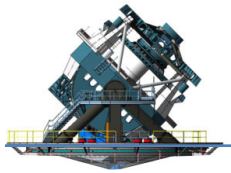
# The Large Synoptic Survey Telescope Camera Data Acquisition System

Gregg Thayer
SLAC National Accelerator Laboratory

21st IEEE Real Time Conference
June 12, 2018

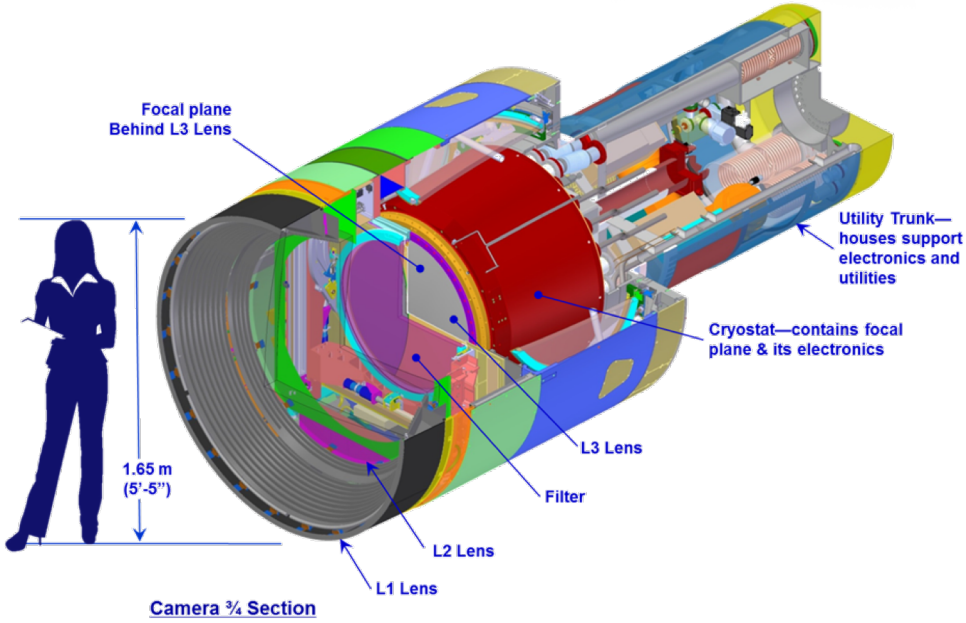
# What is LSST?

- LSST is a telescope that is being built on Cerro Pachón in Chile
  - Will conduct a 10-year survey of the optical sky
    - Scheduled to begin in 2022
  - Has an 8.4m primary mirror and wide field of view
    - One image will cover 49 times the area of the full moon
  - Able to survey the southern hemisphere sky every few nights in 6 wavelength bands
  - Data will be archived at NCSA and made publicly available along with tools to access and study it
    - The final image collection will be 0.5 exabytes
    - The peak compute power in LSST data centers will be 1.8 PFLOPS
- Diverse science goals drive LSST
  - The Nature of Dark Matter and Dark Energy
  - Cataloging the Solar System (moving objects)
  - Exploring the Changing Optical Sky
  - Milky Way Structure and Formation

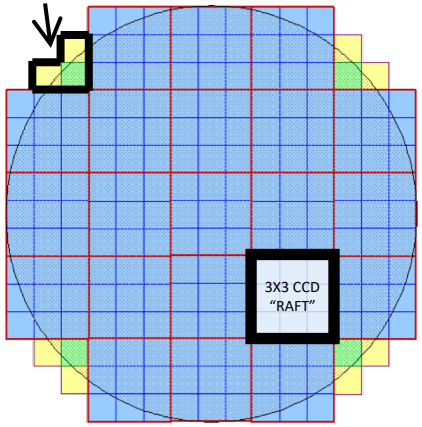
# The LSST Camera

- The largest digital camera constructed for astronomy
  - 3.2 Gigapixels digitized with 18-bit accuracy
  - Each 16-megapixel sensor is read out in parallel by 16 amplifiers
    - 189 science sensors
    - 8 guide sensors
    - 4 wavefront sensors
- 3 science sensors are controlled by a Readout Electronic Board (REB)
  - Three REBs and their associated sensors form a mechanical unit called a Raft
  - There are 21 Science Rafts
- Guiding and wavefront sensors are contained in the four Corner Rafts
- Nominal cadence is 17 seconds
  - 15 second exposure
  - 2 second readout
- Project completion scheduled for mid-2020

