Atlas R&D
 - PetaCache (a large scale, random access, high performance, storage system)
- Use of recent technology developments in telecom industry combined with modern FPGAs/system-on-chip
- The proposed technology is already deployed in test bench applications for ATLAS pixel detector Insertable B-Layer (IBL) project and verified its viability for the IBL readout system for 2015, and potentially also the CSC muon readout upgrade for 2013

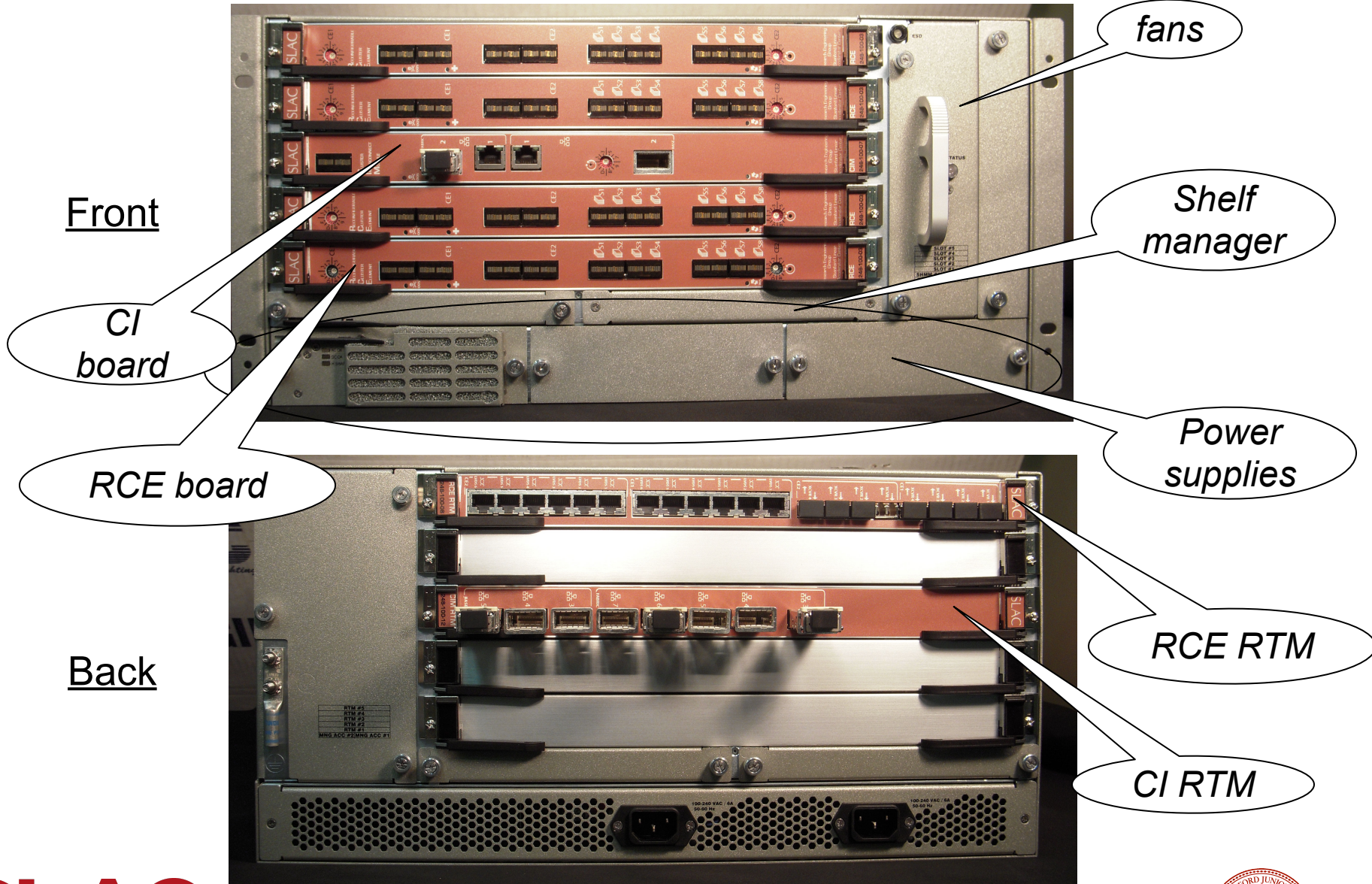
Possible Atlas Upgrade Scenario

- Possible Atlas upgrade for higher luminosity in two phases
 - Phase one ~2015 with a luminosity $\sim 2 \times 10^{34}$
 - Phase two ~2019 sLHC luminosity of $\sim 5 \times 10^{34}$
- RCE for ATLAS is targeting any DAQ upgrade needs from 2013 onwards
 - RCE proposed for ATLAS pixel and CSC muon readout in near term
 - Sufficient capacity for major overhaul of front-end electronics, trigger and DAQ for long term upgrade
- Current favored pixel upgrade option is an additional insertable B-layer
 - New pixel in preparation: FE-I4 for IBL
 - Current pixel DAQ needs to be modified to cope with increased read-out speed of FE-I4 (previous generation: FE-I3)
 - RCE is one path for phase one upgrade

