Electronics Systems Issues for SiD Dataflow & Power

Gunther Haller

Research Engineering Group
Particle Physics and Astrophysics Division
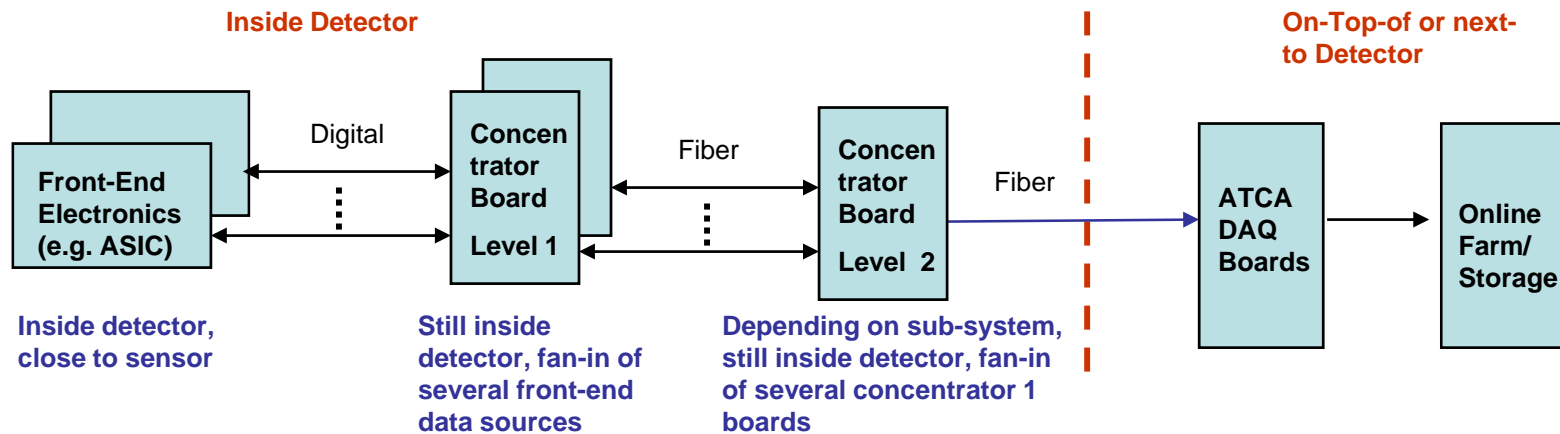**SLAC-Stanford University**

October 26, 2007

Overview

- **Overall dataflow architecture**
- **Expected data rates**
- **Example implementation**
  - ATCA system
  - EM CAL sub-system
- **Power system**

Gunther Haller

haller@slac.stanford.edu

Electronics Architecture

Inside Detector

On-Top-of or next-to Detector

Front-End Electronics (e.g. ASIC) — Digital → Concentrator Board Level 1 — Fiber → Concentrator Board Level 2 — Fiber → ATCA DAQ Boards → Online Farm/Storage

Inside detector, close to sensor

Still inside detector, fan-in of several front-end data sources

Depending on sub-system, still inside detector, fan-in of several concentrator 1 boards

- **Total data rate from each front-end relatively small, thus can combine data from several front-ends to reduce number of connections to the outside of the detector**

- **Front-End ASICs/electronics transmit event data to concentrator 1 boards**
  - **Digital interface (optical or electrical, e.g. LVDS)**
  - **Concentrator 1 boards close to front-end, combining data-streams from several front-end ASICs**
  - **Zero-suppression either at front-end or on concentrator 1 boards**
    - **No additional processing needed at this stage**

- **Event data from concentrator 1 boards are combined in concentrator 2 boards**
  - **Multiplexing of concentrator 1 board event data onto fewer fibers**

- **Event data is transmitted to top or side of detector**
  - **ATCA crate (see later) to process and switch data packets**
  - **Online farm for filtering (if necessary)**

Data-Rates

- Question is what are the data-rates coming from each sub-system?
    - Influences architecture for readout
- Assume zero-suppression of data towards the front-end (ASIC's or concentrator 1 board)
- See table on next slides
    - Mostly driven by noise or background hits

