Data Acquisition

Amedeo Perazzo
Photon Controls and Data Systems Online Manager

SLAC, November 12th 2008
LCLS FAC - X-Ray Systems Breakout Session



Contents

- **Requirements**
  - Configuration, calibration, acquisition, processing, monitoring, storage
- **Architecture**
  - Accelerator Data
  - FEE and DAQ nodes
- **Detector Interface**
  - Cornell and BNL detectors
- **DAQ Hardware Devices**
  - RCE and CIM

Configuration

- **The configuration software allows the operator to**
  - Create profiles of the detector
  - Upload these profiles to the various instruments before data taking starts
- **The configuration data describes**
  - Instrument settings
  - Output data types
- **Instrument configuration saved together with science data**
  - Always possible to correlate configuration with corresponding event data
- **Common PCDS data acquisition software provides base classes for**
  - Configuration objects
  - Science data types

Calibration

- **Same software used to create the configuration profiles for data taking is able to create the configuration settings used for calibration**
  - data taking: configuration remains mostly constant during a run
  - calibration: detector configuration changes often in between cycles
- **Calibration software allows calibration data files to be extracted from a calibration run**
  - Calibration files used during data taking and during the off-line analysis to
    - convert raw data to physical units
    - filter the image data (dark image accumulation, gain calculation, etc.)

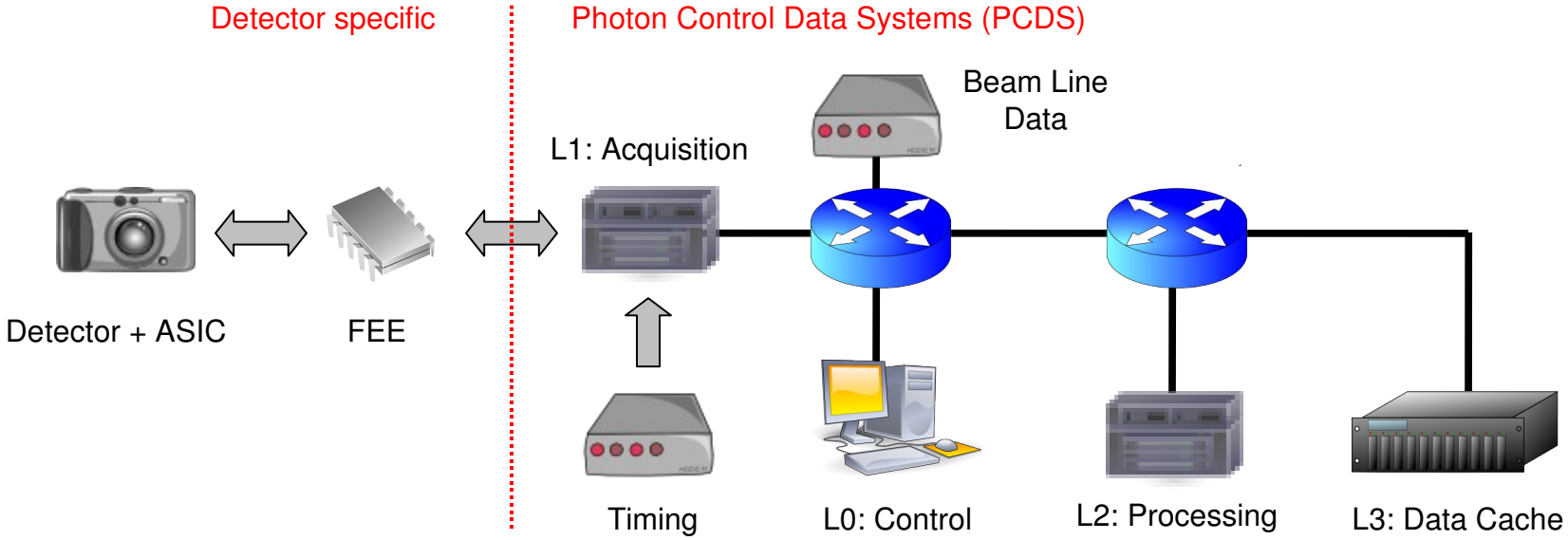
Acquisition & Processing

- **Each event labelled with a pulse ID for later reconstruction of the data**
- **Preprocessing envisioned during the acquisition (readout) stage**
  - Zero-suppression and bad pulse rejection
  - Simplify and speed up downstream data processing
- **Processing stage performs feature extraction on the data**
  - Statistical parameters different waveforms (single shot and averaged)
  - Centroid determination of the image data
  - Information collected at processing stage used for
    - Monitoring
    - Stored together with the data to simplify and expedite off-line analysis stage