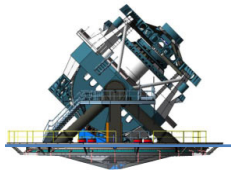



Camera ¾ Section

*4 Corner areas for wavefront sensing (green) and guiding (yellow).*



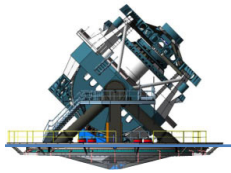

*21 rafts make up the science array*

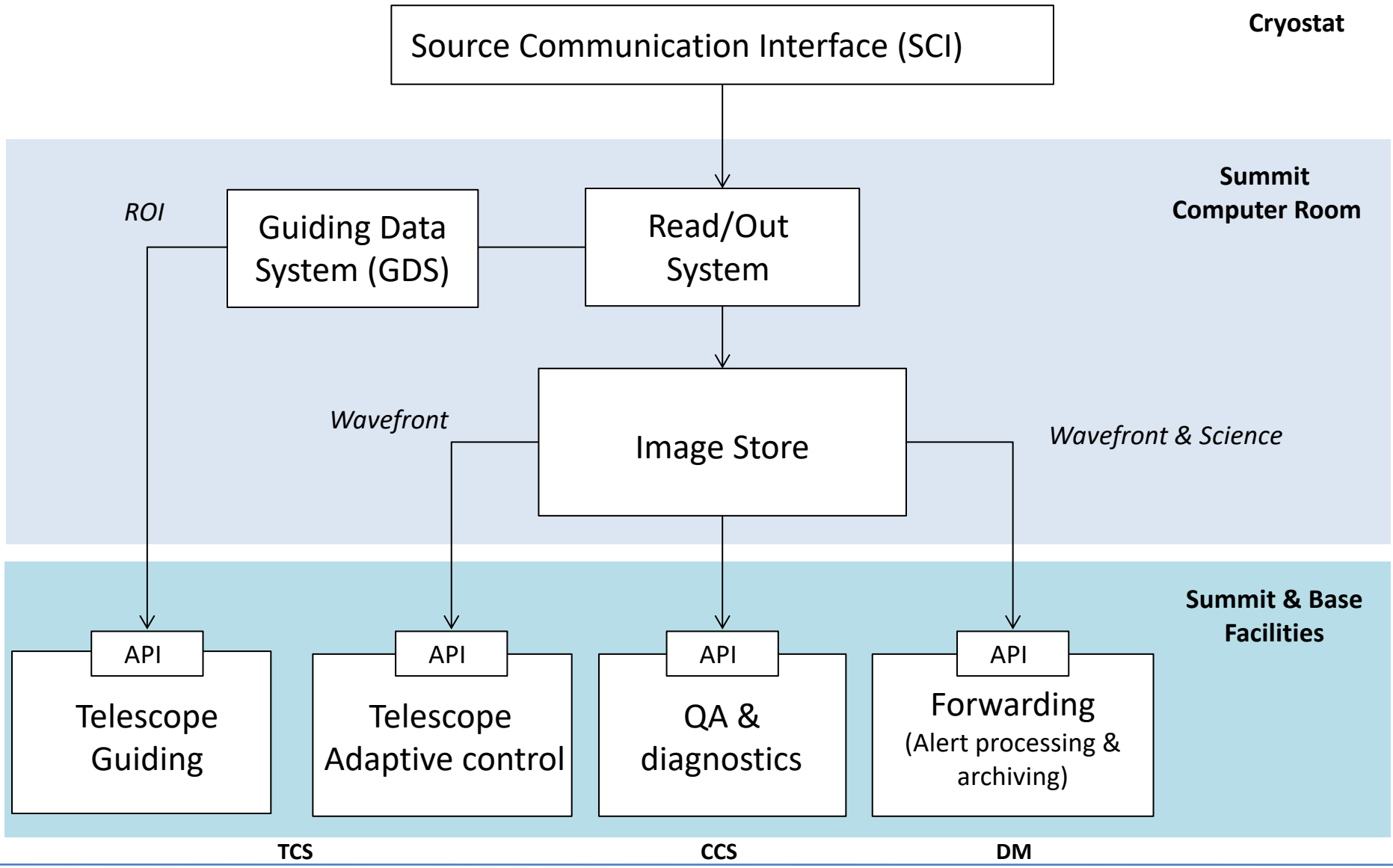


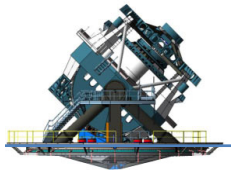
# The Camera DAQ

- Provides access to image data to multiple, concurrent clients at read-out rate, in “real-time”
  - Read-out rate is ~3.6 Gigabytes/second
- Buffers all data for up to 7 days
  - This allows continued observation from the summit in the event of communication loss
    - 2 days covers the Mean Time To Repair (MTTR) for summit/base
    - 7 days covers the MTTR for base/NCSA
- Provides cross-talk corrected data if requested
- Provides access to raft electronics for the purpose of configuration, control, and monitoring
- Provides common system clock & synchronous timing
- Provide windowed guiding data at 9 Hz