Bandwidth into DAQ/Online from each Sub-System

Sub-System	Mean # Hits/Train	#of bytes/hit at level 0	Bandwidth (bits/sec) (5 trains/sec)
Tracker Barrel	$2*10^7$	18*	15G
Tracker Endcap	$8*10^6$	18*	6G
EM Barrel	$4*10^7$	8	13G
EM Endcap	$6*10^7$	8	20G
HAD Barrel	$2*10^7$	8	6G
HAD Endcap	$4*10^6$	8	1.3G
Muon Barrel	$1*10^5$	8	32M
Muon Endcap	$1*10^5$	8	32M
Vertex			10M (dominated by layer 1)
LumCal/BeamCal	tbd		tbd
Total			**~60G**

# of bytes for address: 4 bytes, time: 2 bytes, ADC: 2 bytes
*: tracker assumes nearest neighbor logic, adds 2x8 bytes

- Nominal ~60 Gbits/s data rate (750 Mbyte/s)
  - Need to provide margin, e.g. factor of 4
- Example: DAQ being prototyped for LCLS is very scalable, bandwidth is fine, see later slides

Gunther Haller

haller@slac.stanford.edu

DAQ Sub-System

- **Based on ATCA (Advanced Telecommunications Computing Architecture)**
  - Next generation of "carrier grade" communication equipment
  - Driven by telecom industry
  - Incorporates latest trends in high speed interconnect, next generation processors and improved Reliability, Availability, and Serviceability (RAS)
  - Essentially instead of parallel bus backplanes, uses high-speed serial communication and advanced switch technology within and between modules, plus redundant power, etc

Gunther Haller

haller@slac.stanford.edu

ATCA Crate

- ATCA used for e.g. SLAC LUSI (LCLS Ultra-fast Science Instruments) detector readout for Linac Coherent Light Source hard X-ray laser project
  - Based on 10-Gigabit Ethernet backplane serial communication fabric
- 2 custom boards
  - Reconfigurable Cluster Element (RCE) Module
    - Interface to detector
    - Up to 8 x 2.5 Gbit/sec links to detector modules
  - Cluster Interconnect Module (CIM)
    - Managed 24-port 10-G Ethernet switching
- One ATCA crate can hold up to 14 RCE's & 2 CIM's
  - Essentially 480 Gbit/sec switch capacity
  - SiD needs only ~ 320 Gbit/sec including factor of 4 margin
  - Plus would use more than one crate (partitioning)

Gunther Haller

haller@slac.stanford.edu
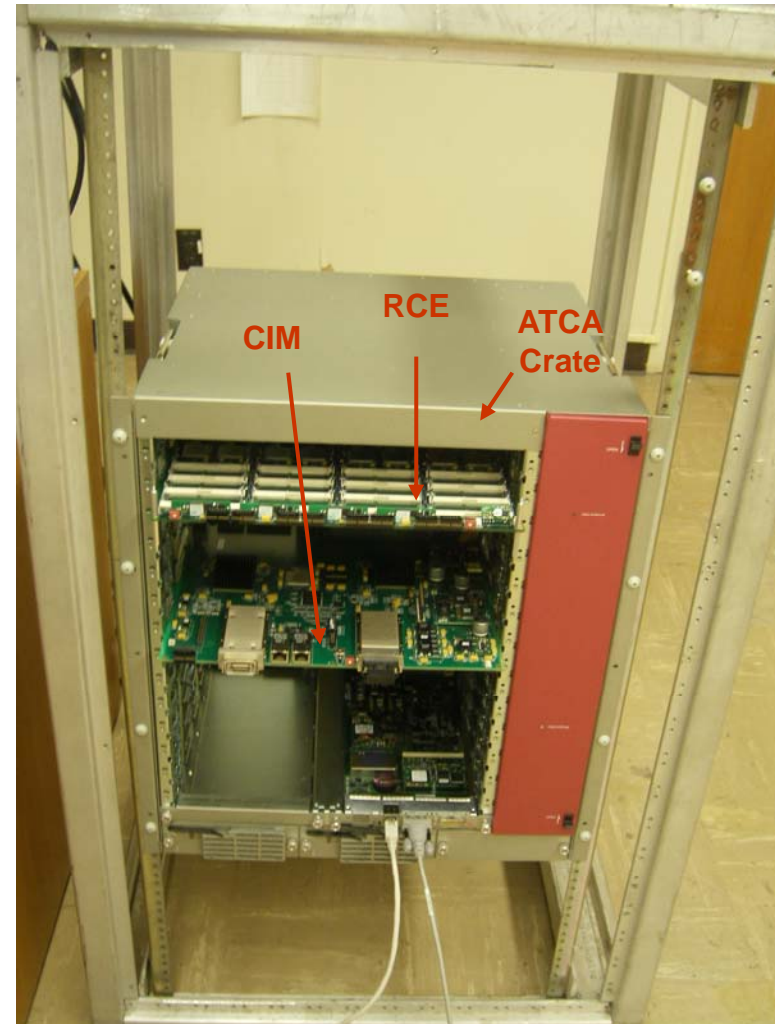
Reconfigurable Cluster Element (RCE) Boards