Monitoring & Storage

- **Individual or summed spectra displayed in real-time**
  - User selectable refresh rate of up to 10 Hz
  - Time-frame of the display user selectable
    - i.e. last n shots, cumulative data set, etc.
  - Possible to zoom on the display
    - options for auto-scaling or user defined ranges
  - Storage and printing capabilities included for the displayed data
  - All data converted from hardware format to physical units before display
- **Data stored to data cache in custom object-oriented format**
  - Data from each individual shot will be uniquely identified, retrievable, and all software controlled and measured parameters correlated with the data
  - Likely more economical store data in hardware word format
    - potentially with data reduction through zero suppression, along with sufficient information to decode it later

Data System Architecture



- **Detector**
  - Experiment specific
  - May be bump-bonded to ASIC or integrated with ASIC
- **Front-End Electronics (FEE)**
  - Provide local configuration registers and state machines
  - Provide ADC if ASIC has analog outputs
  - FEE uses FPGA to transmit to DAQ system

Accelerator Data

- **Timing interface**
  - Timing boards (EVR) receive 120 Hz timing information from accelerator timing generator board (EVG)
    - 1Gb/s optical fibers connect EVG with EVR through custom protocol
  - EVR distributes timing signals to L1 nodes and FEE
- **Beam Line Data**
  - Time-stamped 120 Hz beam quality data information
  - Information contained in raw Ethernet packets
    - low latency network
    - UDP multi-cast
  - Used by L1 nodes to veto events