# The DAQ Components and Clients



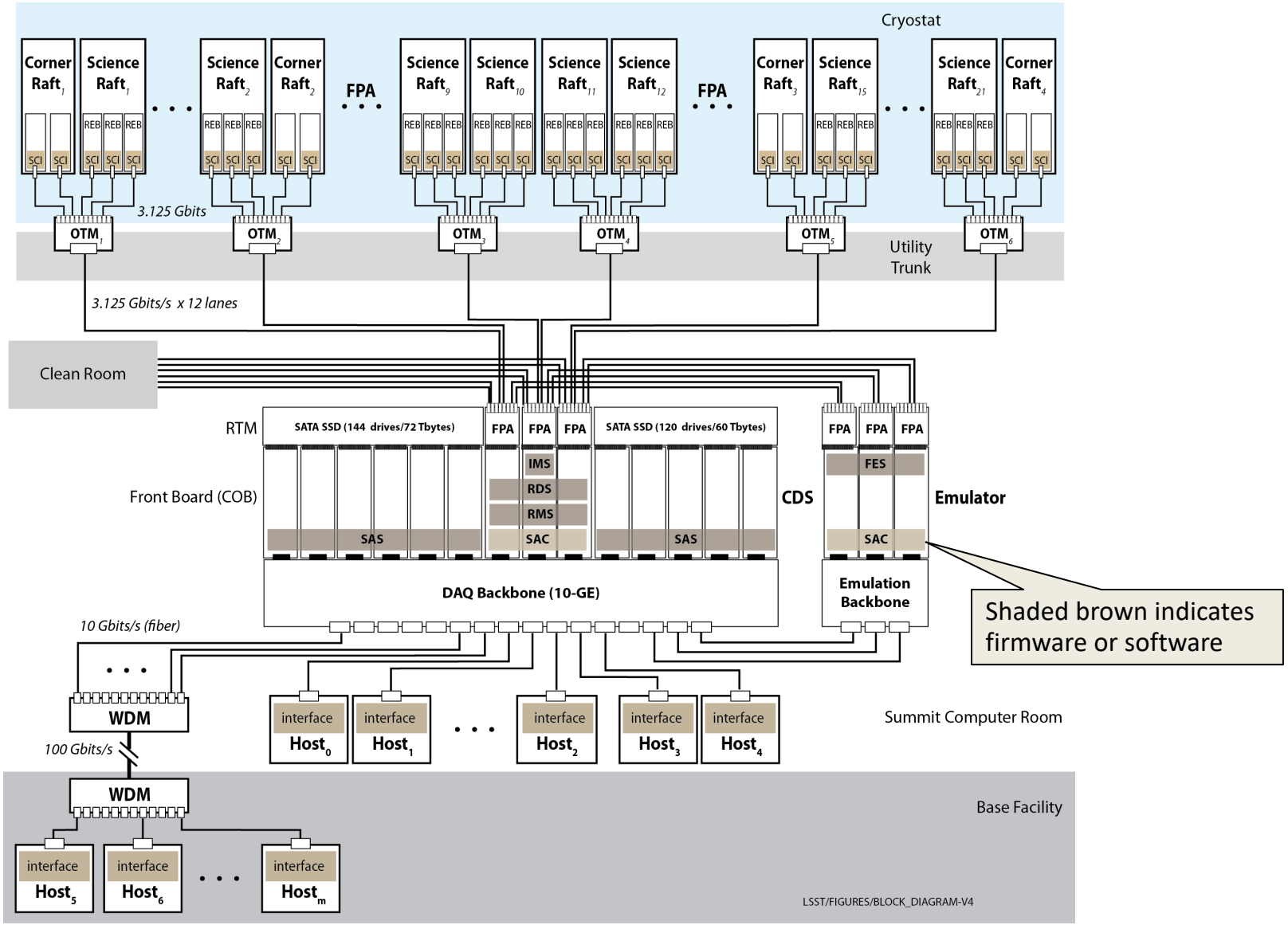




# DAQ System Novelties

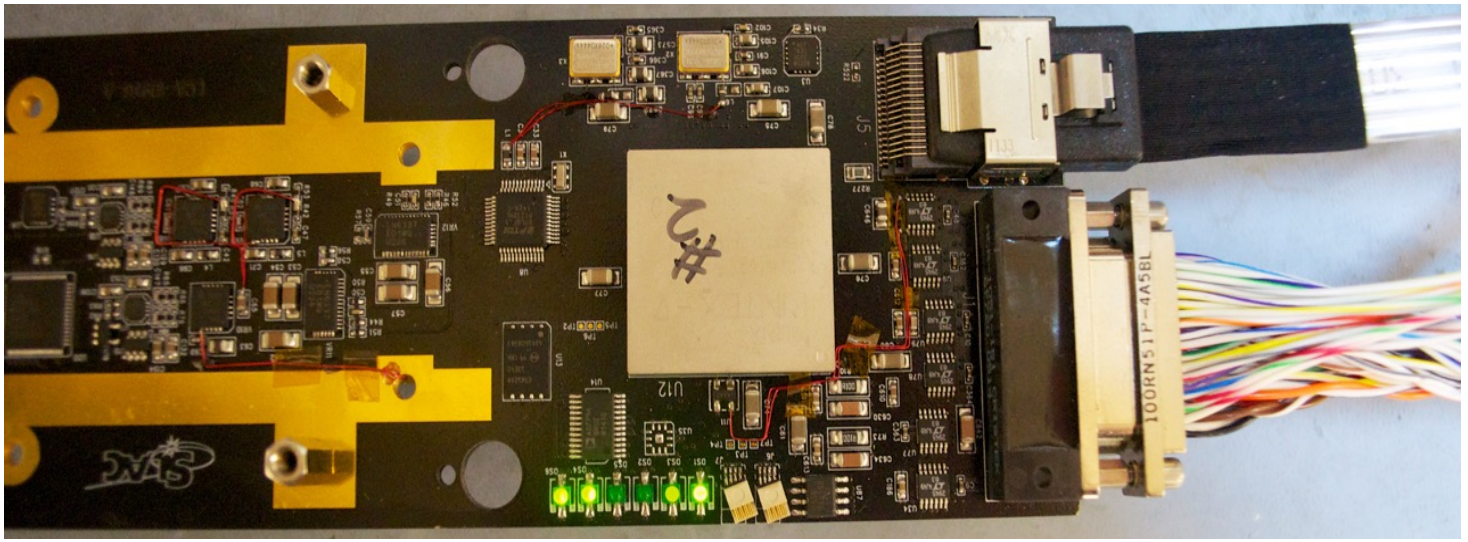
- "Event" size is large at 7.2 GB and rate is low at 0.06 Hz
  - And that data is almost exclusively collected at night…
- However, we must provide the data at the rate it is generated in the front end
- DAQ latency requirements are driven by the Observatory real-time alert latency of 60 seconds
  - ~10 million such alerts are expected to be generated each night
- DAQ client machines are distributed over more than 100 km
  - All clients, and indeed DAQ compute elements, are equal peers on a single 10G Ethernet
- There are 142 high speed serial links which must penetrate the cryostat in a very limited area (6 x ~8 $cm^2$) and then travel 145 meters to the summit computer room