- Addresses performance issues with off-shelf hardware
    - Processing/switching limited by CPU-memory sub-system and not # of MIPS of CPU
    - Scalability
    - Cost
    - Networking architecture
- Reconfigurable Cluster Element module with 2 each of following
    - Virtex-4 FPGA
        - 2 PowerPC processors IP cores
    - 512 Mbyte RLDRAM
    - 8 Gbytes/sec cpu-data memory interface
    - 10-G Ethernet event data interface
    - 1-G Ethernet control interface
    - RTEMS operating system
    - EPICS
    - up to 512 Gbyte of FLASH memory

Rear Transition Module

Reconfigurable Cluster Element Module

Gunther Haller

haller@slac.stanford.edu

Cluster Interconnect Module

- Network card
  - 2 x 24-port 10-G Ethernet Fulcrum switch ASICs
  - Managed via Virtex-4 FPGA
- Network card interconnects up to 14 in-crate RCE boards
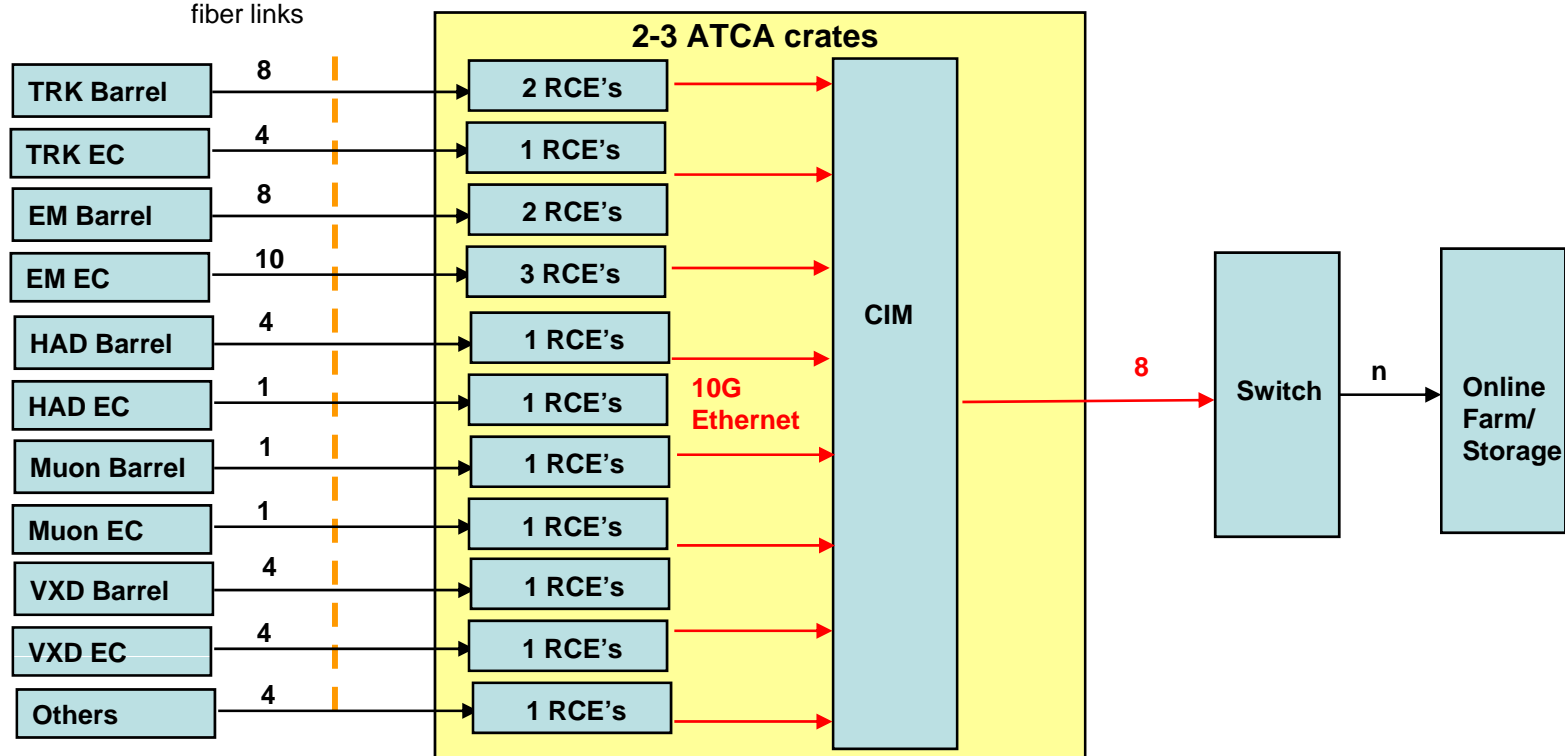- Network card interconnects multiple crates or farm machines

