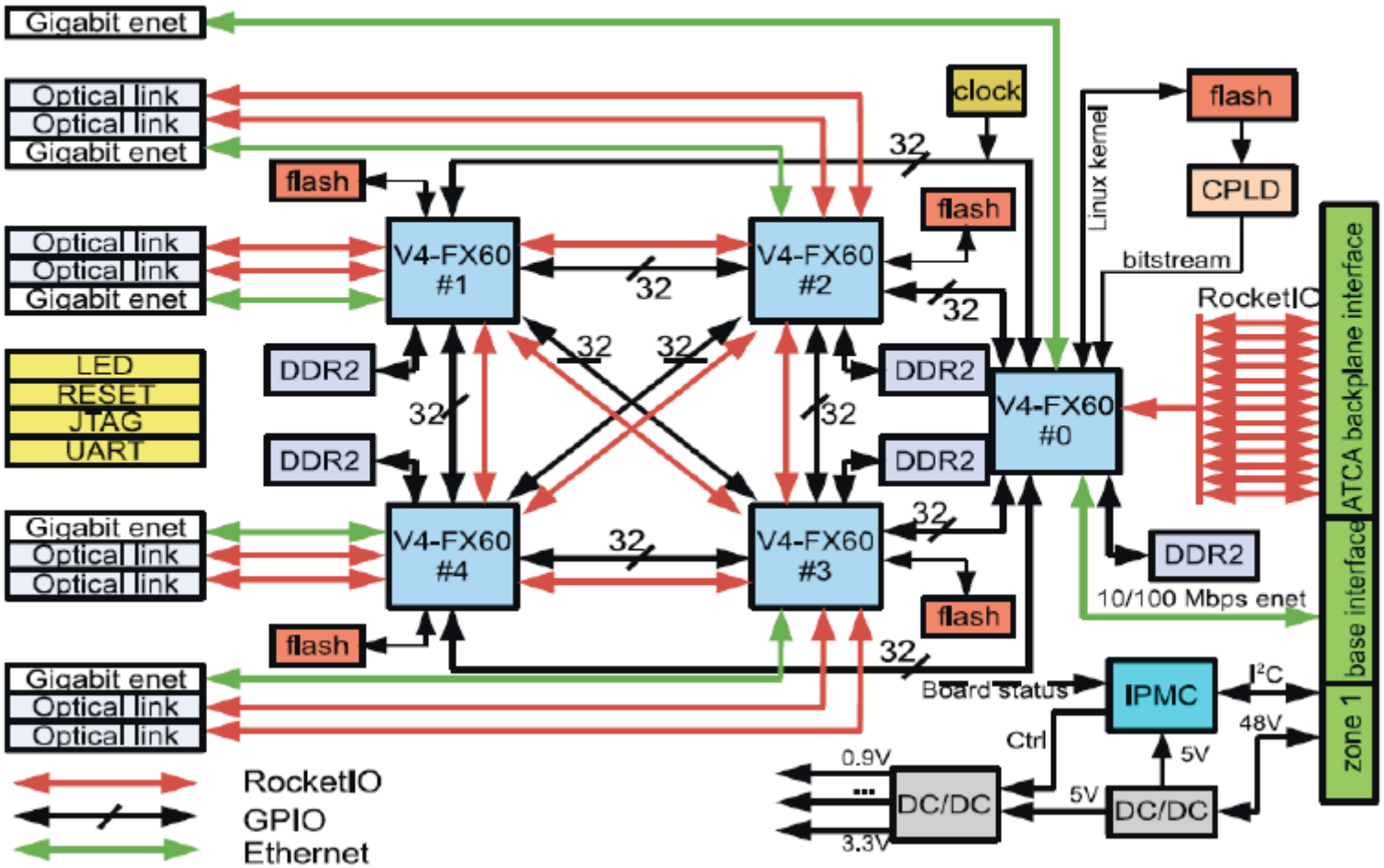
ATCA based PXD DAQ

Sören Lange, Univ. Giessen
PXD DAQ Workshop, Rain am Lech, 22.-23. April 2010

Compute Node Architecture



ATCA Option. Status. Tests done so far.

- **Compute Node PCB, 2nd iteration in testing since ~1 year**
  mass production for Hades experiment starting at IHEP
  18 boards planned for this year
- **Tests done with preliminary results**
  - Booting sequence for 5 FPGAs (by CPLD) o.k.
  - Optical link BER test at 1.6 Gb/s for ~160 hours o.k.
  - Full-mesh ATCA backplane BER test o.k.
  - ATCA power management by ATMEL Microcontroller add-on board
    2nd iteration of PCB produced in Giessen (master thesis)
    temperature readout o.k.
  - GB ethernet by PowerPC on FPGA
    (Linux and VxWorks tested, >0.2 Gb/s achieved)
  - Algorithms:
    1. Track finder for HADES implemented on ML405 board (Virtex-4)
    with real data, factor ~20 faster than PC
    2. Event decoder for Hades (real data in RAM,
    >0.15 Gb/s achieved, DMA size not optimized)
    3. Hough transform helix track finder with lookup tables
    for Panda FPGA implementation ongoing (Ph.D. thesis)
    important work for helix extrapolation of HLT track
    (even if stand-alone PXD+SVD track finder is abandoned)
- **About 1-2 years advantage in timescale**
  vs. PC based option