ATCA packaging standard

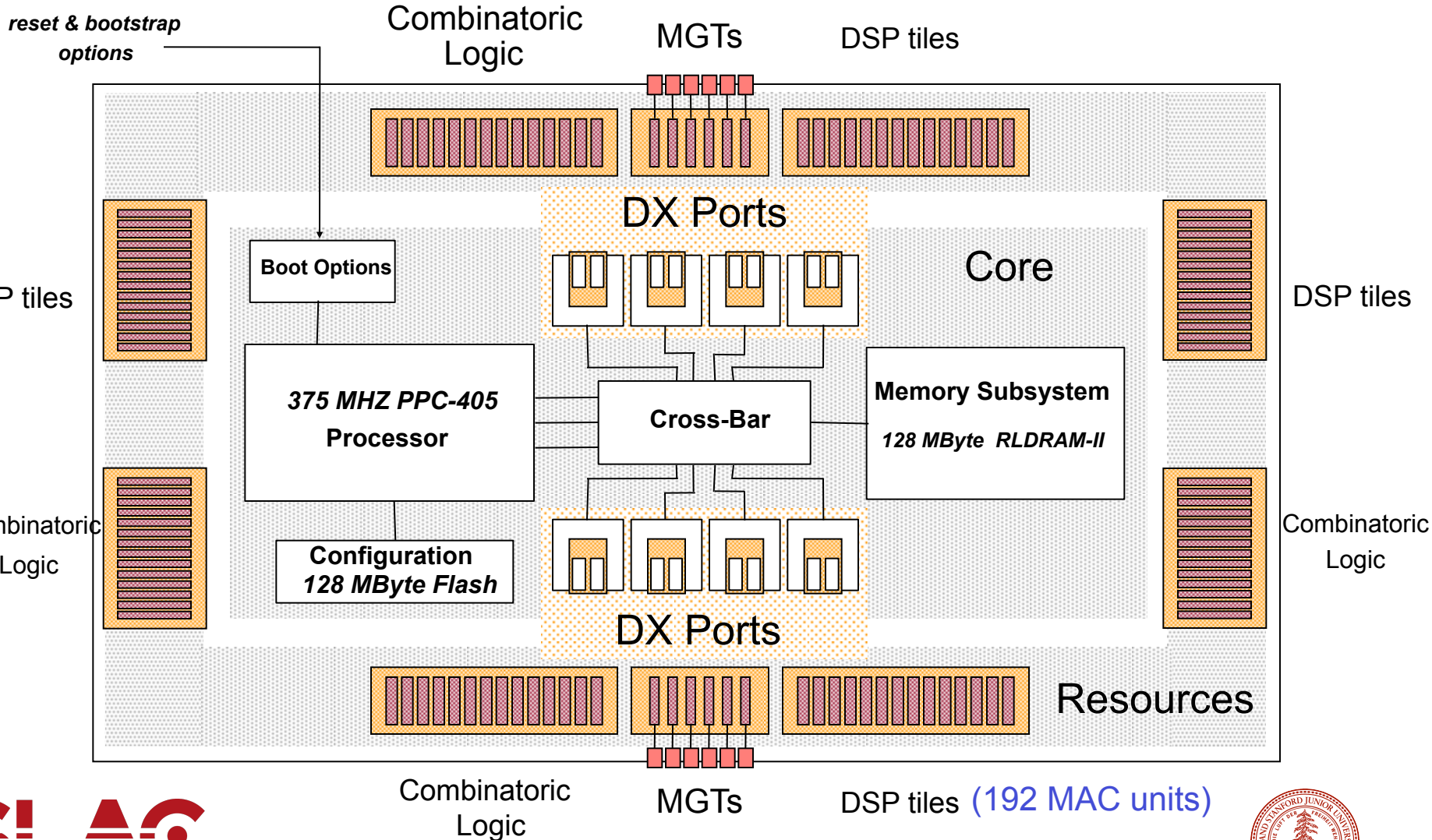
- Advanced Telecommunications Computing Architecture
- Emerging standard in telecommunication
- Attractive features
 - Backplane and packing available as commercial solution
 - Hot swap capability
 - Well-defined environmental monitoring and control
 - Emphasis on High Availability
 - External power input is low voltage DC
 - Allows for aggregation of rack power
- Shelf supports Rear Transition Modules (RTM)
 - RTM and front board interconnected by user-defined connector (zone 3)
 - Allows all cabling on rear - facilitates re-cabling, board swap
- high speed serial backplane with 4.5 TBit/s
 - Protocol agnostic
 - compared to 40 MByte/s for VME

Typical (5 slot) ATCA crate



Reconfigurable Cluster Element (RCE)