DAQ Architecture, Minimum Number of Reconfigurable Cluster Elements

Outside Detector

3 Gbits/sec PGP fiber links

2-3 ATCA crates

Detector	Links	RCE's
TRK Barrel	8	2 RCE's
TRK EC	4	1 RCE's
EM Barrel	8	2 RCE's
EM EC	10	3 RCE's
HAD Barrel	4	1 RCE's
HAD EC	1	1 RCE's
Muon Barrel	1	1 RCE's
Muon EC	1	1 RCE's
VXD Barrel	4	1 RCE's
VXD EC	4	1 RCE's
Others	4	1 RCE's

CIM

10G Ethernet

8 → Switch → n → Online Farm/ Storage

- Could be more 3-G links depending what partitioning is best for on-detector electronics
- Just need to add more RCE's or even a few more ATCA crates
- 1 ATCA crate can connect to up to 14 x 8 Input fibers
- Bandwidth no issue (each ATCA crate can output data to online farm at > 80 Gbit/s)
- No need for data reduction in SiD DAQ, can transfer all data to online processing farm blades
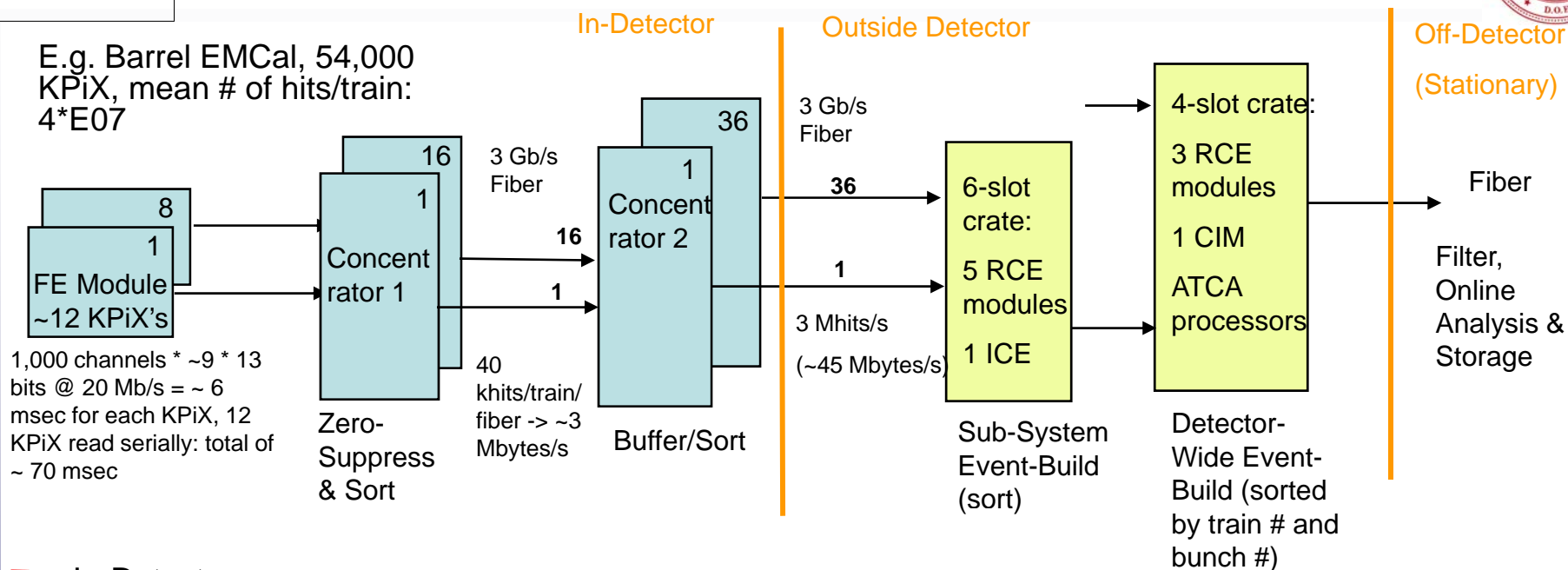
Partitioning