3-dim Algorithm

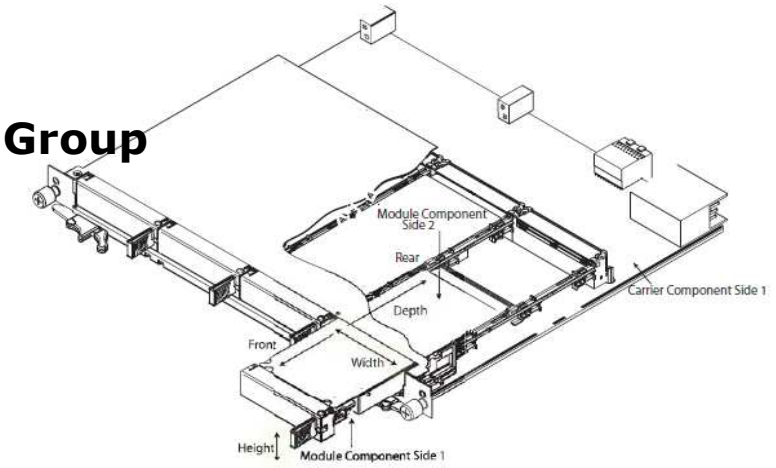
(see extra slides in the INDICO)

Content

- How to get more buffer?
  - new Compute Node PCB
  - more buffer in one ATCA crate
    - 1. event builder (factor 2)
    - 2. a few new thoughts on a PXD stand-alone data reduction (1st pass clean-up)
      > in order to free memory
    - 3. data compression („gzip") for PXD data
      not covered today, but not given up yet
      (note: mpeg-2 gives factor ~2 for SVD 1 data, shown at NARA DAQ Workshop 2003)
- Timelines
- (many) Questions

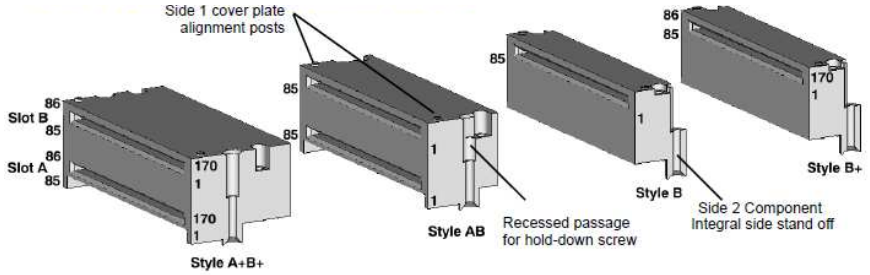
New Compute Node Proposal by IHEP and Giessen

- 2 specific hardware modifications required for Belle-II PXD readout only
  - 4 GB RAM per 1 FPGA
  - high speed optical links >5 Gbit/s (this means: new transceivers and new FPGAs)
- New approach by IHEP group: Carrier Board w/ Advanced Mezzanine Cards follow AMC.0 R2.0 specification of PICMG
  PCI Industrial Computers Manufacturers Group
- formfactor 7.4 x 18.0 cm
- 4 add-on cards per 1 Compute Node



AdvancedMC Connector

- various connector mount types are available for all AMC Connector styles, B, B+, AB, and A+B+.
- fabric Interface
  - 40 signal pairs allocated to the Fabric Interface
- System Management Interface
  - 9 contacts allocated to the System Management Interface
- AMC Clock Interface
  - 5 signal pairs allocated to the AMC Clock Interface
- JTAG Test Interface
  - 5 contacts allocated to the JTAG Test Interface
- Power/ Ground
  - 8 contacts allocated to Payload Power
  - 56 contacts to allocated to Logic Ground
- 2 contacts reserved



Connector Style	Interface to AMC Module	Number of Module Slots	Number of contact positions to Carrier	Number of contact rows on Carrier	Differential pairs	General purpose contacts	Power contacts	Ground contacts
B	Basic	1	85	1	19	11	8	28
B+	Extended	1	170	2	45	16	8	56
AB	Basic	2	170	2	38	22	16	56
A+B+	Extended	2	340	4	90	32	16	112