Current implementation On Virtex-4 FPGA

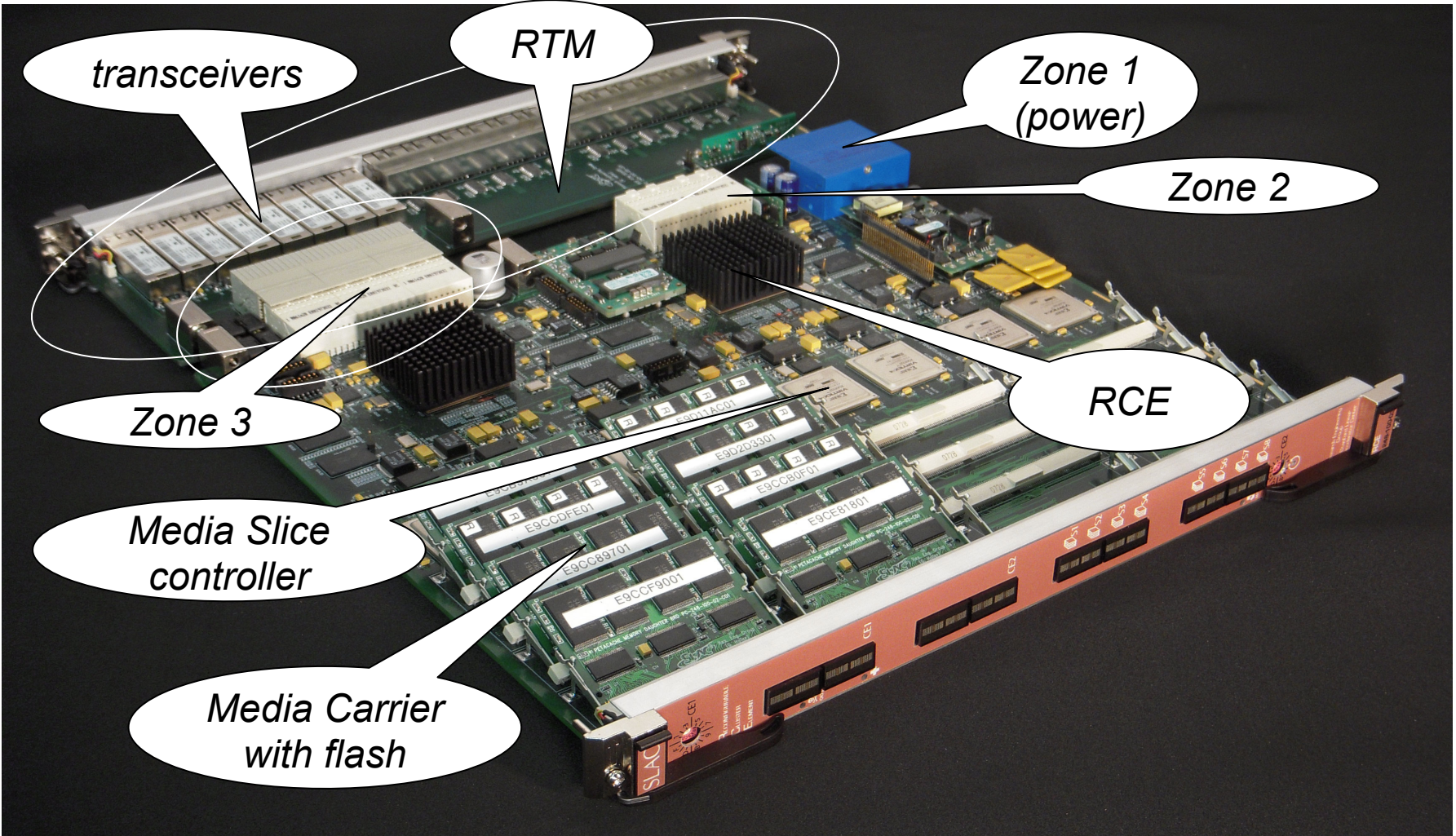


RCE Resources

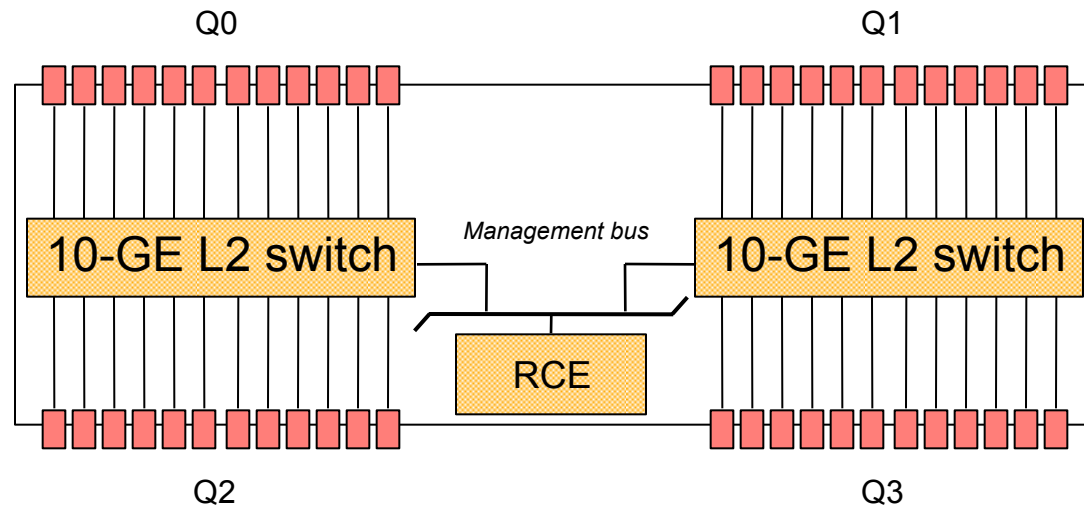
- Multi-Gigabit-Transceivers (MGTs)
 - Up to 12 channels of
 - Serializers/deserializers
 - Input/output buffering
 - Clock recovery
 - 8b/10b encoder/decoder
 - 64b/66b encoder/decoder
 - Each channel can operate up to 6.5 Gbit/s
 - Channels may be bound together for greater aggregate speed
- Combinatoric logic
 - Gates
 - Flip-flops (block RAM)
 - I/O pins
- DSP tiles
 - Up to 192 Multiple-Accumulate-Add (MAC) units

RCE Board

Board shown here with 1TB FlashRAM for PetaCache project

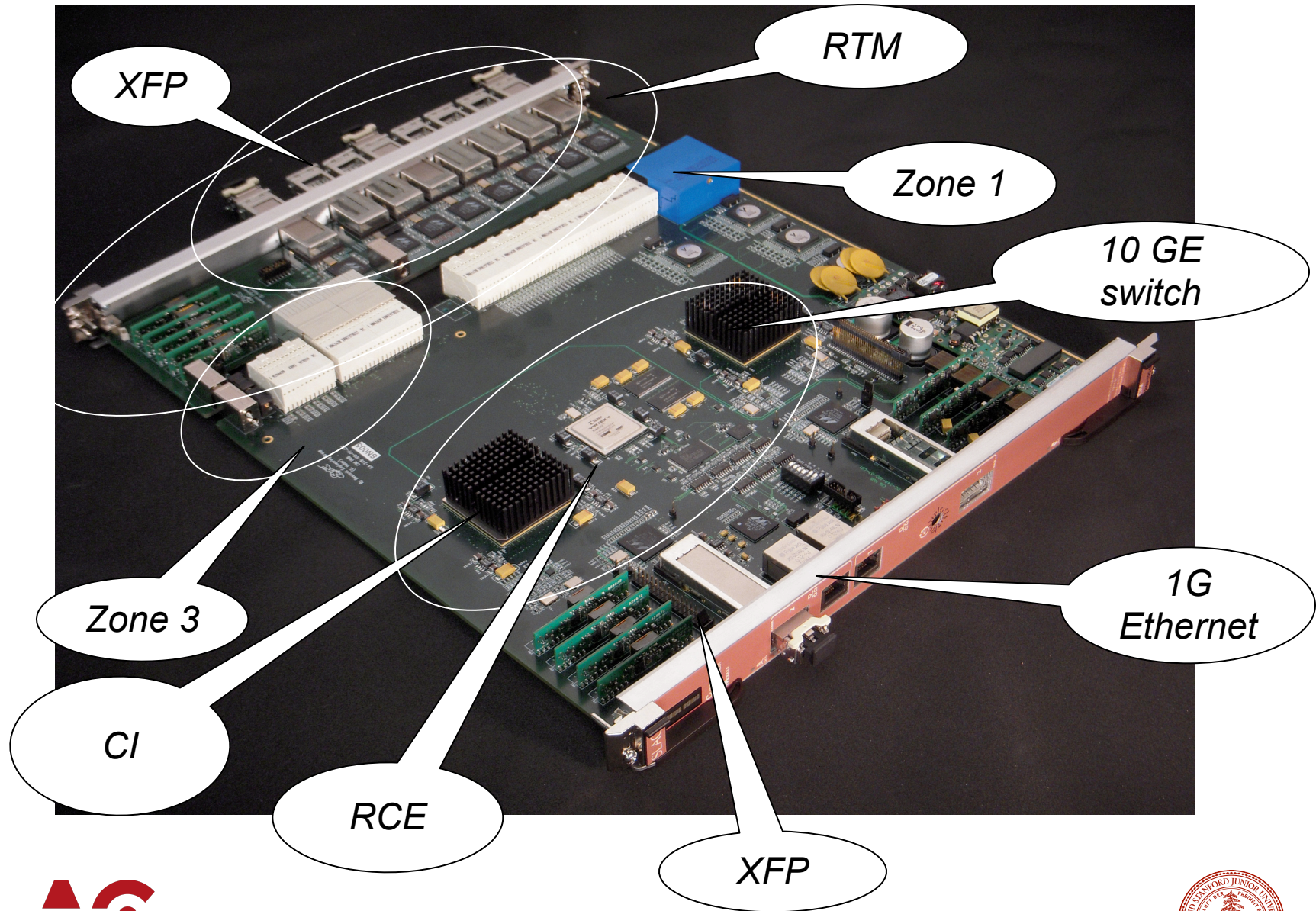


The Cluster Interconnect (CI)



- Based on two Fulcrum FM224s
 - 2x 24 port 10-GE switches
 - Is an ASIC (packaging in 1433-ball BGA)
 - XAUI interface (supports multiple speeds including 100-BaseT, 1-GE & 2.5 GBit/s)
 - Less than 24 watts at full capacity
 - Cut-through architecture (packet ingress/egress < 200 NS)
 - Full Layer-2 functionality (VLAN, multiple spanning tree etc..)
 - Configuration can be managed or unmanaged

Cluster Interconnect board + RTM



Generic RCE Software Infrastructure

- The RCE concept is an integrated hardware/firmware/software infrastructure to allow easy user interface
- RTEMS is used as real-time operating system
 - RCE itself is OS agnostic (Linux available)
 - Open source (real-time Linux is commercial)
 - POSIX compliant API
 - Exception handling
 - BSD TCP/IP network stack + common internet services like DHCP, NFS, telnet, etc
 - Board support package for PowerPC405 was developed at SLAC
- Cross platform tools
 - GNU cross compiler tool chain (ships with RTEMS)
 - Remote (network) GDB
- Object oriented framework in C++
 - Hardware configuration interface
 - Plugin support