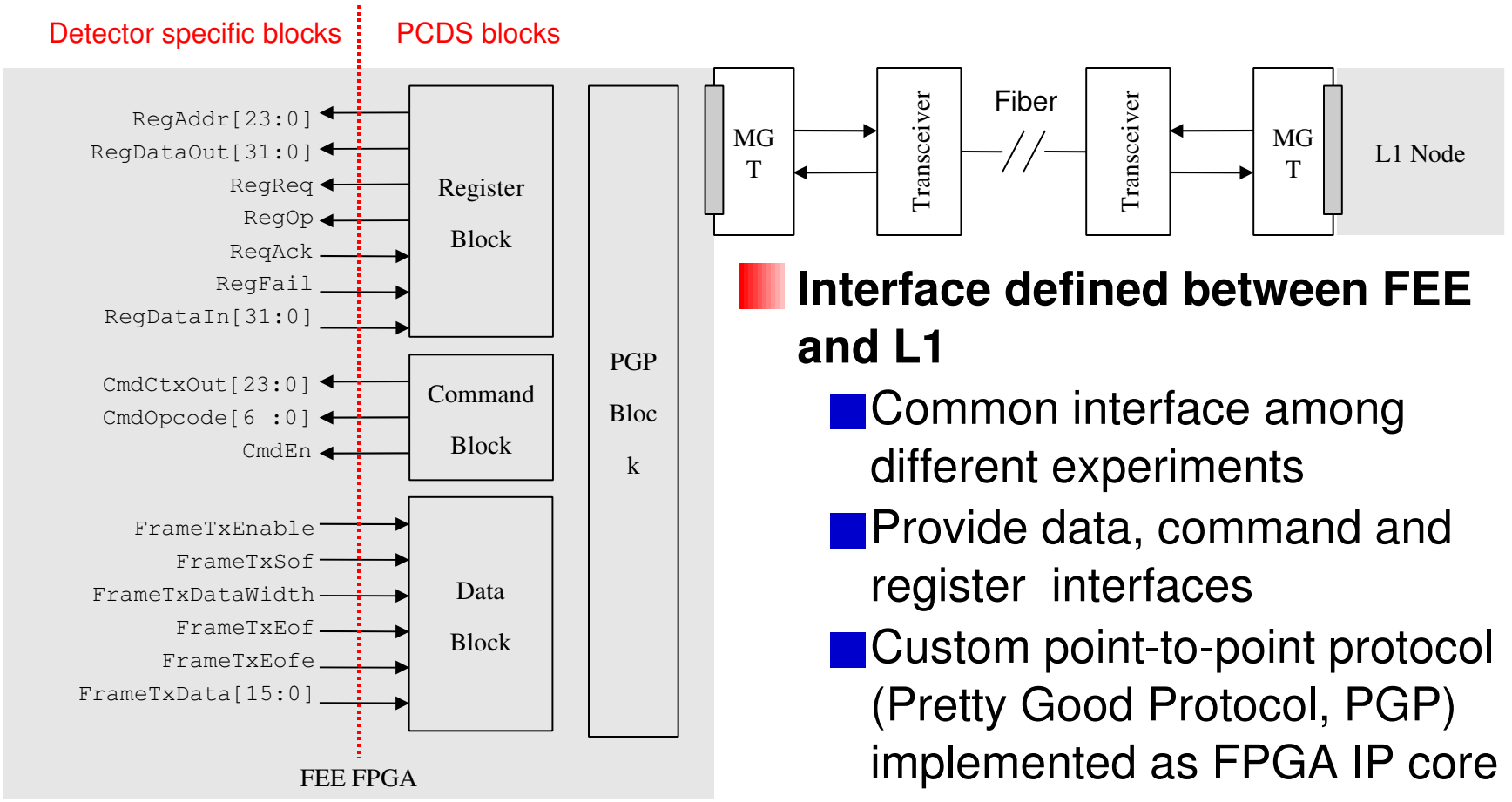
Level 0 Nodes

- **Level 0: Control**
  - DAQ operator consoles
- **Provide different functionalities:**
  - Run control
    - Partition management, data-flow
  - Detector control
    - Configuration (modes, biases, thresholds, etc)
  - Run monitoring
    - Data quality
  - Telemetry monitoring
    - Temperatures, currents, voltages, etc
- **Manage all L1, L2 and L3 nodes in a given partition (i.e. the set of DAQ nodes used by a specific experiment or test-stand)**

Level 1 Nodes

- **Level 1: Acquisition**
  - Receive 120 Hz timing signals, send trigger to FEE, acquire FEE data
  - Error detection and recovery of the FEE data
  - Control FEE parameters
  - Calibration
    - Dark image accumulation and averaging
    - Transfer curve mapping, gain calculation
    - Neighbor pixel cross-talk calculation
  - Event-build FEE science data with beam-line data
  - Image processing
    - Pedestal subtraction using calibration constants, cross-talk corrections
    - Partial data reduction (compression)
    - Rejection using 120 Hz beam-line data
    - Processing envisioned both in software and firmware (VHDL)
  - Send collected data to Level 2 nodes over 10 Gb/s Ethernet

Register Command Data Interface



- **Interface defined between FEE and L1**
  - Common interface among different experiments
  - Provide data, command and register interfaces
  - Custom point-to-point protocol (Pretty Good Protocol, PGP) implemented as FPGA IP core
  - FEE FPGA assumed to be Xilinx Virtex-4 FX family with Multi Gigabit Transceivers (MGT)

Amedeo Perazzo
perazzo@slac.stanford.edu

Level 2 & 3 Nodes

- **Level 2: Processing**
  - High level data processing:
    - Learn, pattern recognition, sort, classify
      - *e.g. combine 10^5 – 10^7 images into 3D data-set*
    - Alignment, reconstruction
  - Currently evaluating different ATCA blades for L2 nodes
  - Send processed data to L3 over 10 Gb/s Ethernet
- **Level 3: Data Cache**
  - Provide data storage
    - Located in server room in experimental hall
  - Off-line system will transfer data from local cache to tape staging system
    - Tape staging system located in SLAC central computing facilities
  - Must be able to buffer up data in local storage during downtimes of staging system
    - Current requirement is ~4 days of data

Reconfigurable Cluster Element (I)

- **The RCE is the most interesting among the different Level 1 node types**
  - SLAC custom made ATCA board
- **Based on System On Chip (SOC) Technology**
  - Currently implemented with Xilinx Virtex 4 devices, FX family
    - Targeting XC4VFX60
  - Xilinx devices provide
    - Reconfigurable FPGA fabric
    - DSPs (200 for XC4VFX60)
    - Generic CPU (2 PowerPCs 405 running at 450 MHz for XC4VFX60)
    - TEMAC: Xilinx TriMode Ethernet Hard Cores
    - MGT: Xilinx Multi-Gigabit Transceivers 622Mb/s to 6.5Gb/s (16 for XC4VFX60)
- **Power consumption**
  - About 72 watts for fully populated board
    - Configuration with 2 RCE per board and 4 flash memory slices per RCE
    - ~30W per board+RTM, ~15W per FX60 (RCE), ~1.5W per FX20 (slice)