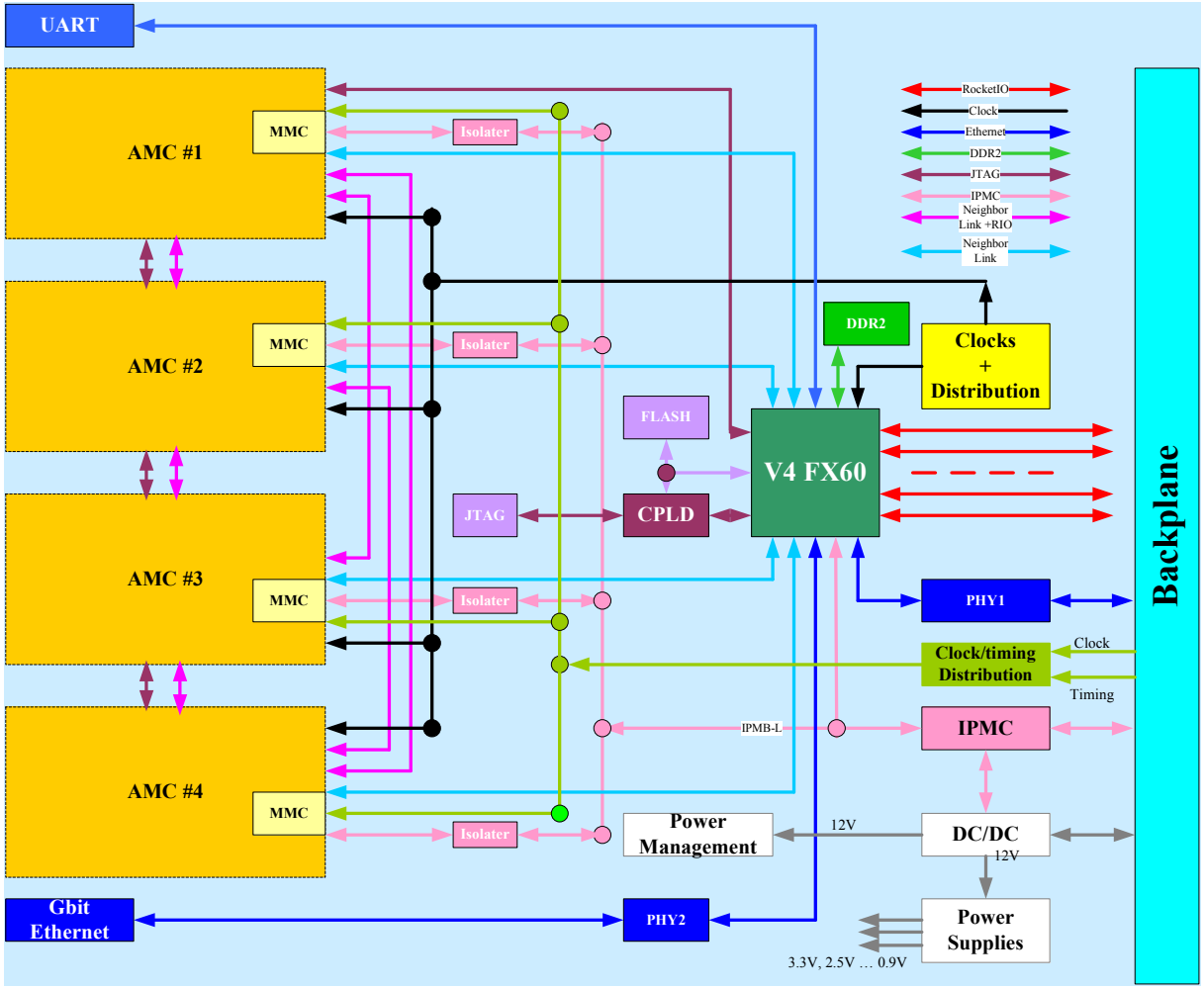
From PICMG AMC.0 R2.0

Development of Carrier Board

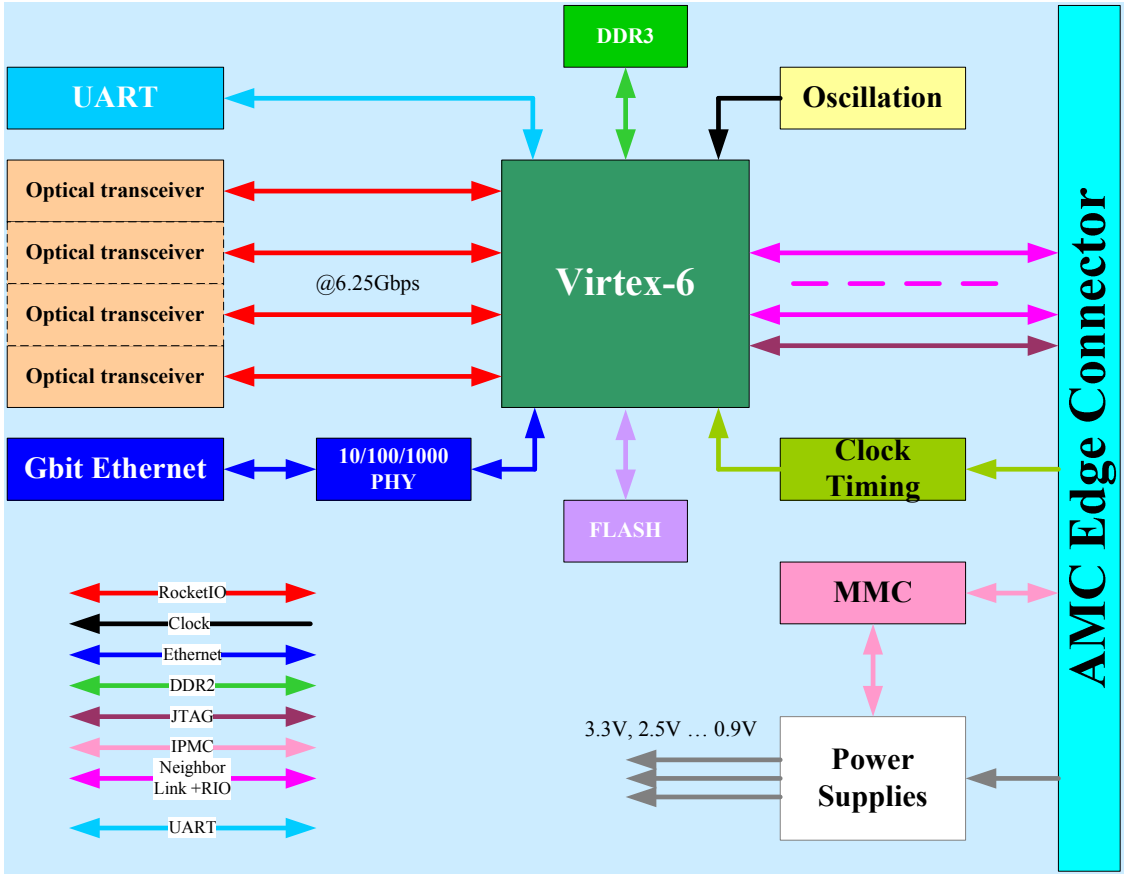
- Carrier Board with high bandwidth switch for neighbour-link
  - Virtex-4 FX60 based – functions test
    - 1 Virtex-4 FX60 FPGA
    - 2GB DDR2 SODIMM 400Mbps
    - 512Mb FLASH Memory
    - 13x RocketIOs @2Gbps to backplane
    - 2x Gbit Ethernet
    - **Schematic is ongoing, PCB will be delivered in June/July**
  - Virtex-6 based – high performance
    - 1 Virtex-6 FPGA
    - 2/ 4GB DDR3 SODIMM 800Mbps
    - 512Mb FLASH Memory
    - 13/26 RocketIOs @6.25Gbps to backplane
    - 2xGbit Ethernet

This PCB will be developed and existing in any case.