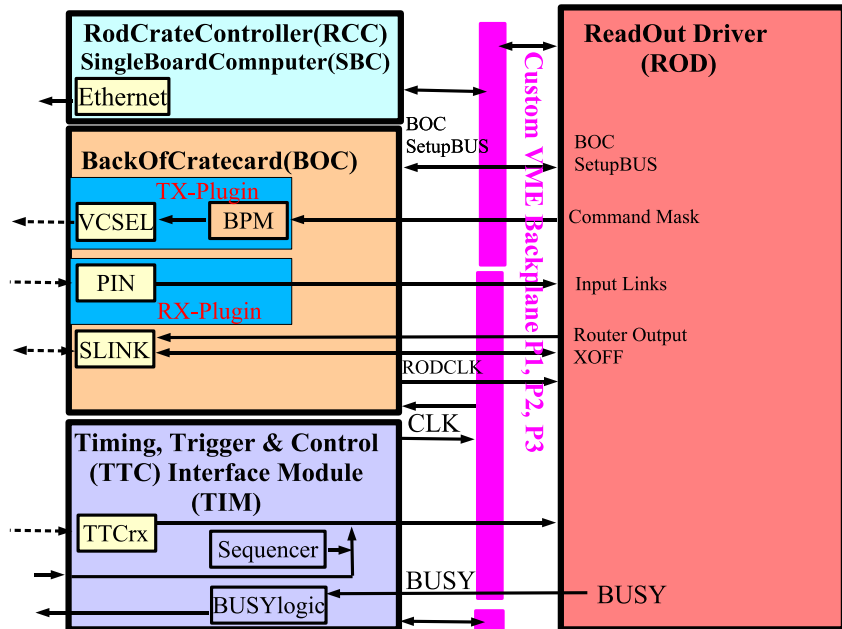
Generic RCE Software Infrastructure

- Network console with interactive, extensible login shell
 - Logging facilities
 - Monitoring commands for memory, network, tasks, etc.
- Core application started at boot from flash memory
 - User code loaded from shared library building on core
 - Task can be loaded and started from network
- No need for Xilinx tools for generic RCE code development
 - Covered by remote network debugger, shell + task starter

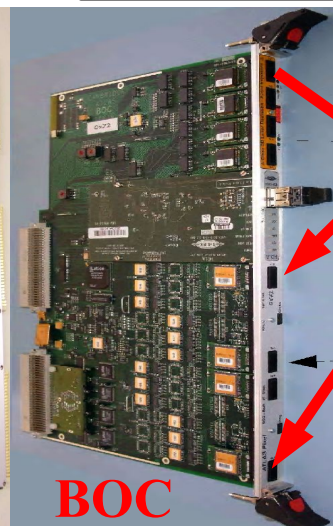
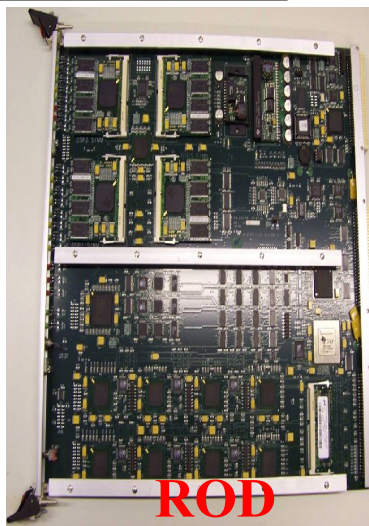
Current Atlas Pixel Readout

- Atlas pixel detector consists of 1744 modules
 - 46080 channels per module
 - One module combines 16 front-end pixel chips (18 columns x 160 rows)
 - Modules send data with 40 and 80 MBit/s depending on location
- VME based DAQ processing boards
 - Read-Out-Driver (ROD) + Back-of-Crate Card (BOC)
 - ROD hosts 4 + 1 DSPs
 - 1 TI C6201@166MHz fixed-point CPU to drive modules + DAQ setup
 - 4 TI C6713@220MHz floating-point CPU for online monitoring + calibration processing each with 256 MB of external SDRAM and 64KByte of CPU cache memory running at CPU clock speed
 - 1 GFLOP per floating-point DSP
 - Plus FPGAs for event fragment building, data formatting
- Optoboard converts electrical to optical
 - For test stands electrical BOC available
- BOC encodes clock/data, decodes in 40MHz streams for ROD

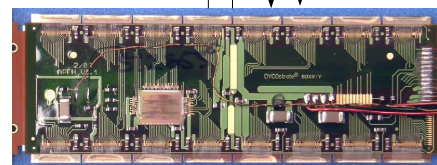
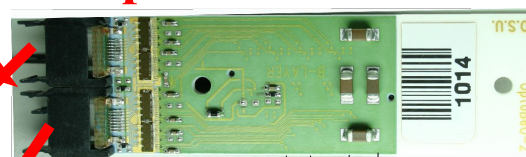
Pixel Readout



- Different data paths for DAQ
 - From via SLINK to upstream event building and higher level trigger
- and calibration data
 - Through DSPs via VME and single-board computer to calibration framework



OptoBoard



Some Current ROD Design Features

- Pixel chip calibration as application example
- DSPs on ROD for processing and histogramming calibration data
 - One ROD serves up to 28 modules
- Example calibration: finding thresholds for each pixel
 - Inject charge with capacitor into pixel
 - 100 data points for obtaining S-Curve
 - Mean and sigma determined by fit on DSP
 - RODs perform calibration in parallel
 - But: up to 16 RODs in one VME crate
 - All calibration data have to go through SBC and one ethernet connection
 - Typical data amount: ~23GB per ROD reduced to 13MByte after fitting
 - Observed VME download speed much less than 40MByte/s
 - Data reduction essential - raw calibration can not realistically read out

Advantages of an RCE Calibration System

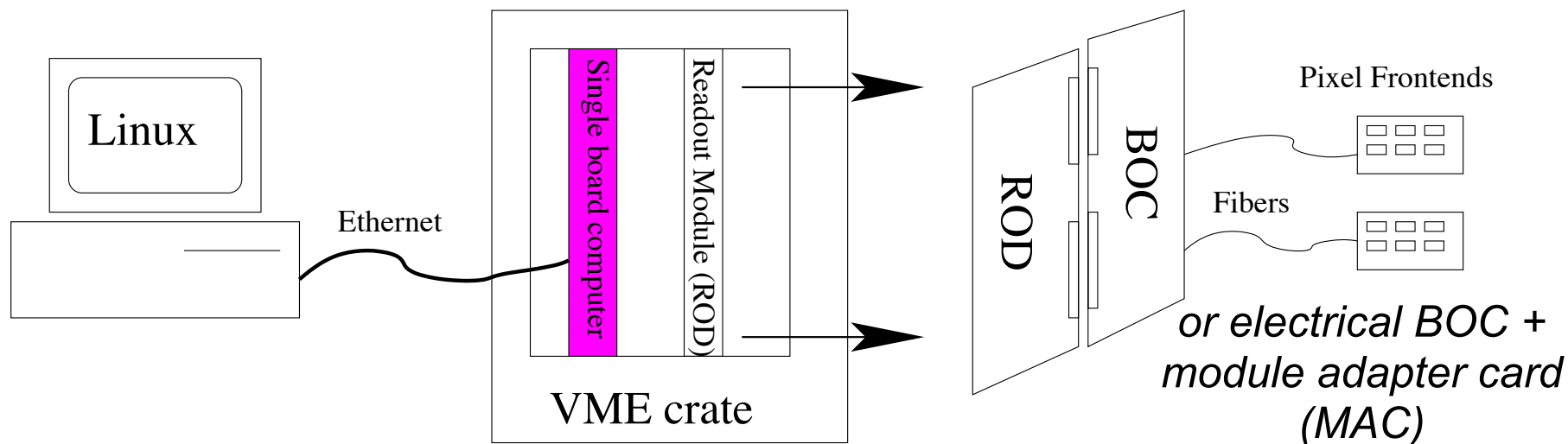
- 10 GBit ethernet data output path removes VME download limitations
- 8 x more memory available per pixel for calibration in next generation RCE boards compared to current ROD
- Flexible and extendible framework was written for RCE
 - Easy to add new calibrations or new hardware like new pixel front-end-chips chips (FE-I4) for IBL upgrade
- PowerPC environment easier to program than the DSP
 - DSPs do not run operating system - limited debugging capabilities
 - Pseudo object oriented C-code with limited modularity
 - Complex memory management necessary for DSP to make use of fast (very limited) internal memory
 - Make as much use of general purpose CPU for data processing rather than FPGA programming
 - RCE offers additional computing resources for highly parallel algorithm like S-curve fitting in form of DSP tiles