Reconfigurable Cluster Element (II)

- **FPGA fabric**
  - Interfaces to:
    - memory subsystems
    - JTAG debug port
    - custom multi-function display
    - various I/O channels
  - Generic DMA Interface (PIC) designed as set of VHDL IP cores
    - Up to 16 PIC channels
  - PIC in conjunction with Multi-Gigabit Transceivers and protocol cores, provide many channels of generic, high speed, serial I/O
    - 10Gb Ethernet
    - PGP
  - PIC in conjunction with TriMode Ethernet Hard Cores also provide commodity network interfaces
    - 1Gb Ethernet

Reconfigurable Cluster Element (III)

- **System Memory Subsystem**
  - 512 MB of RAM (currently 128 MB)
    - Memory controller provides 8 GB/s overall throughput
    - Uses Micron RLDRAM II
- **Platform Flash Memory Subsystem**
  - Stores firmware code for FPGA fabric
- **Configuration Flash Memory Subsystem**
  - 128 MB configuration flash
  - Dedicated file system for storing software code and configuration parameters (up to 16 selectable images)
- **Storage Flash Memory Subsystem (optional)**
  - Up to 1TB per RCE persistent storage flash (currently 256GB per RCE)
    - Low latency/high bandwidth access through I/O channels using PGP
    - Uses Samsung K9NBG08 (32 Gb per chip)