Carrier Board Rev.1



AMC Module 2, Virtex-6 based



Subevent Builder

Why event building?

- Only planned in the ATCA based option
- 1. To have 1st and 2nd layer of PXD correlated and combined for data reduction algorithm
- a. for 1st pass clean up algorithm (if possible)
- b. for extrapolation of HLT track otherwise 2 separate extrapolations
  - to 1st layer
  - to 2nd layer

  i.e. we are loosing the „tracklet“ in PXD for the data reduction
- 2. increase ATCA buffer memory by factor 2 (!) it means: if it can be realized, then we will have 4 GB memory already in the existing PCB solution (!)

PXS Subevent Builder

Note: backplane link is also RocketIO

40 FPGAs (data receive optical links) + 10 FPGAs (routing to backplane)

40 FPGAs for storing complete events

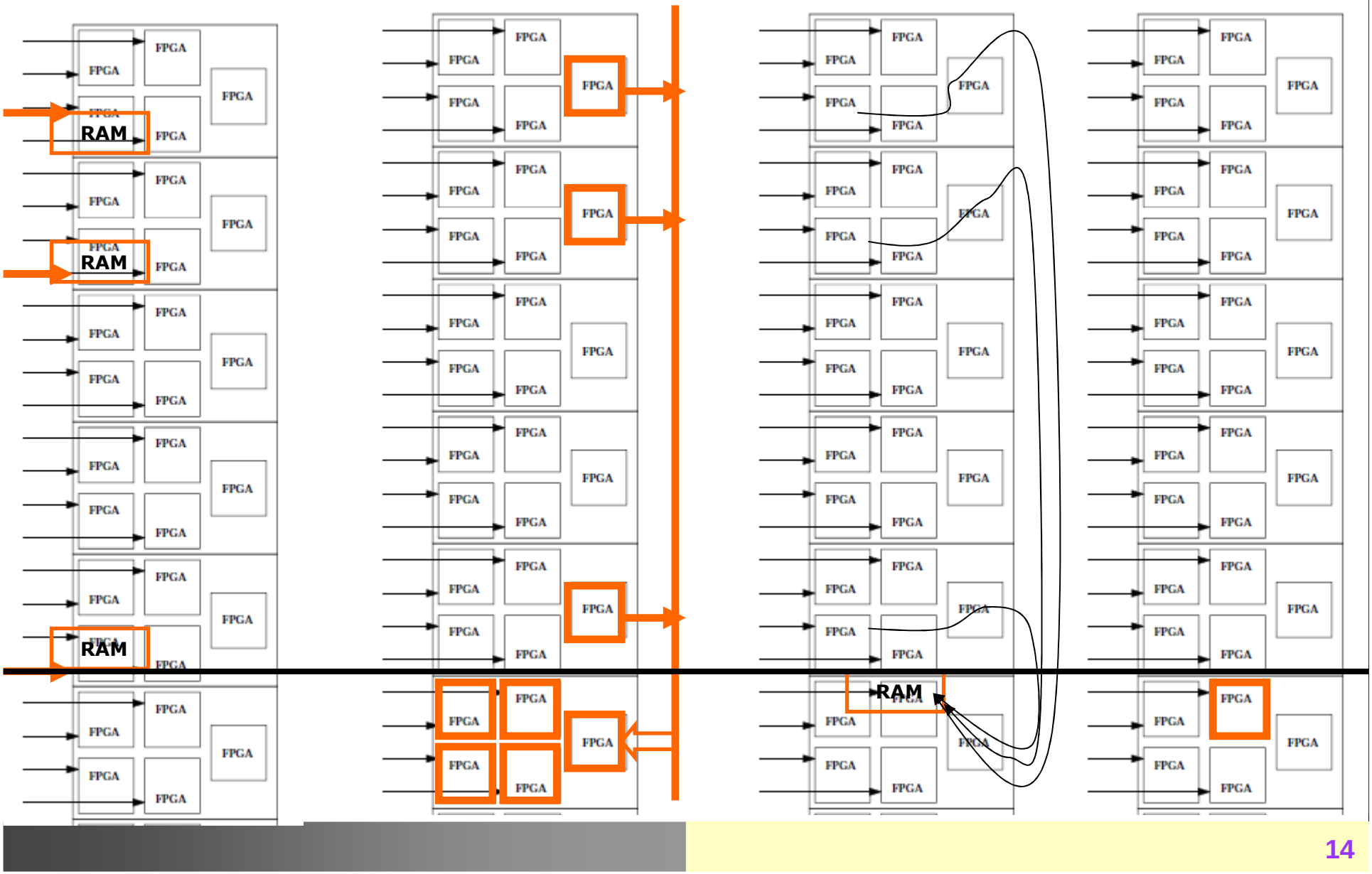
optical links here not connected

receive data
by optical links

Notify all
bookkeeper FPGAs

Move (copy and delete)
via ATCA backplane

wait for HLT