RCE Calibration Framework

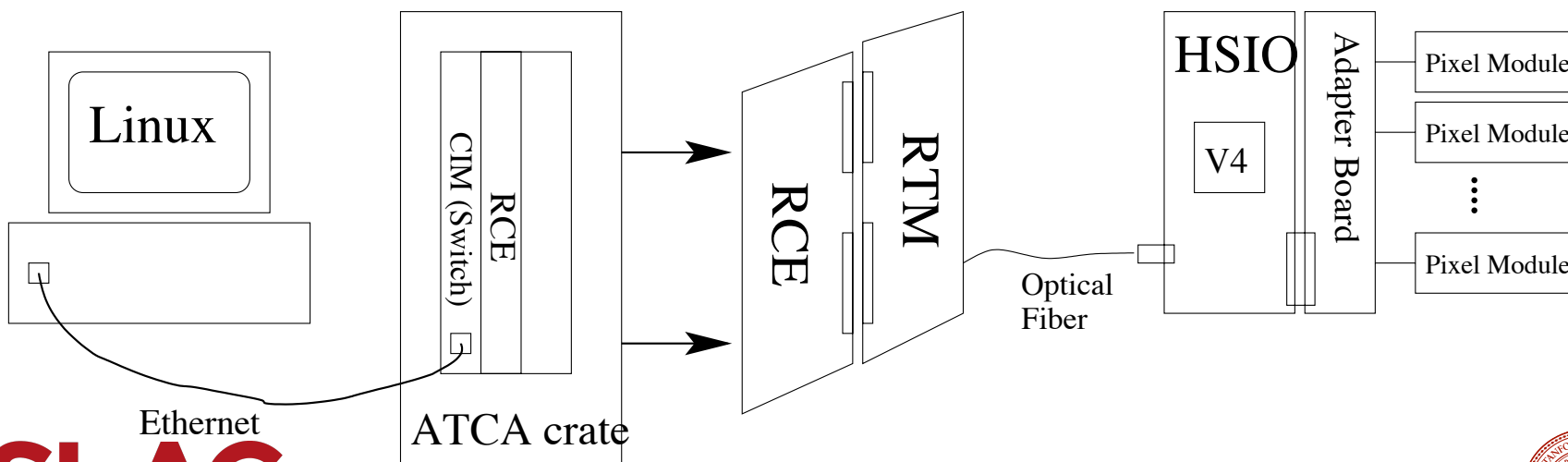
- Modular C++ code
- Completely redesigned framework for calibration scans
- Known good algorithms/code from DSPs reused for RCE implementation
- PowerPC 405 (no FPU)
 - Not a limitation for calibration application
 - FPU undesirable in real-time environment
 - FPU available through Auxiliary Processor Unit (APU)
 - Processing and fitting code was ported to fixed-point representation
- Successfully integrated RCE into Atlas TDAQ infrastructure
 - Process control and communication software ported to RTEMS

Application of RCE to Pixel Calibration

VME based Test Stand



RCE based Test Stand

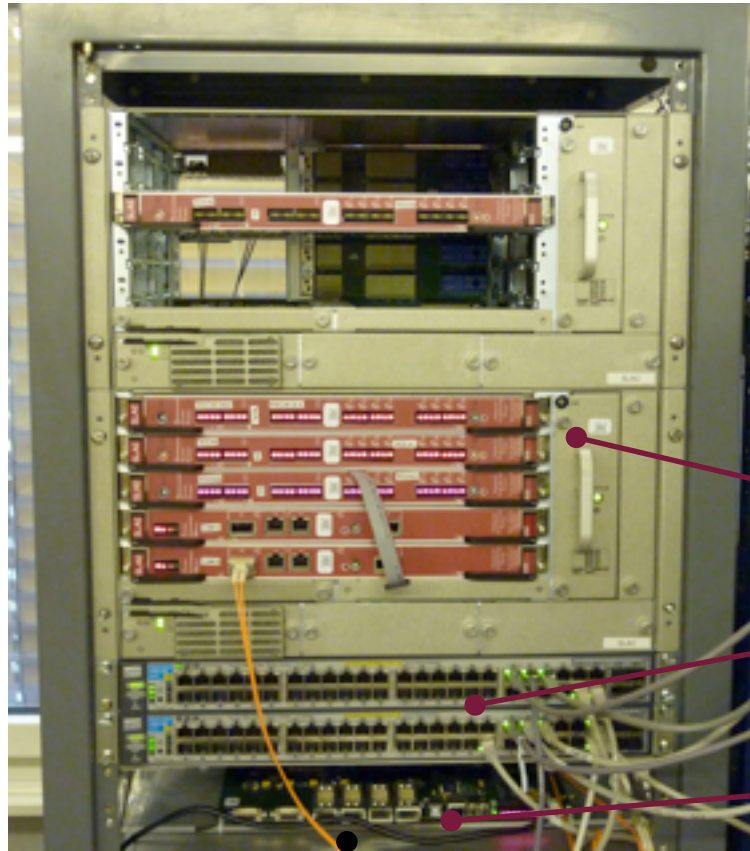


Multi-Channel Readout Board

- Realized as RCE + High Speed I/O (HSIO) Board
- HSIO is a generic DAQ board for a large number of signal types
 - Virtex-4 FPGA for processing
 - Combination with RCE allows fast progress of application creation
 - RCE RTM is connected to HSIO board by 3.125 Gbit/s fibers
 - Up to 8 channels for pixel chip readout
- Existing VDHL firmware infrastructure exists in form of the Pretty Good Protocol (PGP) for RCE and HSIO communication
 - Data rate 2.5 Gbit/s
- Read-out speed for FE-I3 pixel chips is 40 MBit/s for calibration
 - Serial bitstream
 - Commands to pixel module are sent with a 5 MBit/s bit stream
- Firmware upgrade operational to increase speed to 320MBit for FE-I4 testing

CERN RCE Test Stand

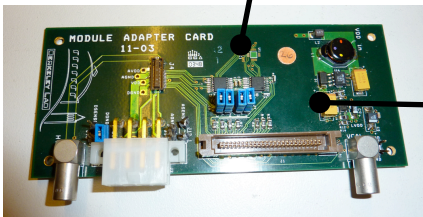
2 x Dell PowerEdge 2950 Servers
(infrastructure + SW development)