# Block diagram of the DAQ System



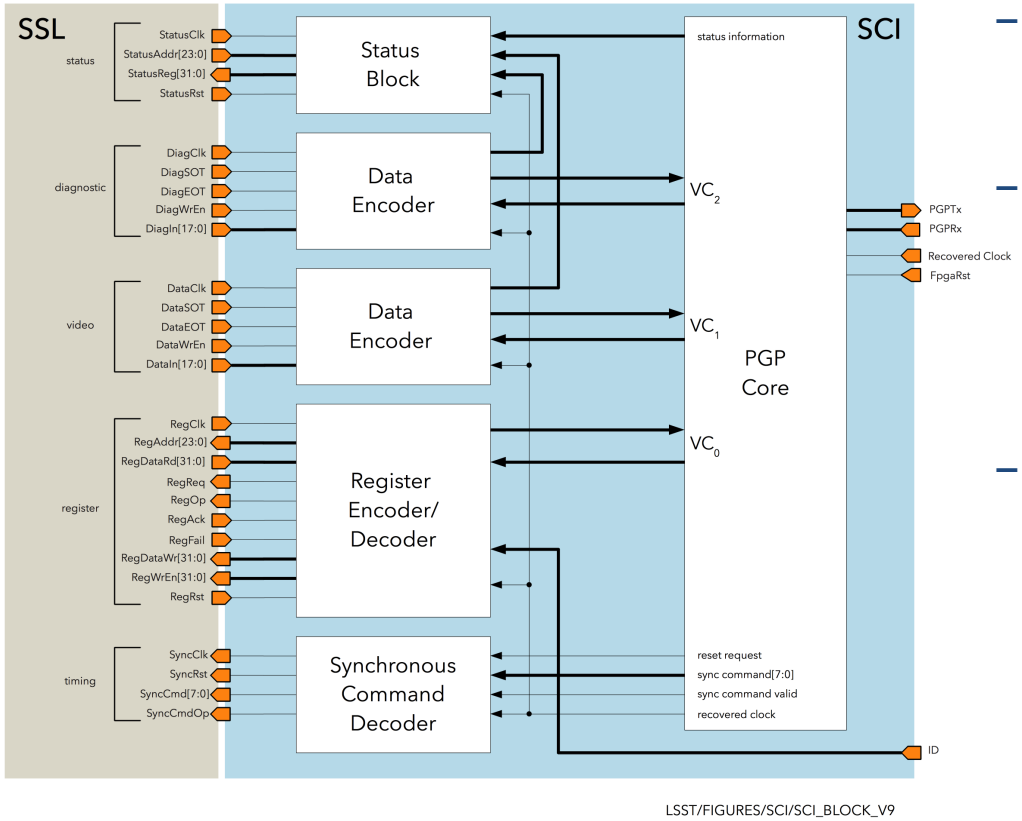

Shaded brown indicates firmware or software

# The Raft Electronics Board



- Each REB, regardless of type (Science, Guider, or Wavefront) contains an identical FPGA
- FPGA firmware is partitioned into two components separated by a controlled interface
  - The Source Communication Interface (SCI) provided by the DAQ
  - The sensor specific logic provided by the electronics group
- The DAQ specifies the Data connector and its pinout

# The Source Communication Interface

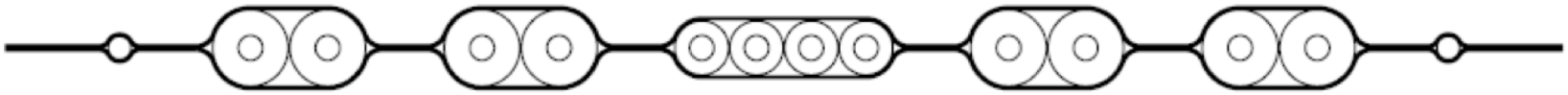



LSST/FIGURES/SCI/SCI_BLOCK_V9

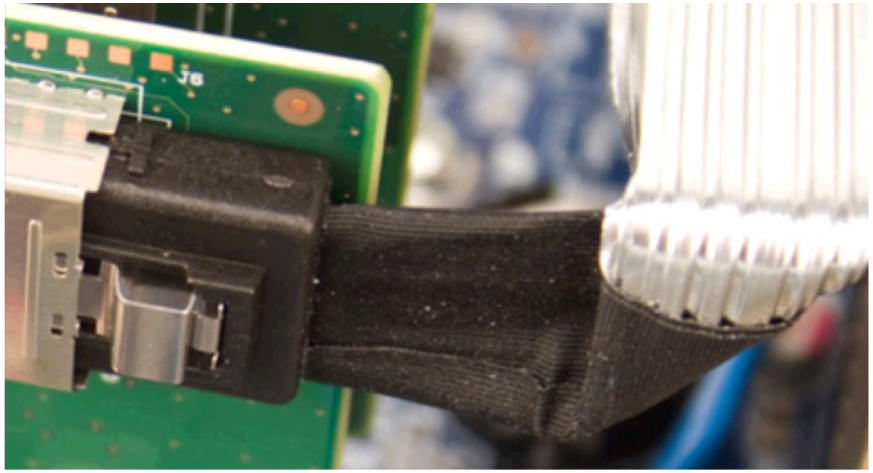
See Stefano Russo's poster 446 for some of what goes in the SSL box!

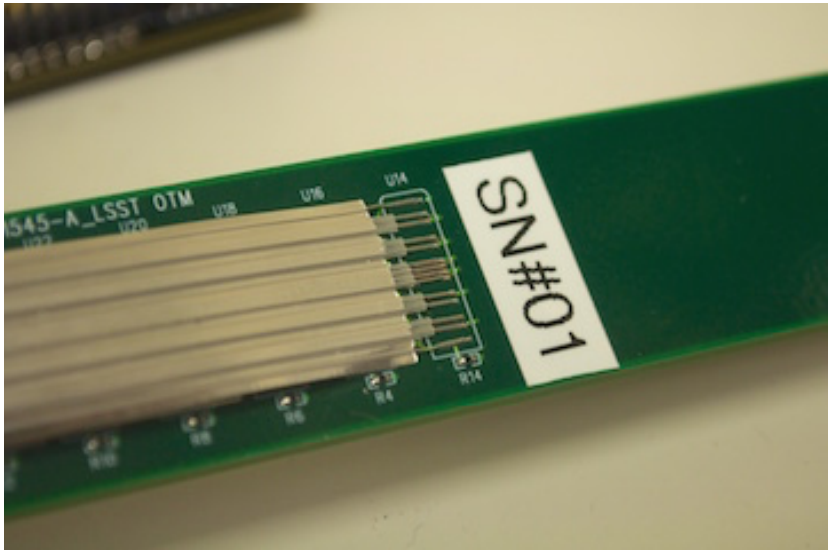
- A small, lightweight, self-contained set of firmware sharing the FPGA with the Sensor Specific Logic
- Employs high speed (3.125 Gbits/sec) point-to-point, serial communication protocol (PGP)
  - Protocol assures QOS through four Virtual Channels
- The SCI has four principal functions
  - Provide a stable, system reference clock and synchronous commands
    - This clock is the recovered clock from the serial data link
  - Reliably encode and transmit 18-bit pixel data to the CDS
  - Provide a mechanism for external configuration, control and monitoring of the REB's sensor electronics