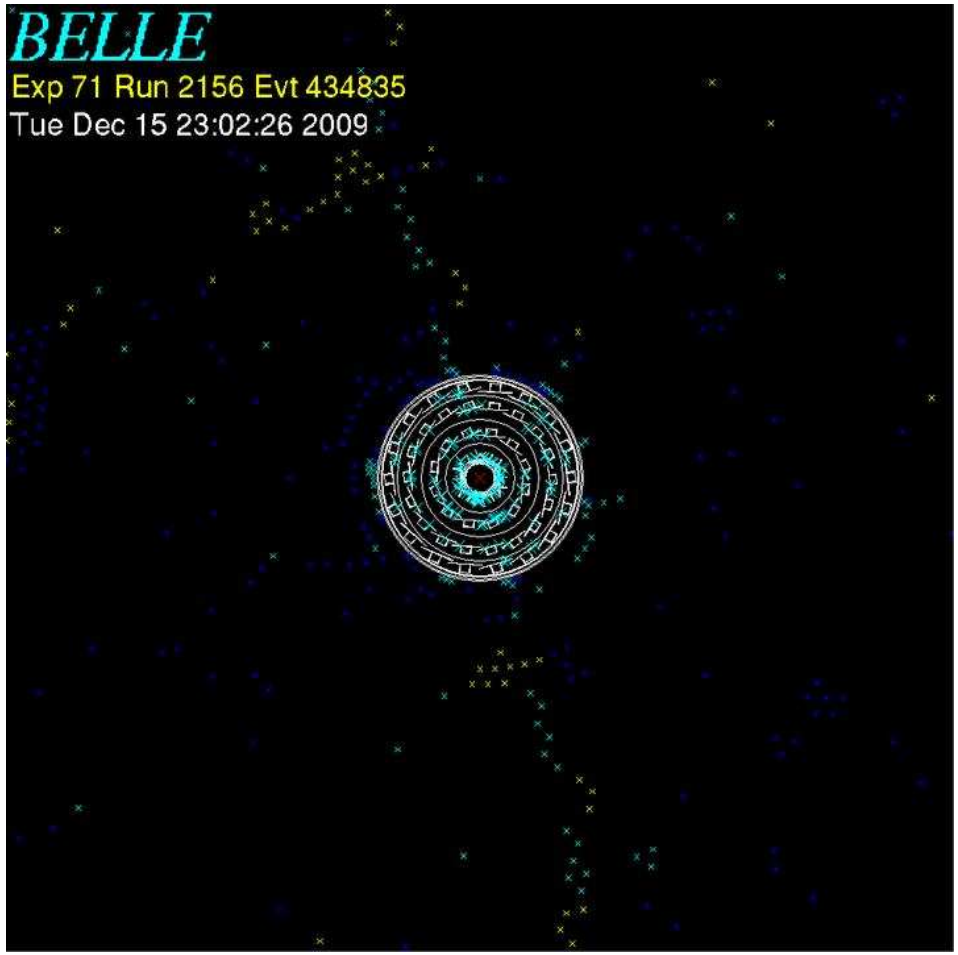
- The master FPGAs might also be able to provide a BUSY signal (RAM is full) to Nakao-san.

HOW TO SYNCHRONIZE DATA FRAGMENTS?

- There are 3 ways
  - 8-bit event number (also in SVD data)
  - 64-bit time stamp (Nakao-san, 19.11.2009)
  - PXD counter synchronized with the grand system clock (beam clock 508.89 MHz, divided by 5 on DHH), written into the data on frontend (Hans Krüger, DAQ MiniWorkshop Giessen 07.08.2009)
- guideline
  - for PXD subevent building, we use the PXD counter to identify event fractions, which belong to each other
  - For PXD+HLT matching, we use a lookup table [PXD counter :: event number], synchronize by event number, and write the event number into the data
    this lookup table needs to be synchronized
  - 64-bit timestamp will not be in the PXD events (unless DAQ group requires it for the global event building)

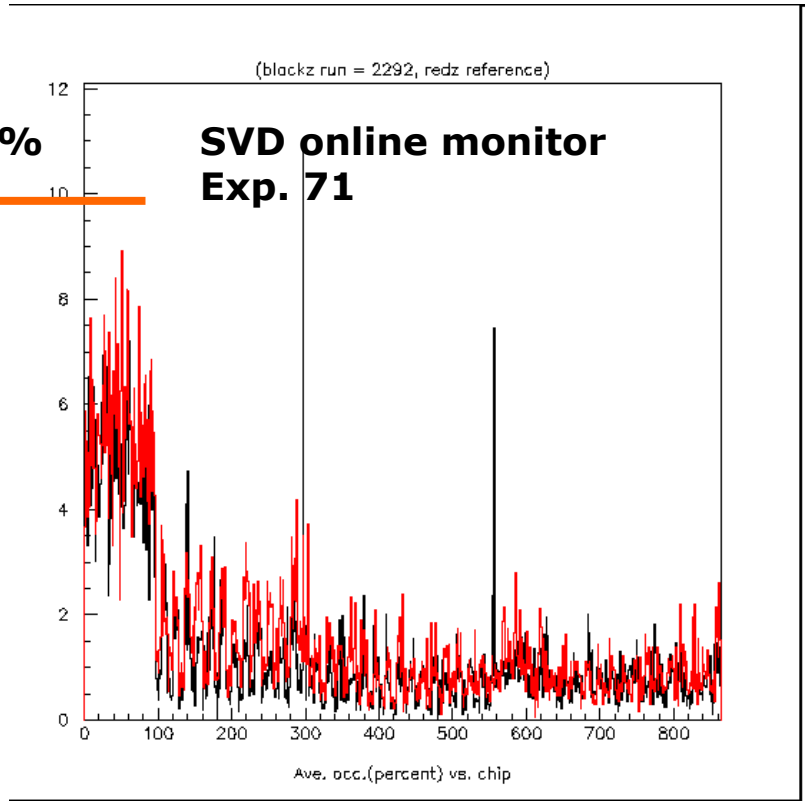
Can PXD clean-up events in a 1st pass stand-alone?