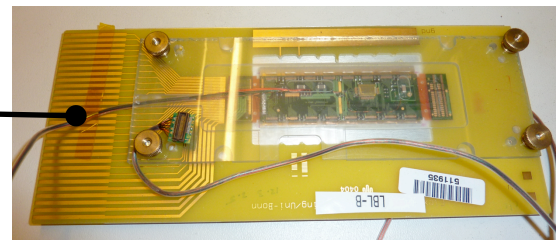
ATCA with RCE/CIM

2X HP ProCurve 3500 switches with 48 x 1
GE ports and 2 x 10 GE ports

HSIO Board



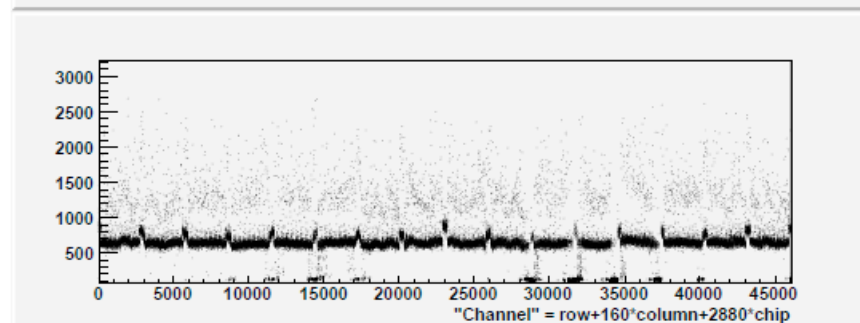
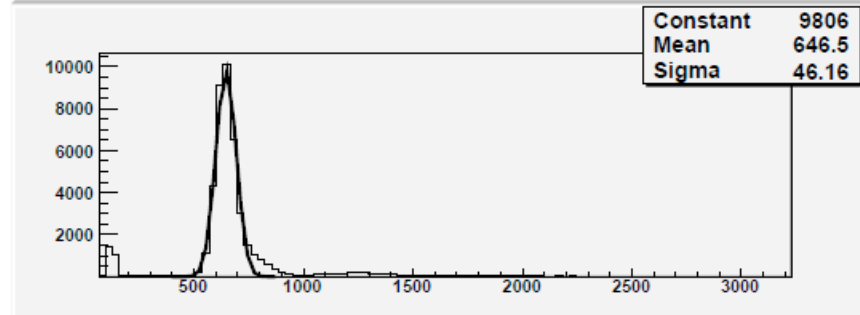
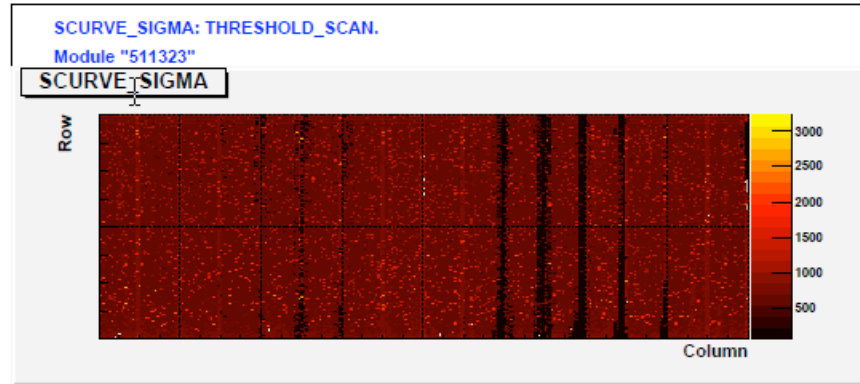
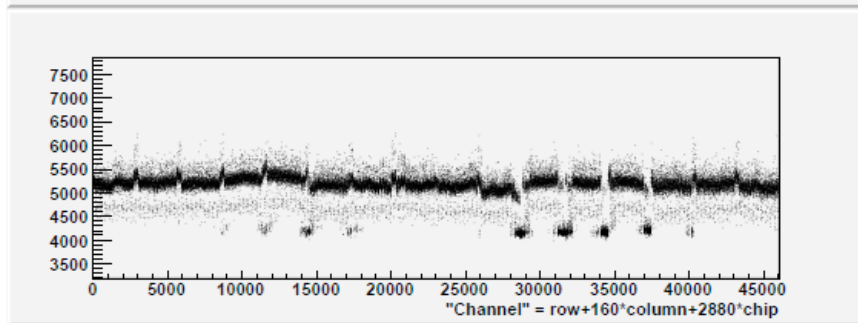
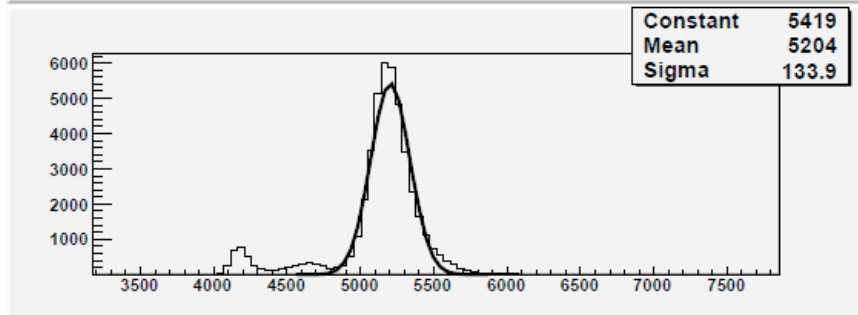
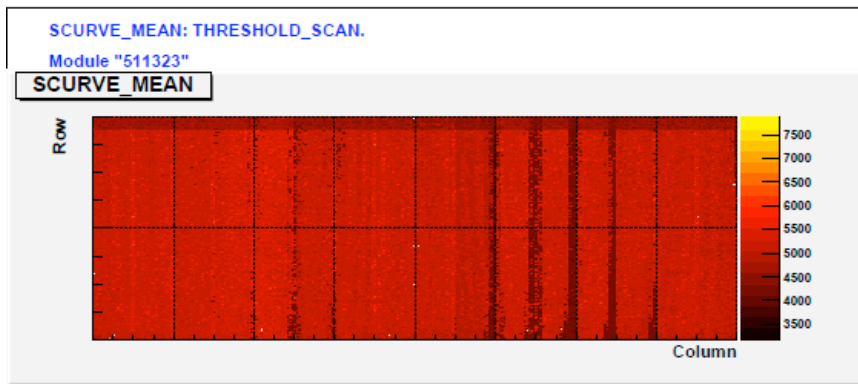
Module Adaptor Card