- Although 2 or 3 ATCA crates could handle all the SiD detector data
  - could use one crate for each sub-system for partitioning
    - 2 to 14 slot crates available
    - E.g. one 2-slot crate for each sub-system
    - Total of 1 rack for complete DAQ

EM Barrel Example

In-Detector · **Outside Detector** · **Off-Detector (Stationary)**

E.g. Barrel EMCal, 54,000 KPiX, mean # of hits/train: 4*E07

FE Module ~12 KPiX's (8, 1)

1,000 channels * ~9 * 13 bits @ 20 Mb/s = ~ 6 msec for each KPiX, 12 KPiX read serially: total of ~ 70 msec

Concentrator 1 (16, 1) — Zero-Suppress & Sort

3 Gb/s Fiber — 16 — 1 — 40 khits/train/fiber -> ~3 Mbytes/s

Concentrator 2 (36, 1) — Buffer/Sort

3 Gb/s Fiber — 36 — 1 — 3 Mhits/s (~45 Mbytes/s)

6-slot crate: 5 RCE modules 1 ICE — Sub-System Event-Build (sort)

4-slot crate: 3 RCE modules 1 CIM ATCA processors — Detector-Wide Event-Build (sorted by train # and bunch #)

Fiber — Filter, Online Analysis & Storage

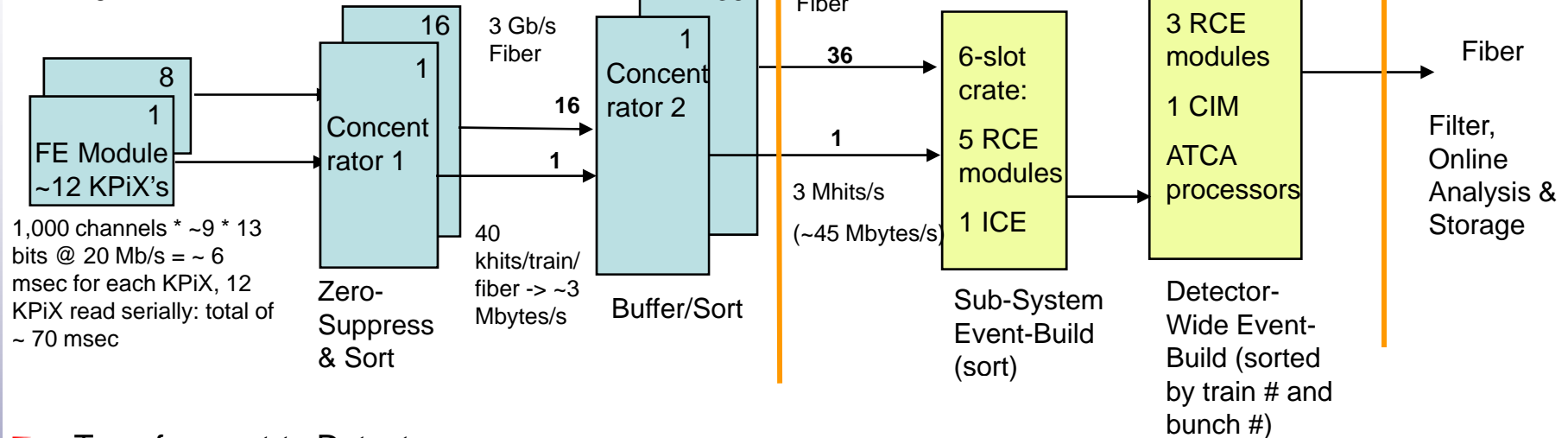- In-Detector
  - KPIX ASIC as front-end (1,024 channels, serial datain/clock/dataout LVDS interface)
  - Concentrator 1 (FPGA based): zero-suppress. Sort total 740 hits/train/Kpix -> 2.8 Mbytes/s for 96 KPIX's (720 hits/train/KPIX * 5 trains/s * 96 KPIX * 8 bytes)
  - Concentrator 2 (FPGA based): Sort total of ~45 Mbytes/s
    - Total out of detector: 1.6 Gbytes/sec