Even a factor 2 would be very effective.


BELLE
Exp 71 Run 2156 Evt 434835
Tue Dec 15 23:02:26 2009

SVD occupancies

- Kibayashi-san,
  VERTEX2005, NIMA 569(2006)5
- Occupancy of SVD2 @ 15 nb-1 s-1
  (5 kHz trigger rate)
- Radius Occupancy
  1.5 cm beampipe
  2.0 cm ~10%
  4.35 cm ~4%
  7.0 cm ~2%
  8.8 cm ~2%



Occupancy by background estimates for Belle-II

- **arXiv:0810.4084v1 [hep-ex] 22 Oct 2008**
  sBelle Design Study Report
- **occupancy is proportional to**
  - shaping time,
  - the area of each readout channel and
  - $1/r^2$
- **Current occupancy in SVD layer-1**
  seems over-proportional
  if this is ~ a * $1/r^2$
  what is a?
  1 cm? 10 cm ?
- **however, synchrotron radiation**
  is not coming from the IP
  why scale with $1/r^2$??

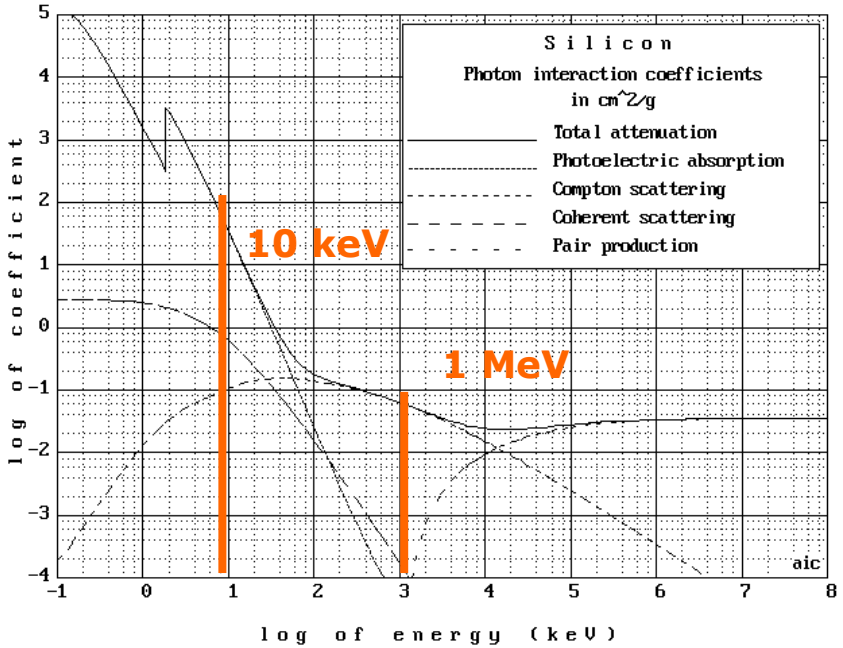
TABLE I: Relative contributions for individual detector components in the present KEKB / Belle

	SVD *(A , B)*	CDC	PID	ECL	KLM *(Barrel , E.C.-Fwd , E.C.-Bwd)*
Beam-Gas (HER)	0.56 , 0.42	0.25	0.45	0.40	0.15 , 0.00 , 0.25
(LER)	0.20 , 0.13	0.40	0.40	0.40	0.12 , 0.20 , 0.00
Touschek	0.04 , 0.03	0.10	0.10	0.10	0.03 , 0.05 , 0.00
SR Backscattering	0.18 , 0.40	0.24	–	–	–
SR Upstream	0.02 , 0.02	–	–	–	–
Radiative Bhabha	–	0.01	0.05	0.10	0.70 , 0.75 , 0.75

Fraction in Nano-Beam option expected smaller

Si Absorption Coefficient (Photon Energy Dependance)