# Cables between SCI & OTM



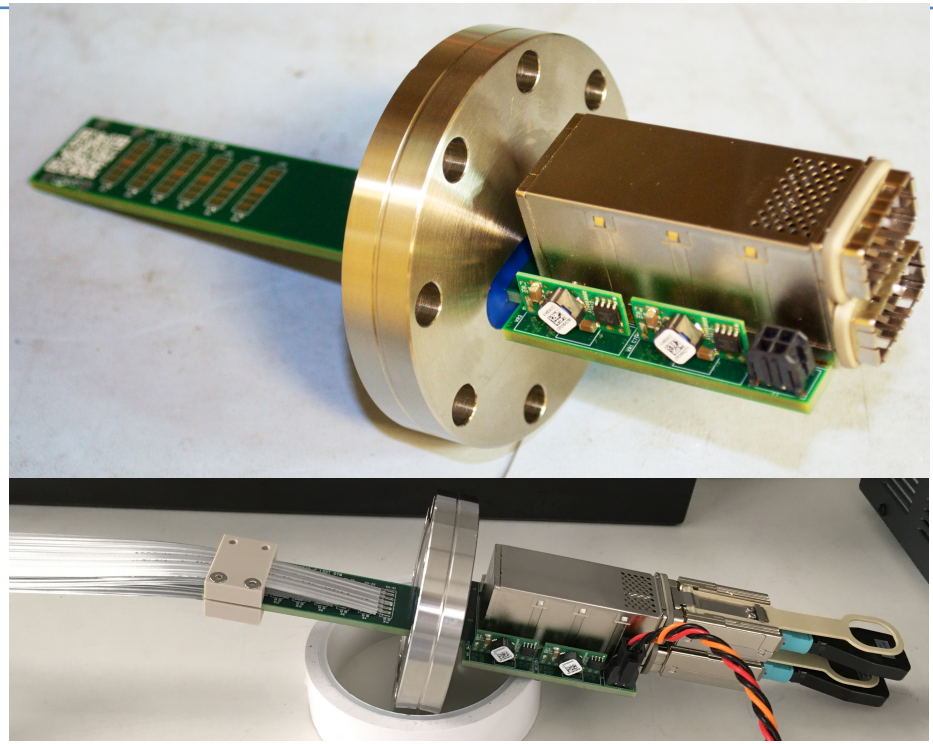
- 3M Twin Axial cable
  - Longitudinal aluminum shield
  - 2 ground wires connected to shield
  - 4 pairs for high speed signals
    - Impedance : 100 Ohm
    - Propagation delay : 4.95 ns/m
    - Intra-pair skew : <10 ps @ 3 m
    - Attenuation : -4.9 dB/m @ 10 GHz
  - 4 single ended “sideband” signals
  - 1 mm minimum bend radius
- High speed pairs (PGP primary & redundant)
- 4 single ended wires (used as 2 LVDS pairs)
- Terminations
  - REB side uses Mini-SAS connector
  - OTM side soldered directly to OTM