EM Barrel Example

E.g. Barrel EMCal, 54,000 KPiX, mean # of hits/train: 4*E07

8
1
FE Module ~12 KPiX's

1,000 channels * ~9 * 13 bits @ 20 Mb/s = ~ 6 msec for each KPiX, 12 KPiX read serially: total of ~ 70 msec

16
1
Concent rator 1

Zero-Suppress & Sort

3 Gb/s Fiber

16

1

40 khits/train/ fiber -> ~3 Mbytes/s

36
1
Concent rator 2

Buffer/Sort

3 Gb/s Fiber

36

1

3 Mhits/s

(~45 Mbytes/s)

6-slot crate:

5 RCE modules

1 ICE

Sub-System Event-Build (sort)

4-slot crate:

3 RCE modules

1 CIM

ATCA processors

Detector-Wide Event-Build (sorted by train # and bunch #)

Fiber

Filter, Online Analysis & Storage

- Top-of or next-to Detector
  - Readout to outside-Detector crates via 3 Gbit/s fibers
    - Single 6-slot crate to receive 36 fibers: 5 RCE modules + 1 Cluster Interconnect Module (CIM)
  - Total out of EM Barrel partition: 1.6 Gbytes/s
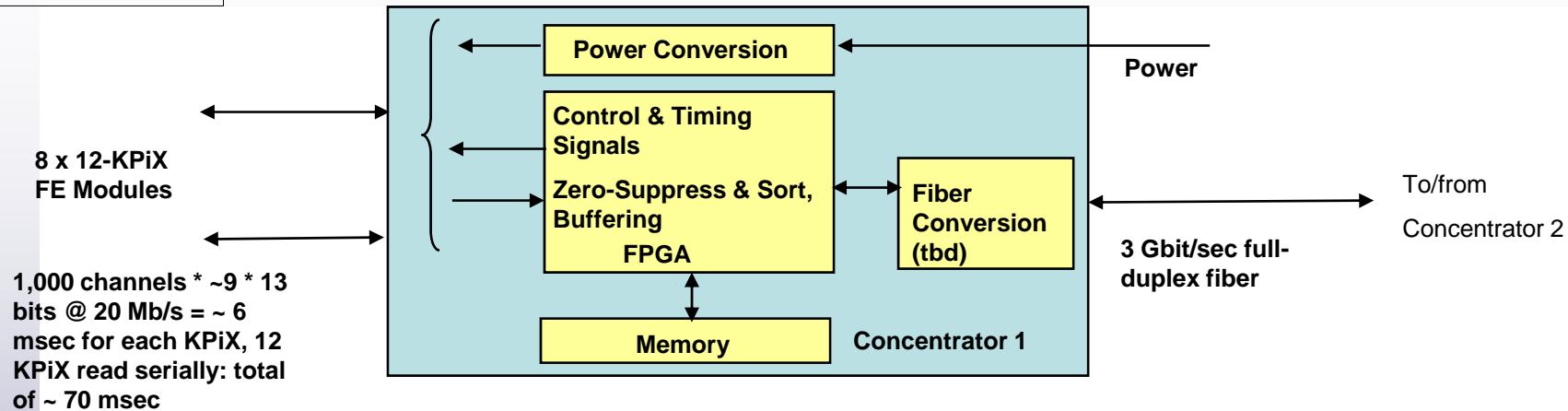   - Available bandwidth: > 80 Gbit/s (and is scalable)
  - Sorting, data reduction
- Can be switched into ATCA processors for data-filtering/reduction or online farm
  - A few 10-G Ethernet fibers off detector
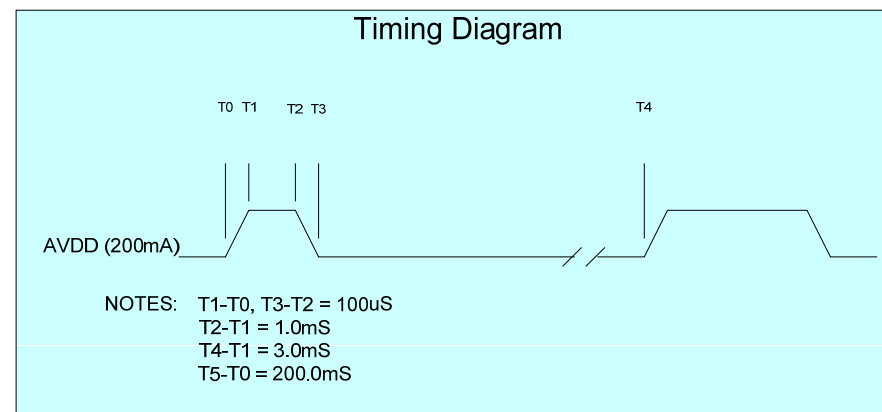
Concentrator-1