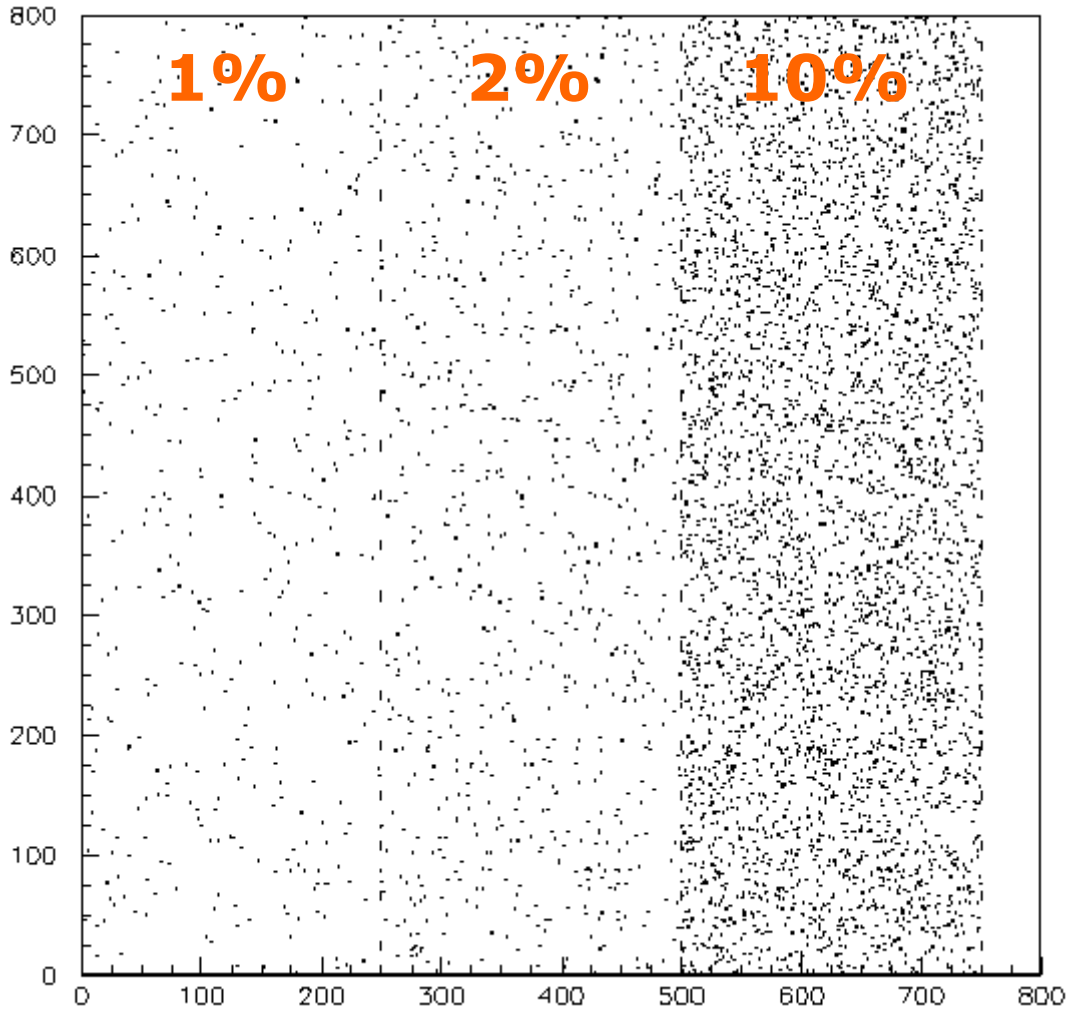
Synchrotron radiation is low $E_\gamma \simeq 10$ keV (?)



- $I = I_0 \exp(- \mu \rho x)$
- $\rho = 2.33$ g/cm3
- $E\gamma = 10$ keV
  $\mu = 10^1$ cm^2/g
  intensity on 1/e in 0,04 cm
  = 400 um
- $E\gamma = 1$ MeV
  $\mu = 10^{-1}$ cm^2/g
  intensity on 1/e in 4 cm

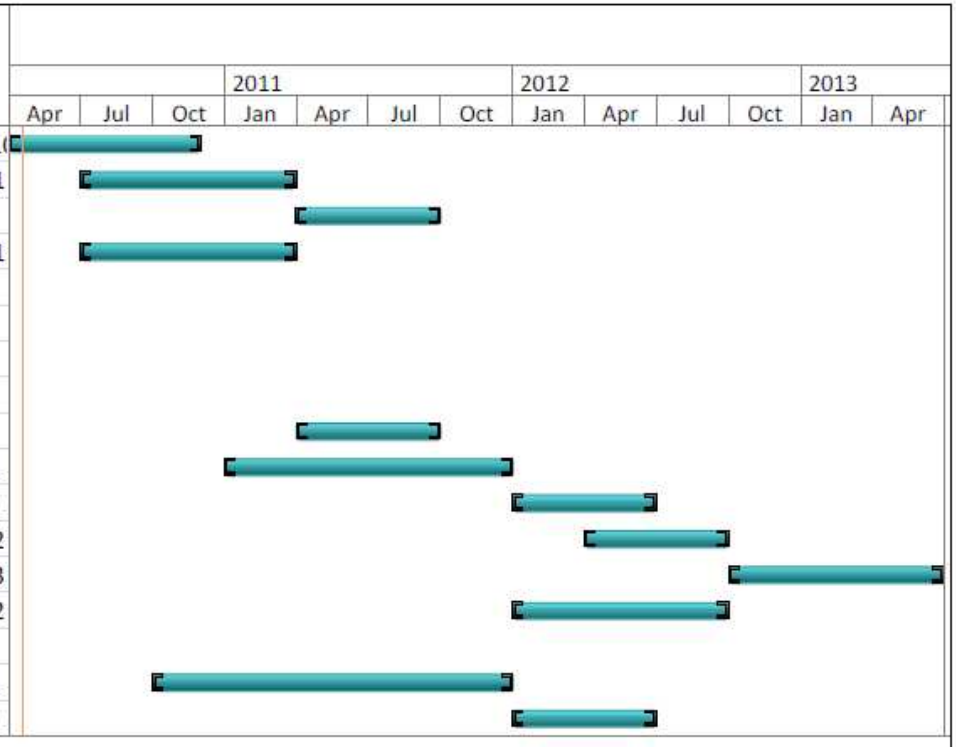
- 1st pass clean-up would increase ATCA buffer memory even more by removing data (free memory)
- We would like to spend some development time on studying this approach
- 2 ideas:
  - remove single pixels in 1st PXD layer?
    maybe low E_γ photons generate
    less 2-pixel clusters?
  - ϕ match and z match
    of 1st and 2nd layer
    even factor 2 would help!
    (would increase buffer from 5 sec to 10 sec)