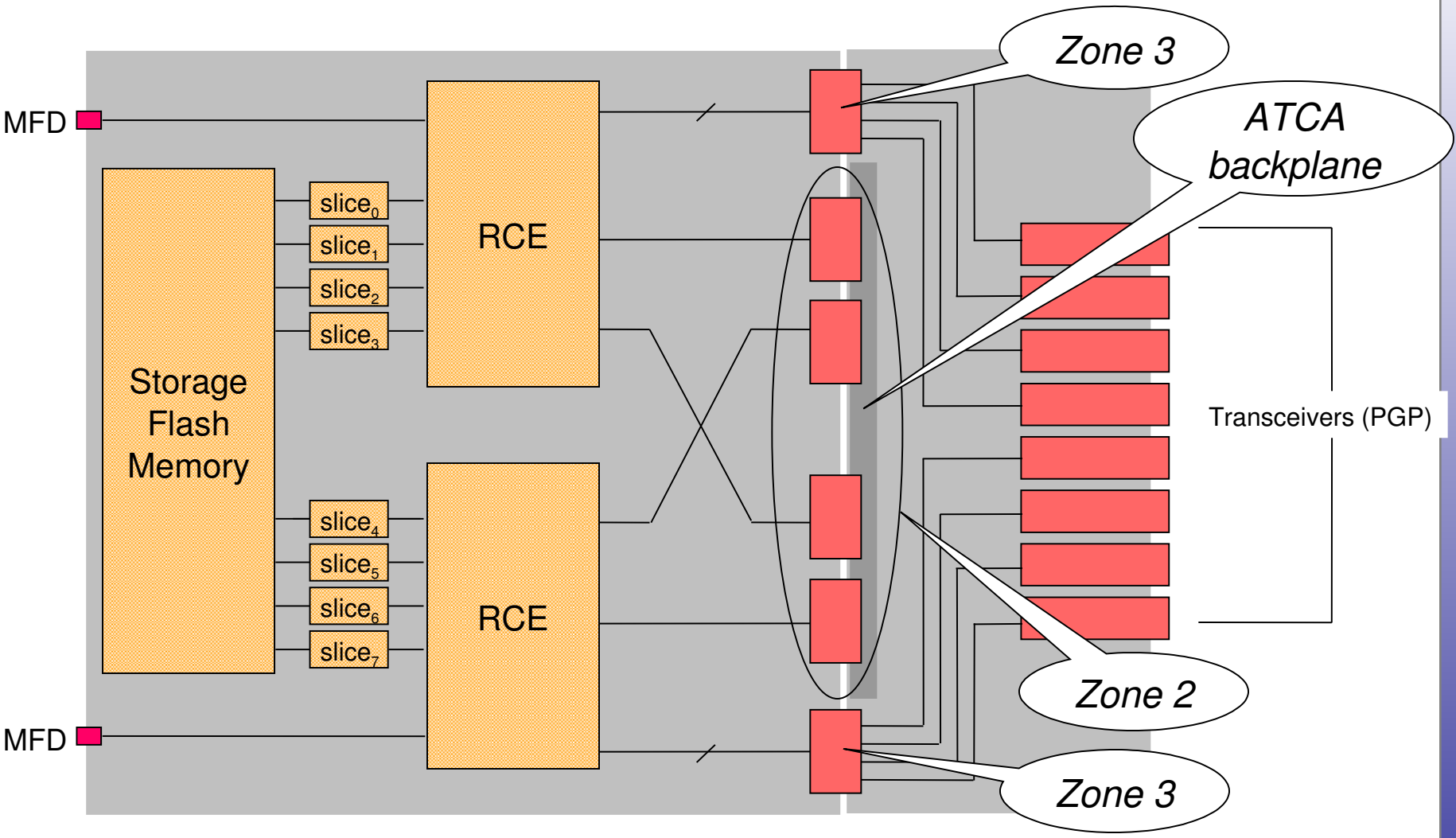
Pretty Good Protocol

- **Pretty Good Protocol (PGP)**
  - Serial point-to-point connectivity
    - Physical interface is 2 LVDS pairs/lane
  - Small footprint
  - Features: clock recovery, full duplex, reliable frame transmission and reception, deterministic (and small) latency
  - Implemented as IP protocol core interfaced to PIC
    - Extensible in both bit-rate and number of lanes
  - Used to interface the RCE to its storage flash memory subsystem
    - Four lanes at 3.125Gb/s (~1GB/s per RCE)
  - Used to interface the RCE to the front-end electronics