Power Conversion

Power

Control & Timing Signals

8 x 12-KPiX FE Modules

Zero-Suppress & Sort, Buffering

FPGA

Fiber Conversion (tbd)

To/from

Concentrator 2

3 Gbit/sec full-duplex fiber

1,000 channels * ~9 * 13 bits @ 20 Mb/s = ~ 6 msec for each KPiX, 12 KPiX read serially: total of ~ 70 msec

Memory   **Concentrator 1**

- Input Data Rate
  - 8 x 20 Mbit/sec
- Buffer Memory
  - 16 bytes from each KPiX channel (time/amplitude for up to 4 samples/channels)
  - Before suppression: 96 KPiX x 1,000 channels x 16 bytes x 4 samples: 6 Mbytes/train
  - After zero-suppression: < 800 hits/train/KPiX => 384 khits/sec x 8 bytes (ID/amplitude/time)
    - => 2.8 Mbytes/s
- Output Data Rate
  - Unsuppressed: 30 Mbyte/sec or 2.4 Gbit/sec
  - Suppressed: 22 Mbit/s
  - Run standard 3 Gbit/sec link, no issue even without suppression
- IO to concentrator 2
  - Full-duplex fiber
  - Control/monitoring/event data on same link
- Concentrator 2 board similar (not shown)

Power Supply Timing (use EMCAL KPIX as example)

- **Timing**
    - Period = 200mS
    - AVDD is pulsed internal to KPiX for 1.0mS
    - DVDD = DC
- **AVDD – per KPiX**
    - 200mA peak
    - 10 mW average
- **DVDD**
    - 2mA average
    - 10mW average

Timing Diagram

T0 T1   T2 T3   T4

AVDD (200mA)

NOTES: T1-T0, T3-T2 = 100uS
        T2-T1 = 1.0mS
        T4-T1 = 3.0mS
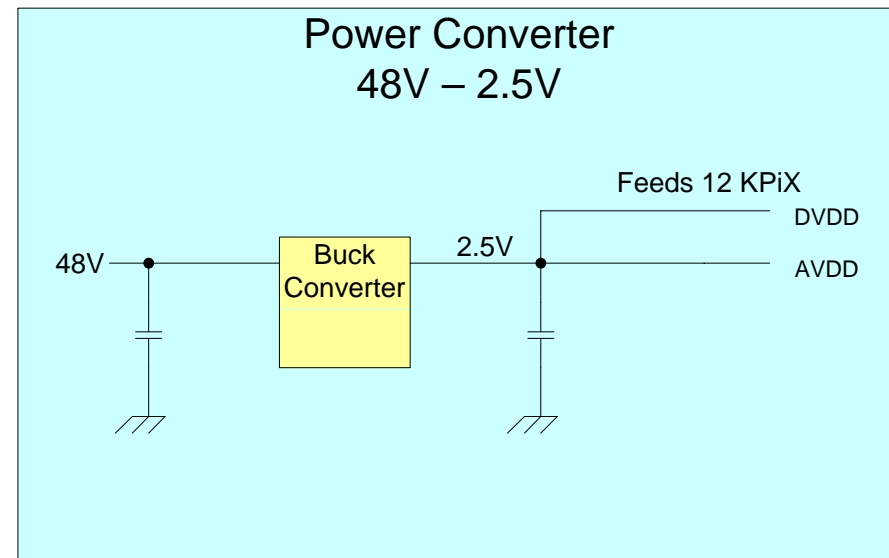   T5-T0 = 200.0mS