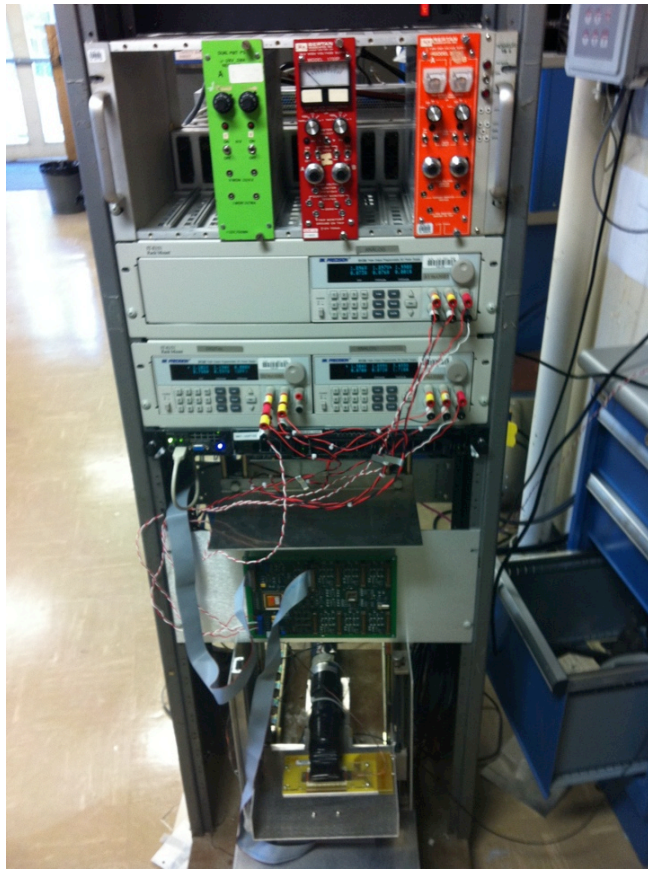
Pixel Module

Threshold Scan Calibration

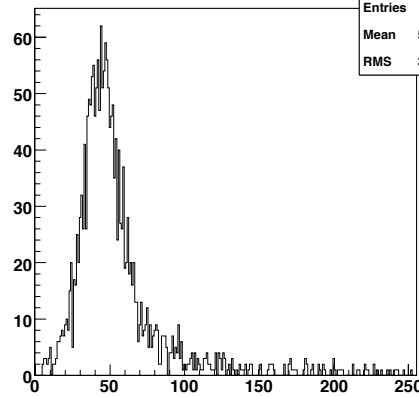
S-Curve Fit Results (mean, sigma) obtained from RCE
RCE readout is integrated into an existing test stand GUI



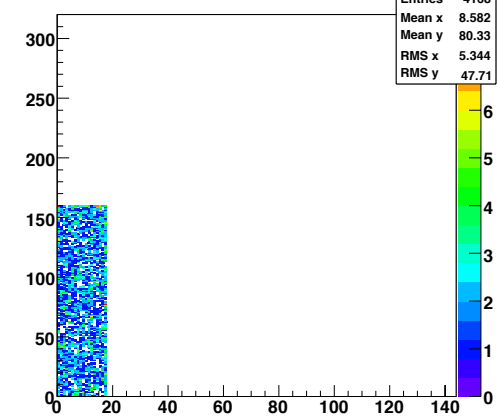
Cosmic Telescope at SLAC



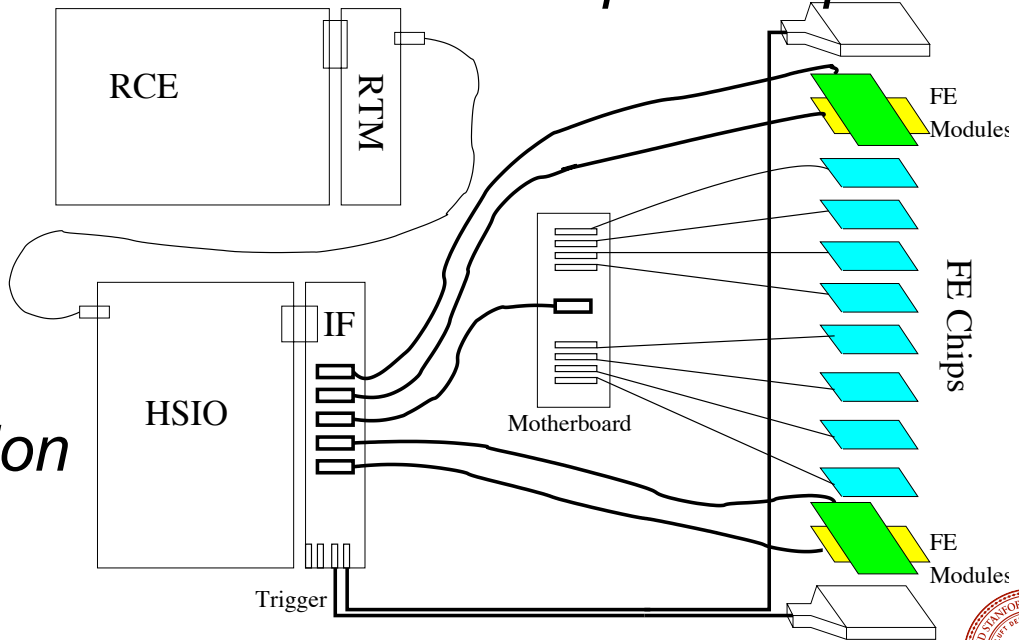
ToT module 5



Occupancy module 5



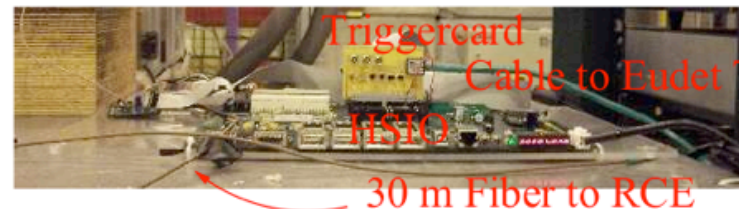
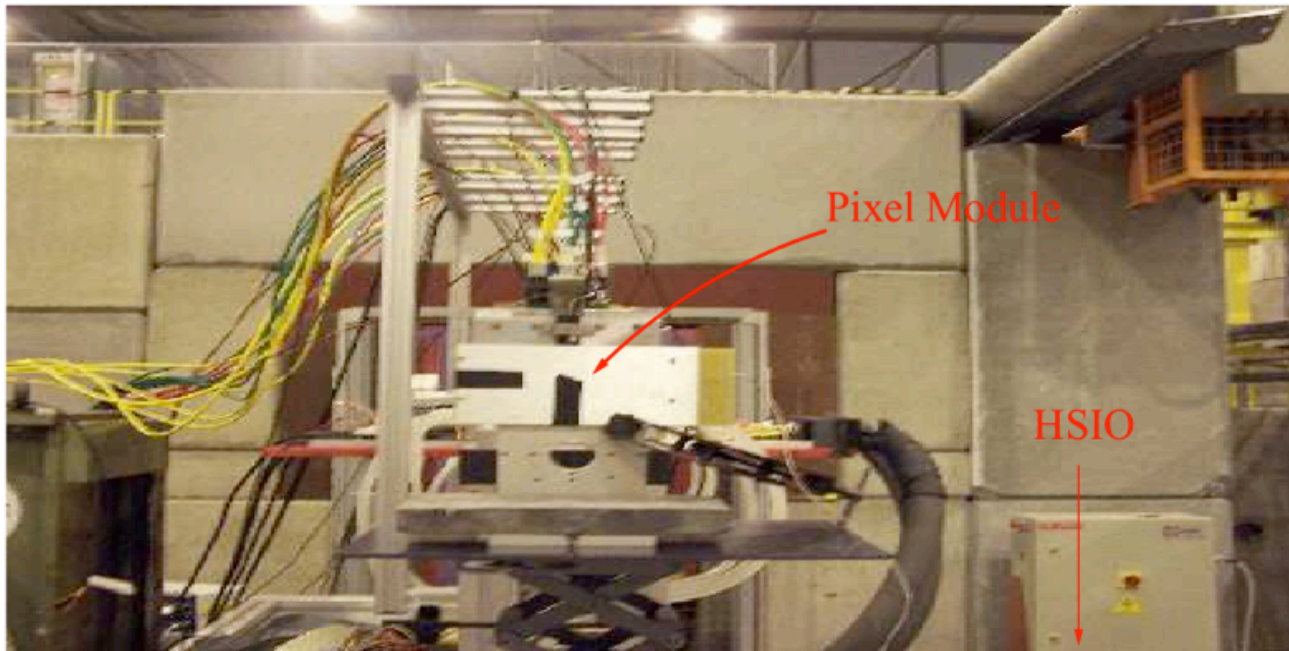
Cosmic Telescope Setup