    - Up to four channels at 3.125Gb/s (~1GB/s per RCE or ~2GB/s per RCE board)

RCE Software

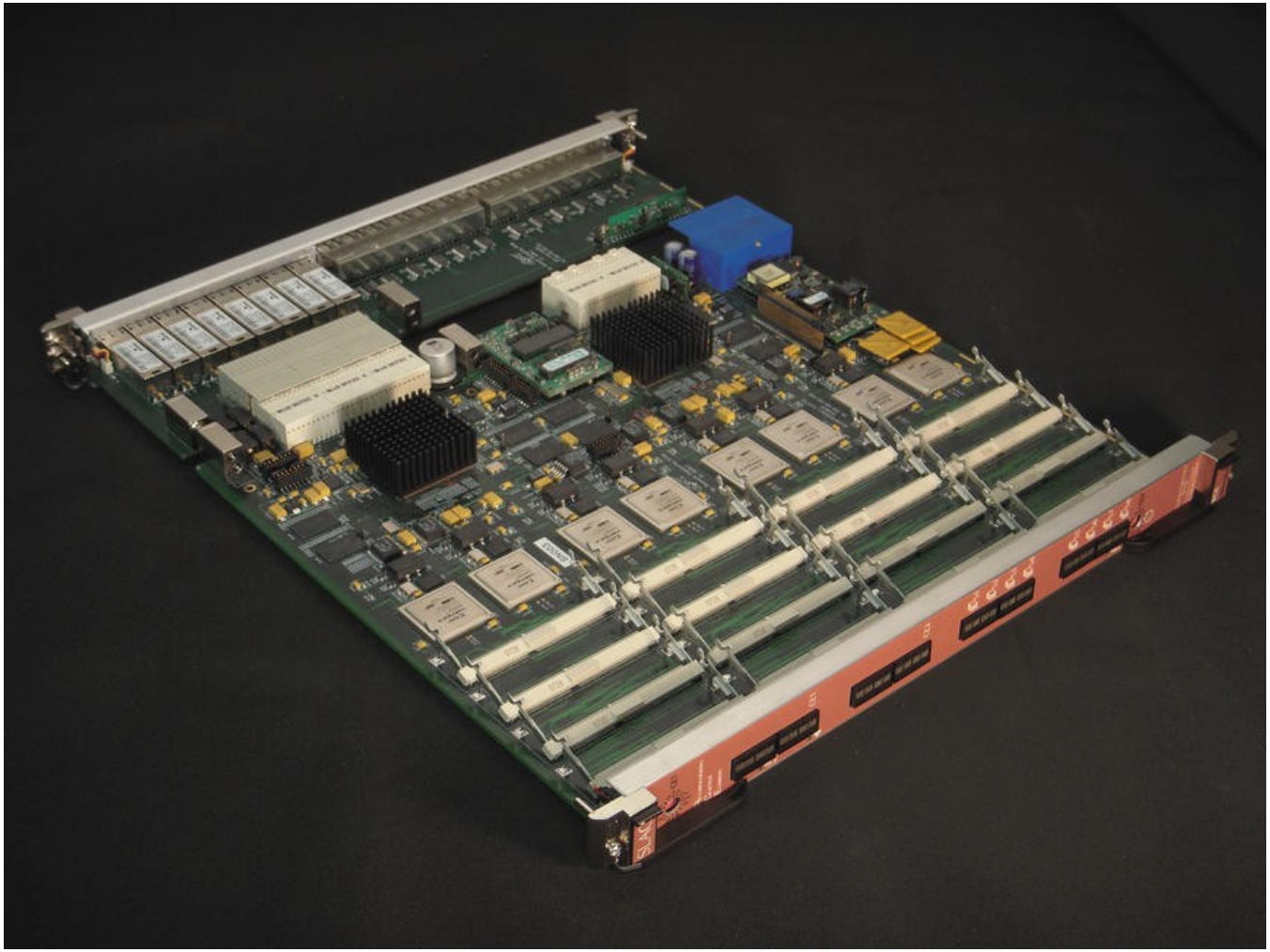
- **Software**
  - Ported open source Real-Time kernel
    - Adopted RTEMS: Real Time Operating Systems for Multiprocessor Systems
  - Written BSP mainly in C++
    - Plus some C and assembly
  - Written 10Gb Ethernet driver and PGP drivers for bulk data
  - 1Gb management interface driver
  - Built interface to RTEMS TCP/IP network stack (BSD)
  - Developed specialized network stack for zero-copy Ethernet traffic