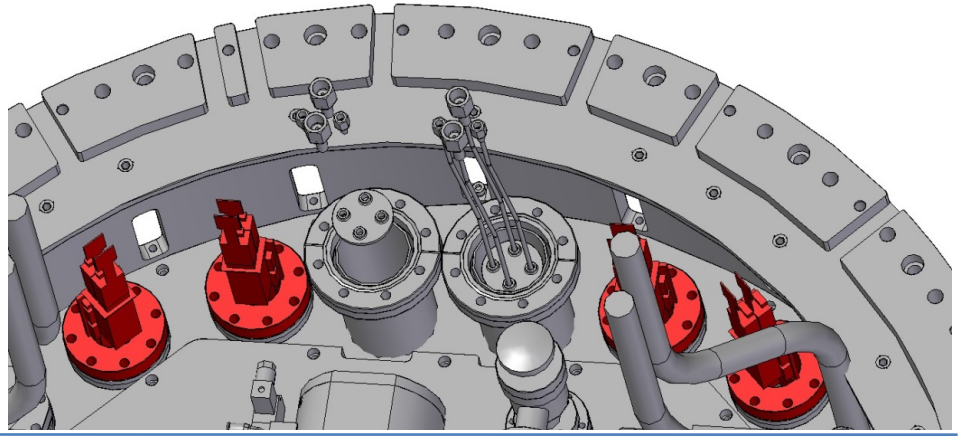




# The Optical Transition Module

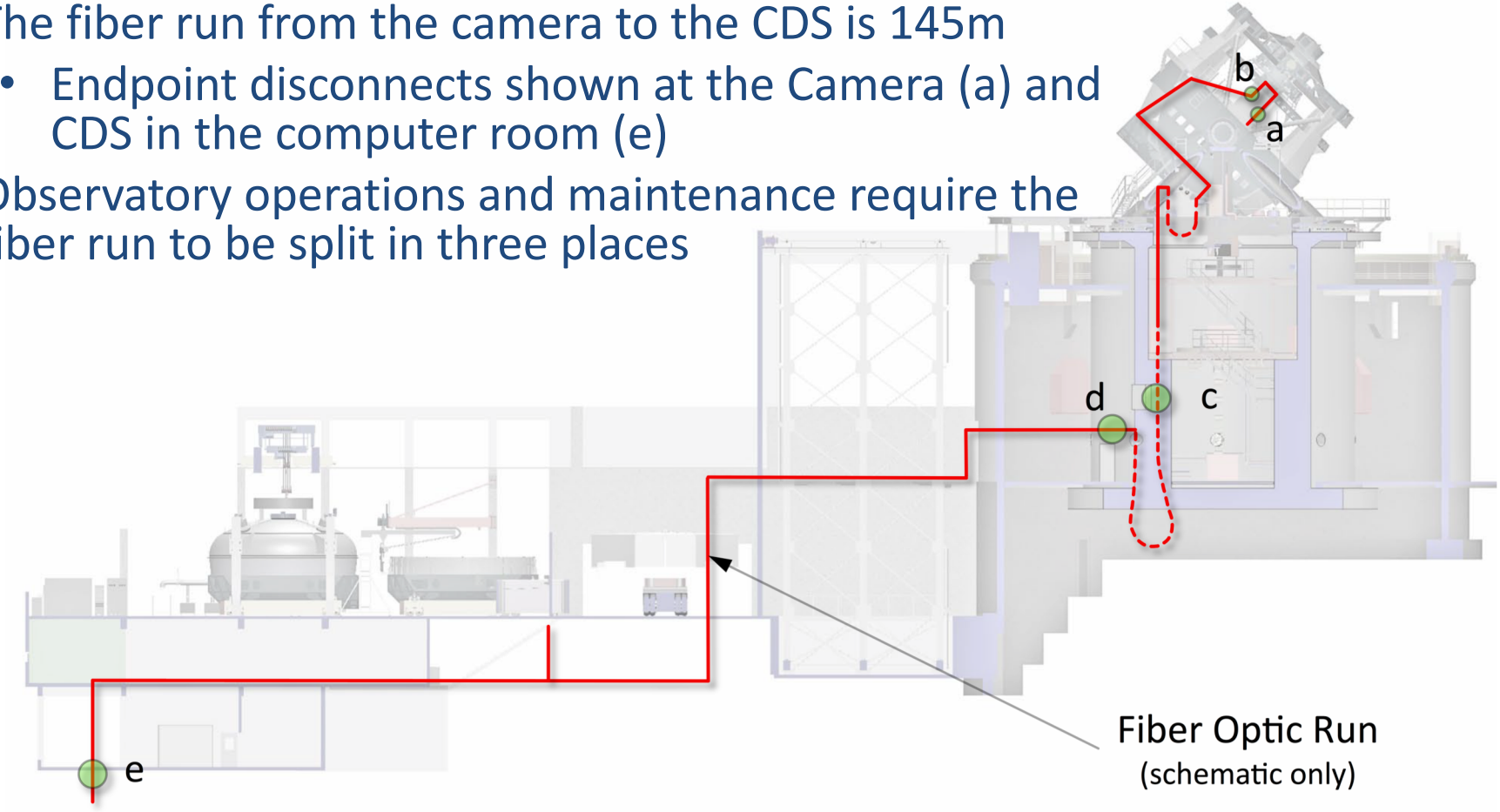
- Two purposes
  - Converts signals between copper and fiber
  - Penetrates the cryostat
- The PCB board is potted into a 3 3/8” Conflat vacuum flange
  - The vacuum side connects to 12 SCIs over copper
  - The air side has two transceivers
    - One is active, the other a cold spare
      - connected to the SCI primary and redundant links, respectively
    - Selection is done by powering one or the other
  - Each transceiver accepts a 24-fiber MTP connector
- 6 OTMs serve the focal plane’s 71 SCIs
  - 12 cables pass through the Camera’s Utility Trunk to the Camera/Telescope Interface

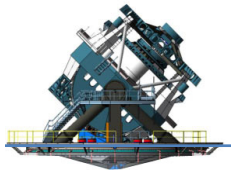




# Fiber run between camera & summit computer room

- The Telescope provides the fiber between the Camera and the summit computer room
  - Camera DAQ specifies the optical properties
- The fiber run from the camera to the CDS is 145m
  - Endpoint disconnects shown at the Camera (a) and CDS in the computer room (e)
- Observatory operations and maintenance require the fiber run to be split in three places