*Triggered DAQ Application
ATCA in different lab*

CERN Test Beam Setup

- Small (portable) ATCA crate with RCE/CIM + one Linux box as infrastructure server
- Test beam setup with (current) FE-I3 pixel chip
- Pion beam for EuDet telescope



Project Time Scale and Milestones

Achieved Project Milestones

- October 2008: RCE/CIM proposed for Atlas luminosity upgrade
 - Work started on test stand for pixel modules
- June 2009 RCE proof of concept demonstrating read-out of FE-I3 pixel modules
- Spring 2010: relevant Atlas TDAQ software (IS, OH,IPC) ported to RTEMS
- June 2010 most calibrations implemented for FE-I3 on RCE
- Preparation for FE-I4 readout well underway

Next Milestones

- End of October 2010: plan to have full set of calibrations available
- October/November: ready for FE-I4 testing
- End of 2010: demonstrate concept for full integration of RCE platform into Atlas Pixel DAQ software
 - Seamless integration of calibration only, co-existence of ROD and RCE based systems
 - DAQ path needs next RCE generation with TTC interface

Next Generation RCE/CIM

- Generation 2 is based on Virtex-5 (estimated availability ~ spring 2011)
- CIM board eliminated. Functionalities distributed into one type of RCE board
 - Full mesh backplane
- New RCE/CIM board
 - 1.5 TBit/s switching
 - 96 serial input channels
 - 12 RCEs
 - RCE on mezzanine board
 - Switching between generation 1 (Virtex-4) and generation 2 (Virtex-5)
 - PPC440@450Mhz with APU support
 - More memory from 128MB to 2-4 GByte
 - Firmware framework for APU interfacing
- New bootstrap interface for remotely booting + enhanced diagnostics
- CERN compliant TTC timing interface
- S-link plugin for interface to existing DAQ infrastructure

Some Additional Resources

- RTEMs:
 - <http://www.rtems.org>
- ATCA Home Page:
 - <http://www.advancedtca.org>
- RCE/CIM project page at CERN (under construction):
 - <https://savannah.cern.ch/projects/rcedevelopment>

Acknowledgements

Many thanks to the RCE core development team

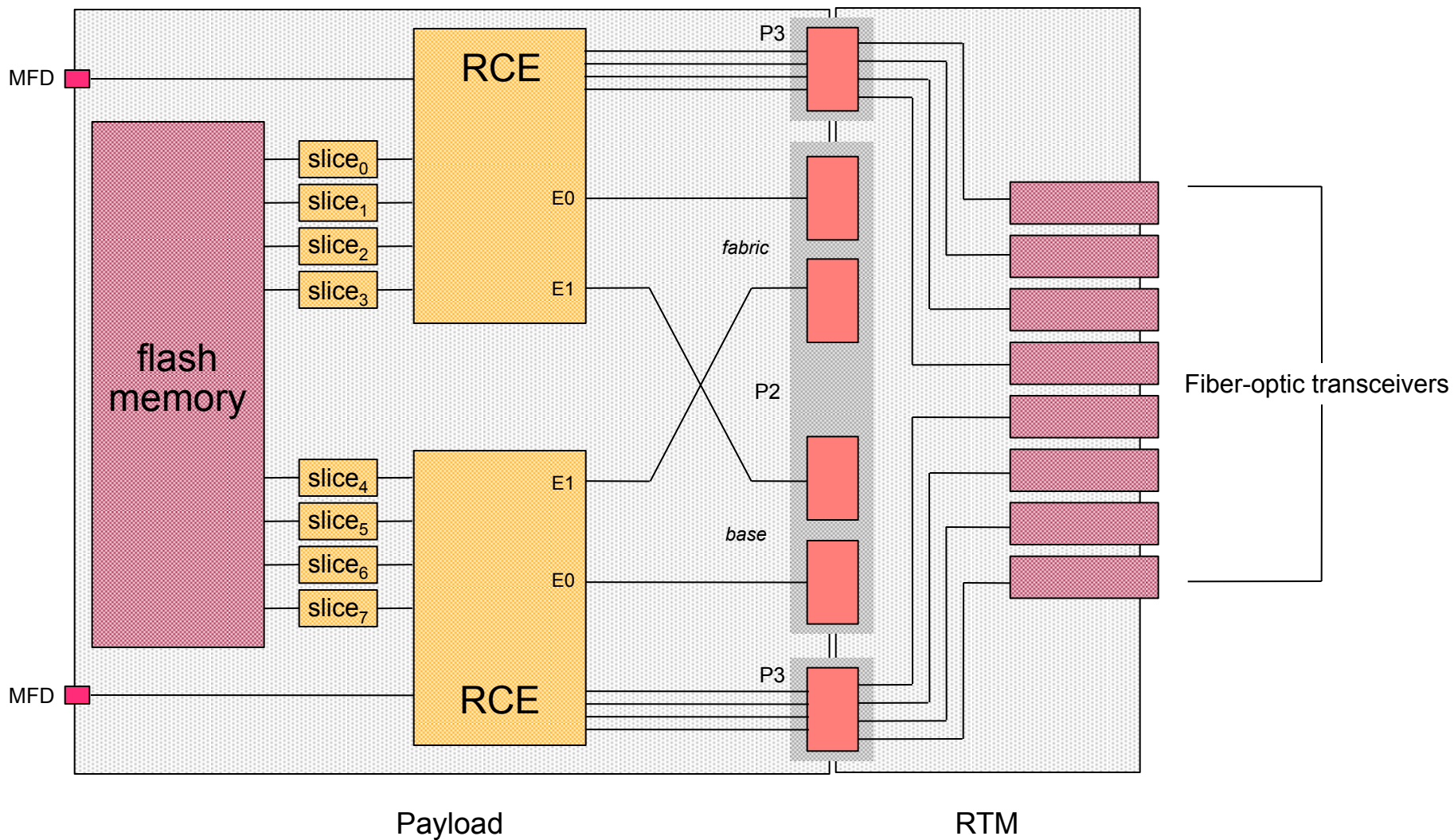
Mark Freytag
Gunther Haller
Ryan Herbst
Chris O'Grady
Amedeo Perazzo
Eric Siskind
Matt Weaver

Conclusion

- Based on analysis of previous DAQ systems the RCE/CIM was developed at SLAC
 - based on ATCA packaging standard
 - generic building blocks adaptable to a broad range of applications
- Generic software Infrastructure framework based on RTEMS is available for RCE development
- Parts of Atlas pixel ROD/DAQ software have been adapted to run on RCE
- Fully functional calibration test stands at SLAC and CERN have been set up
- Cosmic telescope at SLAC as example of RCE based DAQ system
- RCE was successfully used in CERN test beam
- Aiming for full, seamless integration of RCE into the Atlas Pixel DAQ system
- Much of the R&D are also common to potential other DAQ upgrade needs such as the muon system

Additional Slides

RCE Board and RTM



CIM and RTM