RCE Board Block Diagram



Amedeo Perazzo
perazzo@slac.stanford.edu

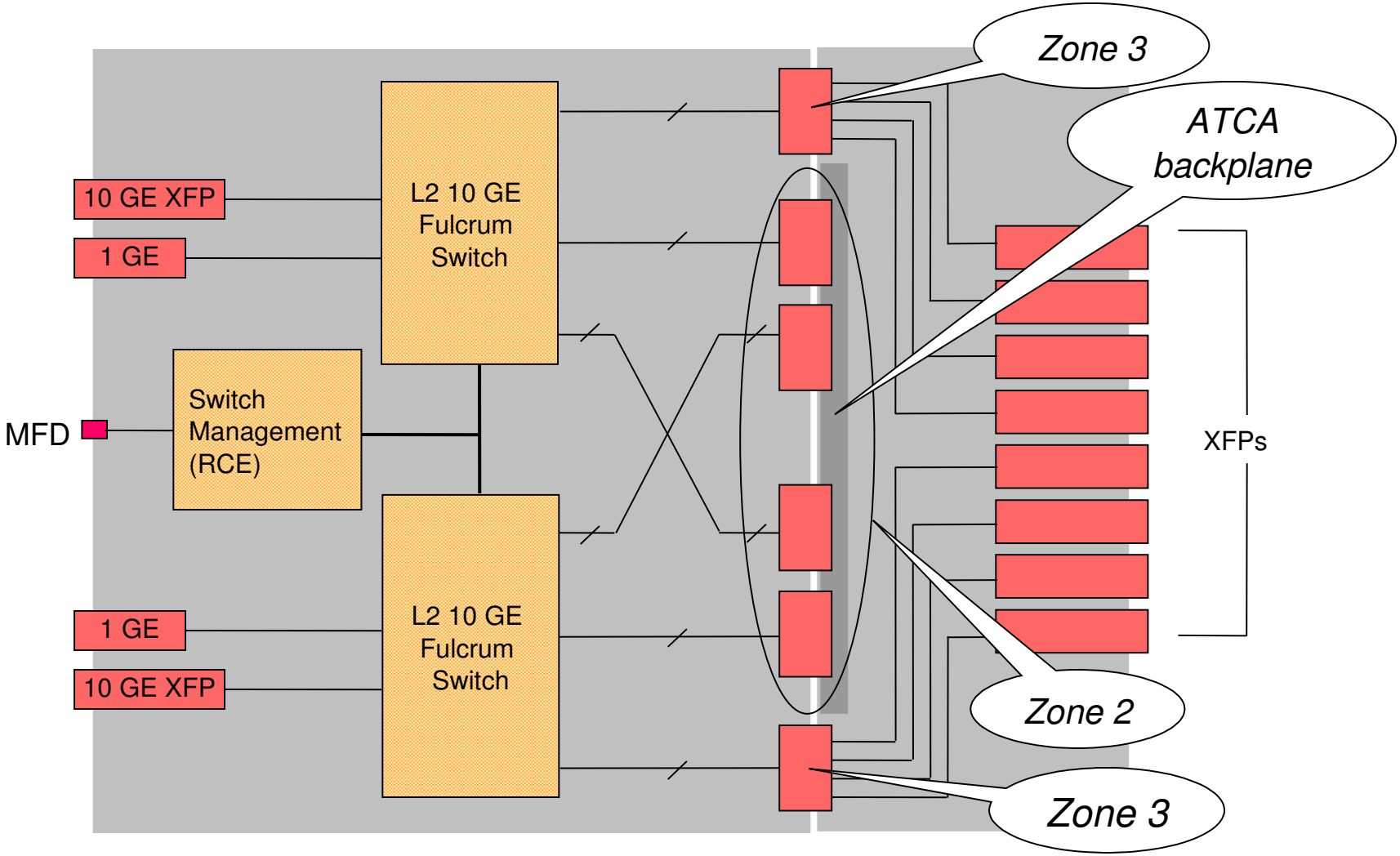
RCE Board with RTM



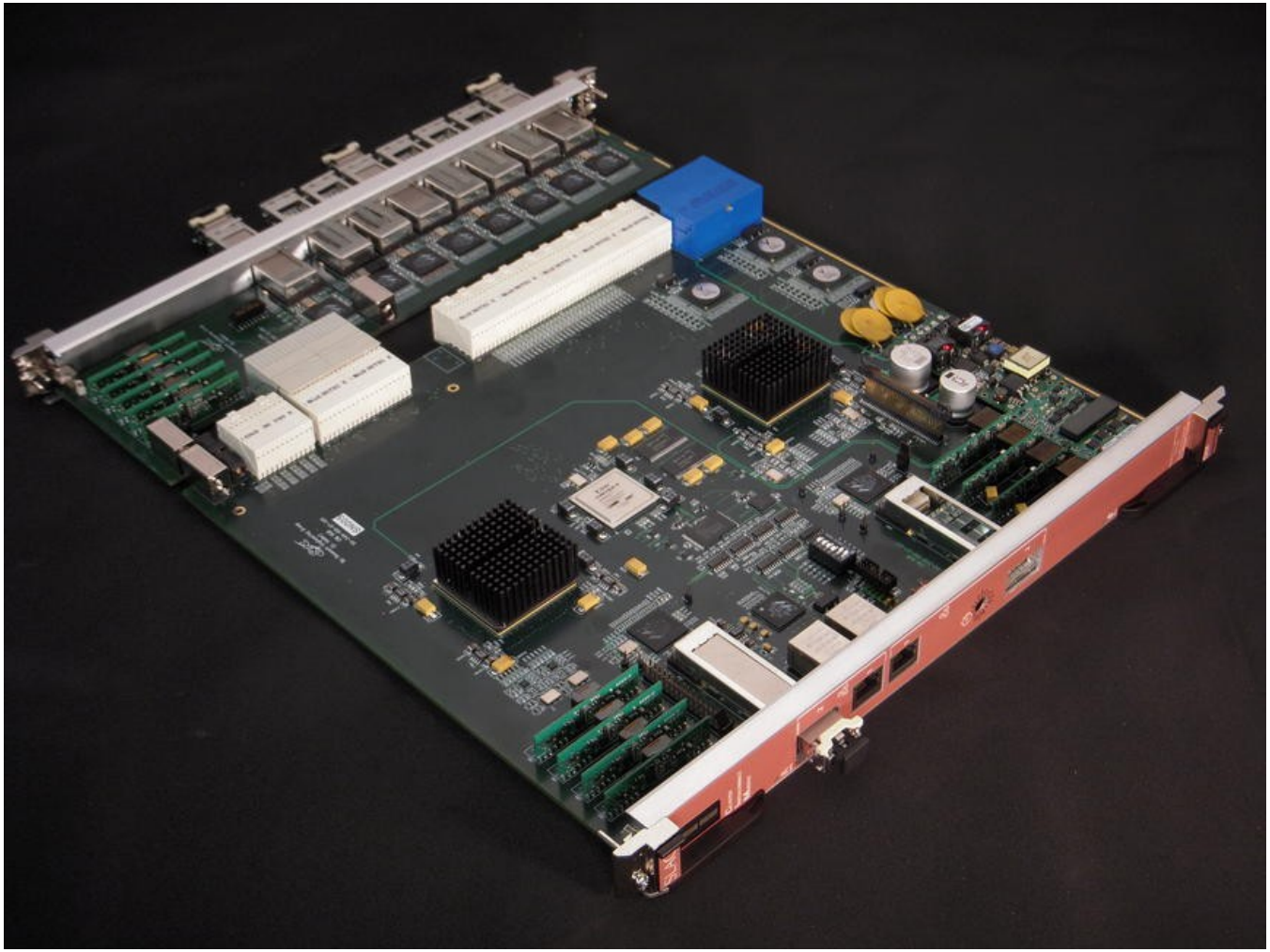
Cluster Interconnect Module

- **ATCA network card**
  - SLAC custom made board
  - Based on two 24-port 10Gb Ethernet switch ASICs from Fulcrum
    - Up to 480 Gb/s total bandwidth
  - Managed via Virtex-4 device
    - Currently XC4VFX12
  - Fully managed layer-2, cut-through switch
  - Interconnect up to 14 in-crate RCE boards (i.e. 28 RCEs)
  - Interconnect multiple crates for additional scalability
  - Power consumption: ~50W per CIM (~100W per board)
- **Fully configurable**
  - Designed to optimize crates populated with RCE boards
    - Ability to use ATCA redundant lanes for additional bandwidth if desired
    - Ability to use 2.5Gb/s connections in place of standard 1Gb/s Ethernet
  - At the same time may be configured to connect standard ATCA blades

CIM Board Block Diagram



CIM Board with RTM



Front-end Development Board

ATCA Chassis

Front

Rear

High Speed Digitizer (AMO)

- **Acqiris DC282 high-speed 10-bit cPCI Digitizer**
  - 4 channels
  - 2-8 GS/s sampling rate
  - Acquisition memory from 1024 kpoints to 1024 Mpoints (optional)
  - Low dead time (350 ns) sequential recording with time stamps
  - 6U PXI/CompactPCI standard, 64 bit, 66 MHz PCI bus
    - Sustained transfer rate up to 400MB/s to host SBC

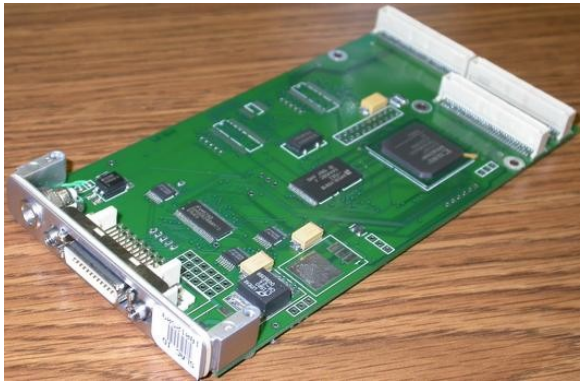
Mega-pixel Camera (AMO)

- **Adimec camera**
  - 1024 x 1024 pixel resolution