Gunther Haller

haller@slac.stanford.edu

Power Converter Block Diagram (located on concentrator 1 board)

- Example:
  - Distribute 48V via concentrator 2 boards to concentrator 1 boards
- On concentrator 1 board:
  - Input Power
    - 48 Volts
  - Output Power
    - 2.5 Volts @ 2.5Amps peak
    - 240mW average
  - High frequency buck
    - > 1.0MHz switching
    - 1.0uH- 10uH air
      core inductor
    - AVDD droop < 100mV
    - 48 volt droop < 5 volts
  - Efficiency > 70%
  - Can run higher input V (e.g. 400V) if needed

Power Converter
48V – 2.5V

Feeds 12 KPiX

48V   Buck Converter   2.5V   DVDD

AVDD

Gunther Haller

haller@slac.stanford.edu

Power System

- Power for 96 KPiX is about 2 watts. At 70% efficiency the input power is 1.3*2=2.6 watts input.

- The capacitance on the input of the converters should smooth charging period over the 200mS.

- Set the input capacitor for a 5 volt drop during AVDD peak power. Letting the voltage to drop would minimize the capacitor size.

- The average current is to one concentrator 1 board is 2.6 watts/48 volts = 0.055 amps.

- Concentrator 2 boards could distribute power to concentrator 1 boards
  - 16 Concentrator 1 board for each concentrator 2 boards
  - 0.88A to each concentrator 2 board

- Wire resistance and power in cable for 20 meters (10m distance, x 2 for return)
  - AWG   Ohms/20 meters   voltage drop   power loss in wire
  - 26   2.66   2.34   2W
  - 22   1.06   0.88   0.77W

- Total of 36 cables into detector (for 36 concentrator-2 boards)
  - Total power in all 36 cables: ~30W with 22-AWG (less if larger or parallel wires)
  - Total power from supply: ~ 1.5kW (or about 30A at ~50V) (plus concentrator 1 and 2 power)
  - Plus add concentrator 1 and 2 power (~700W for EMCAL)

Power System (con't)

■ As an example, table below assumes KPIX-based front-end for most sub-systems

Sub-System	# of sensors	#of pixels/sensor	# of KPiX (or equivalent)	Power for front-end (70% eff)
TrackerBarrel	5,788	1,800	10,000	250W
Tracker Endcap	2,556	1,800	2 * 3,500	200W
EM Barrel	91,270	1,024	54,000	1500W
EM Endcap	23,110	1,024	2 * 18,000	520W
HAD Barrel	2,800	10,000	27,000	800W
HAD Endcap	500	10,000	2 * 10,000	500W
Muon Barrel	2,300	100	5,000 (64-CH KPiX)	100W
Muon Endcap	2,800	100	2 * 1,600	100W
Vertex			tbd	tbd
LumCal			tbd	tbd
BeamCal			tbd	tbd

■ Add power for concentrator 1 and 2 boards (EMCAL is highest, ~700W)
  ■ Concentrator board mainly contains FPGA for sorting

Gunther Haller
haller@slac.stanford.edu

Summary

- **Event data rate for SiD can be handled by current technology, e.g. ATCA system being built for LCLS**
    - SiD data rate dominated by noise & background hits
    - Can use standard ATCA crate technology with e.g. existing SLAC custom cluster elements and switch/network modules
- **No filtering required in DAQ. Could move event data to online farm/off-line for further filtering/analysis**
    - Still: investigate filtering in ATCA processors
- **Power distribution at higher (48V to 400V) voltages to reduce wiring volume**

Gunther Haller

haller@slac.stanford.edu