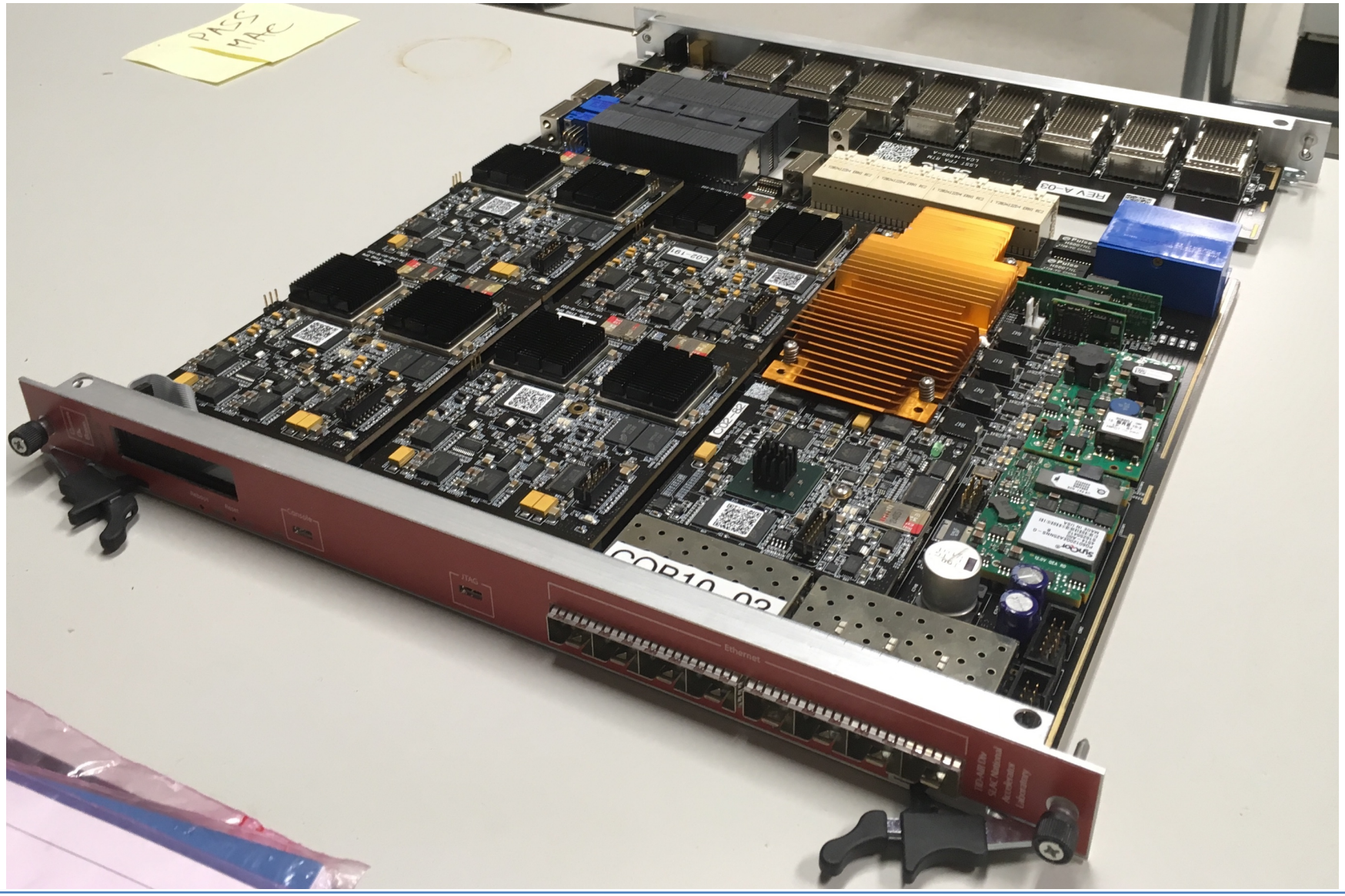




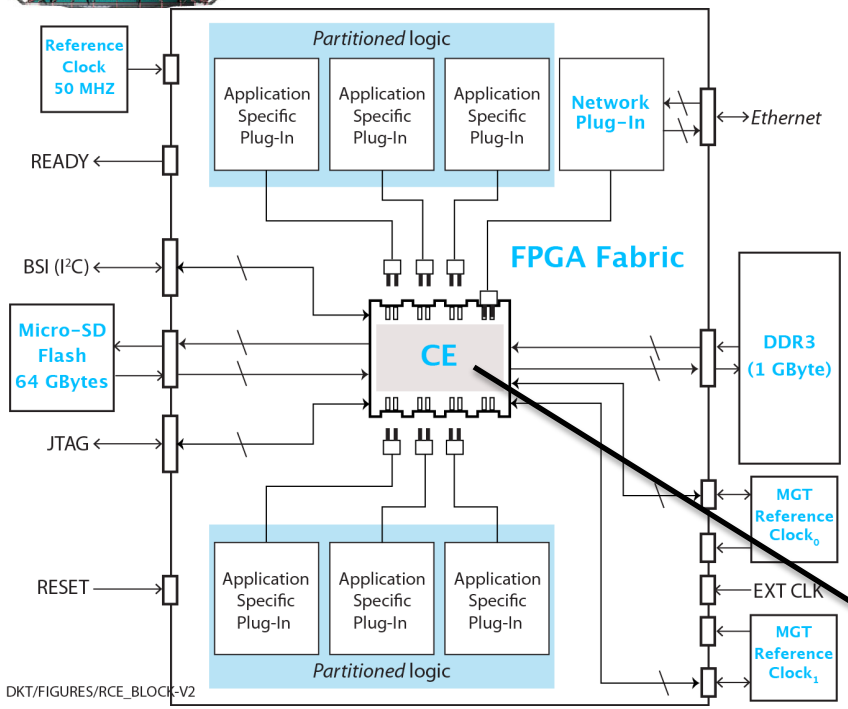
# The Camera Data System

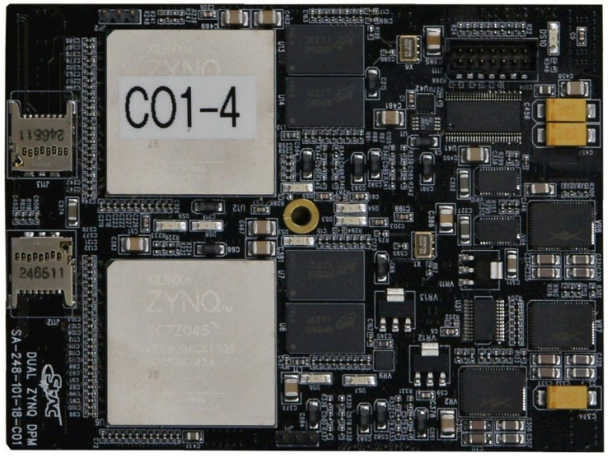
- The CDS is a 14 slot ATCA shelf with a full mesh backplane
- The Front Boards are custom carrier boards (COBs)
  - Each carries 8 Reconfigurable Cluster Elements (RCEs)
    - Ours are Xilinx Zynq-7000 SoC based
- COBs are partitioned into two sets defined by their RTM
  - 3 service the FPA (FPA RTM)
    - 8 CXP transceivers, one connected to each RCE
  - 11 service the Image Store (SSD RTM)
    - 24 SSDs, three connected to each RCE
- The system clock is distributed to all RCEs from a master over the ATCA Clock Synchronization Bus
  - This is used as the reference clock for the Zynq GTX transceivers
  - When recovered in the SCI, this clock is shared by all REBs
- Serialized commands are also fanned out to all RCEs this way
  - From there they are sent to the SCIs with fixed latency