    - KODAK KAI-01050 Image Sensor
  - 5.5 um pixel size
  - 120 full frames per second
  - Industry standard Camera Link interface
  - 63 dB dynamic range
  - 5.6 Lux sensitivity
  - Low readout noise (14 e-)

Readout SBC (AMO)

- **L1 Node for AMO**
- **Concurrent Technologies PP512**
  - Dual 2.5 GHz Core Duo processors
  - Up to 6 GB DDR2 SDRAM
  - Compact flash site
  - Two PMC/XMC slots
  - Three 1Gb/s Ethernet interfaces
  - 6U PXI/CompactPCI standard, 64 bit, 66 MHz PCI bus

PMC Sites (AMO)



- **Micro-Research Finland PMC-EVR-200**
  - Receives accelerator timing information (clock, fiducial, op-codes) via fiber from Machine Event Generator (EVG)
  - Transmit electrical triggers to cameras/ digitizers
- **Camera Link Framegrabber**
  - Industrial standard for cameras control/readout
  - Up to 250 MB/s image data transfer rate

Conclusions

- **LCLS common custom DAQ hardware devices fully prototyped**
  - RCE, CIM and front-end development board
- **Interface with FEE defined for Cornell and BNL detectors**
  - Uses common (among different detectors) communication protocol
    - Implemented as IP cores interfaced to detector specific user logic in FEE FPGA
  - Continue developing detector specific user logic for ASIC operations
- **Interface with off-line system defined**
- **Continue core software development**
  - DAQ partition management, user interface, monitoring, configuration, calibration framework and data-flow
  - Core software will be ready and commissioned for AMO before LUSI starts