# COB

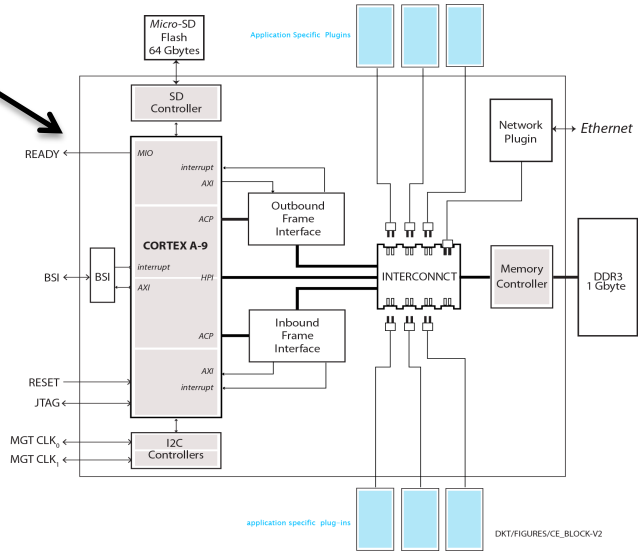


# The Reconfigurable Cluster Element



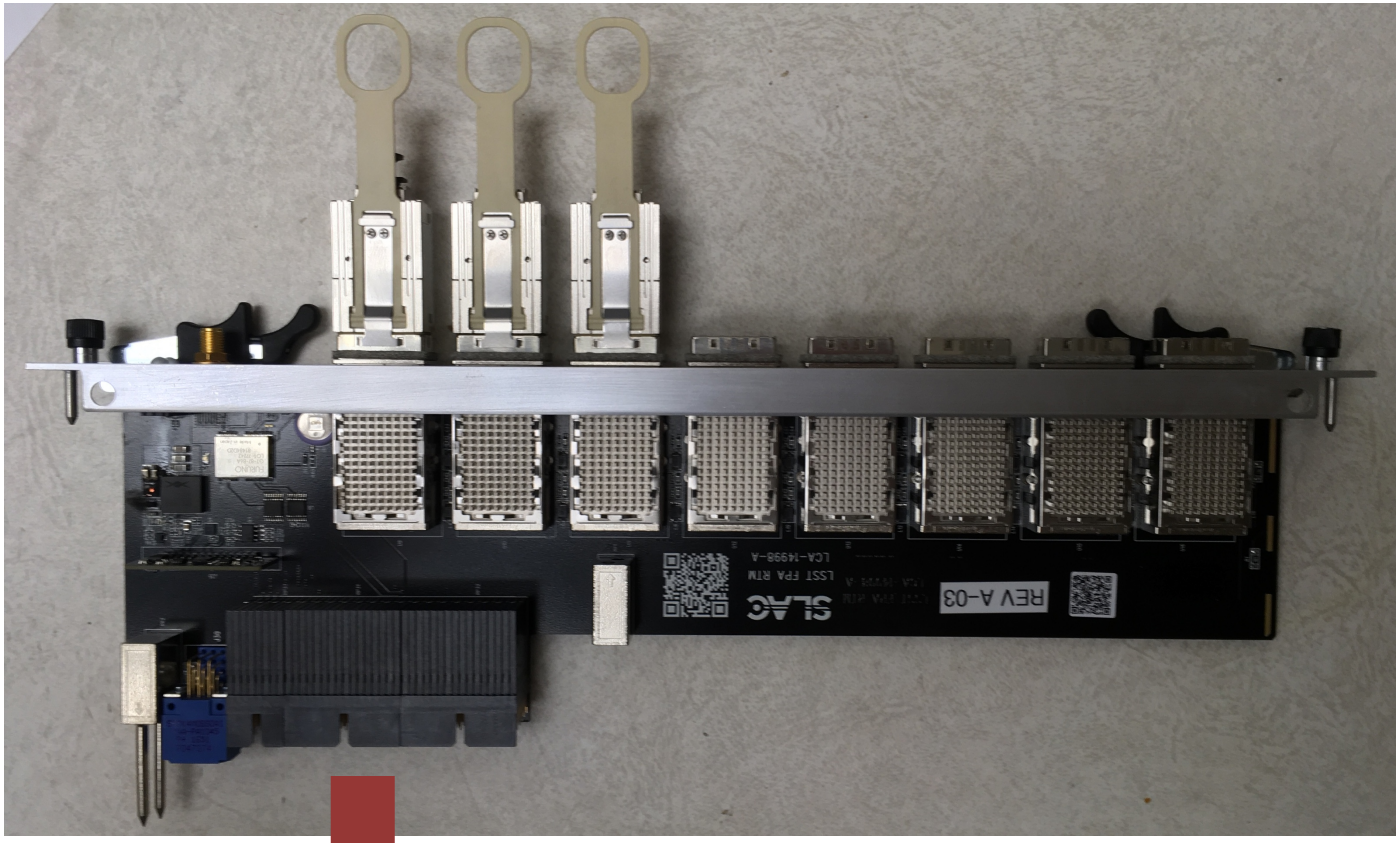




- Based on System-On-Chip technology
- The RCE is a generic computational node
- It's an FPGA, it's a DSP, it's a processor
  - Dual-core (A-9) ARM processor (@1 GHz)
  - 1 GB of DDR3 RAM
  - Up to 64 GB of flash
  - Memory subsystem has > 6 GB/s of I/O capacity
  - Real-Time kernel (RTEMS)
  - Full TCP/IP stack



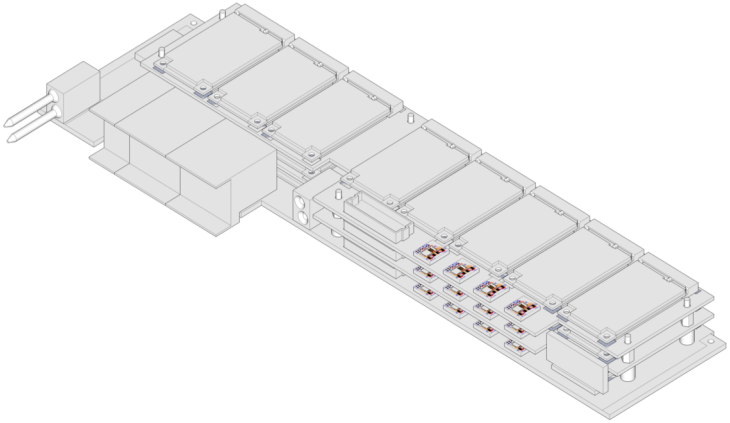
# FPA-RTM

To (up to) 9 SCI's on 1 to 3 rafts

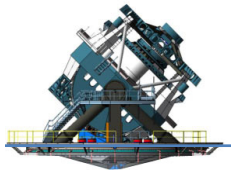



To COB (its RCEs)

# SSD-RTM




To COB (its RCEs)



# Status and Summary

- All the hardware for the DAQ has been prototyped, and full production runs will occur this year
- Scaled down versions of the CDS and its associated software have been used by other LSST groups in their development for many years
- Work remains on the final implementation of the Image Store, Guider Data System, and Crosstalk Correction

- The LSST Camera DAQ has features that distinguish it from other large experiment's DAQ systems
  - The large geographic distribution of its components
  - The need to buffer many days worth of data
  - The high speed serial data are generated in a cryostat and need to get out

# End of Presentation