PXD Occupancy, by random number generator



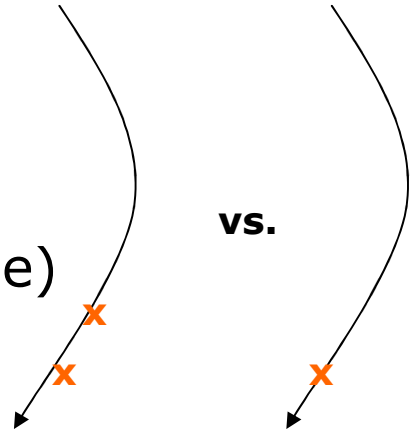
Timelines

ID	Task Name	Start	Finish
1	Data Receive	Thu 01.04.10	Wed 01.12.1(
2	Subevent Builder	Thu 01.07.10	Thu 31.03.11
3	Subevent Sorting (if required)	Thu 31.03.11	Fri 30.09.11
4	Data Reduction	Thu 01.07.10	Thu 31.03.11
5	1. 1st pass clean-up (phi/z match 1st/2nd layer)		
6	2. helix Extrapolation for HLT track		
7	3. alignment/calibration parameter handling		
8	4. event building (bookkeeping)		
9	Interface to HLT farm	Fri 01.04.11	Fri 30.09.11
10	Data Output (GB Ethernet), TCP/IP Stack	Fri 31.12.10	Sat 31.12.11
11	Prototype System (Readout of 1 Half-Module)	Sun 01.01.12	Sat 30.06.12
12	PCB Mass Production	Sun 01.04.12	Sun 30.09.12
13	1 Full System (Test at KEK)	Mon 01.10.1:	Sun 30.06.13
14	Offline Implementation of Data Raduction	Sun 01.01.12	Sun 30.09.12
15			
16	new PCB (Carrier-Board, AMC Card w/ Virtex 6)	Fri 01.10.10	Sat 31.12.11
17	Testing of new PCB (Softcore PowerPC)	Sun 01.01.12	Sat 30.06.12



Remarks about PC-based PXD DAQ

- A backup system is always a good idea.
- But PC based system might not be easier.
  ("PC is easy" argument does not apply here)
- PCI express must be studied
  RocketIO > FPGA
  vs.
  RocketIO > FPGA > PCI Express > RAM > Pentium
  even if multiple lanes (e.g. interrupt handling)
- has not been demonstrated yet by a prototype system
- is more expensive, factor ~2
  (advantage: KEK funds)
- no PXD subevent builder
  before data reduction
  (i.e. 1 HLT track w/ 2 PXD hits will not be possible)
  maybe achievable reduction factor smaller



Backup Slides

Modification of CN

- **If we increase # of links from 44 to 88:**
  no problem (in 1 ATCA shelf there are 112 links)
- **If we keep # of optical links to 44**
  optical links are tested for 1.6 Gbit/s
  we need <5 Gbit/s
  - change FPGA
    VIRTEX-4 FX60-10 $ 904,-
    VIRTEX-4 FX60-11 $1131,-
  - change optical link transceiver
    FTLF8528P2BCK $140,-
    FTLF8519P2BNL $ 45,-
- **Price per 1 compute node**
  increases by ~20%
  (from $ 8100,- to $ 9995,-)
  > must reduce # of CN from 14 to 11 to keep budget

Summary of Virtex-4 Family Features

- Three Families — LX/SX/FX
  - Virtex-4 LX: High-performance logic applications solution
  - Virtex-4 SX: High-performance solution for digital signal processing (DSP) applications
  - Virtex-4 FX: High-performance, full-featured solution for embedded platform applications
- Xesium™ Clock Technology
  - Digital clock manager (DCM) blocks
  - Additional phase-matched clock dividers (PMCD)
  - Differential global clocks
- XtremeDSP™ Slice
  - 18 x 18, two's complement, signed Multiplier
  - Optional pipeline stages
  - Built-in Accumulator (48-bit) and Adder/Subtracter
- Smart RAM Memory Hierarchy
  - Distributed RAM
  - Dual-port 18-Kbit RAM blocks
    - Optional pipeline stages
    - Optional programmable FIFO logic automatically remaps RAM signals as FIFO signals
  - High-speed memory interface supports DDR and DDR-2 SDRAM, QDR-II, and RLDRAM-II.
- SelectIO™ Technology
  - 1.5V to 3.3V I/O operation
  - Built-in ChipSync™ source-synchronous technology
  - Digitally controlled impedance (DCI) active termination
  - Fine grained I/O banking (configuration in one bank)
- Flexible Logic Resources
- Secure Chip AES Bitstream Encryption
- 90-nm Copper CMOS Process
- 1.2V Core Voltage
- Flip-Chip Packaging including Pb-Free Package Choices
- RocketIO™ 622 Mb/s to 6.5 Gb/s Multi-Gigabit Transceiver (MGT) [*FX only*]
- IBM PowerPC RISC Processor Core [*FX only*]
  - PowerPC 405 (PPC405) Core
  - Auxiliary Processor Unit Interface (User Coprocessor)
- Multiple Tri-Mode Ethernet MACs [*FX only*]

Summary of Virtex-5 FPGA Features

- Five platforms LX, LXT, SXT, TXT, and FXT
  - Virtex-5 LX: High-performance general logic applications
  - Virtex-5 LXT: High-performance logic with advanced serial connectivity
  - Virtex-5 SXT: High-performance signal processing applications with advanced serial connectivity
  - Virtex-5 TXT: High-performance systems with double density advanced serial connectivity
  - Virtex-5 FXT: High-performance embedded systems with advanced serial connectivity
- Cross-platform compatibility
  - LXT, SXT, and FXT devices are footprint compatible in the same package using adjustable voltage regulators
- Most advanced, high-performance, optimal-utilization, FPGA fabric
  - Real 6-input look-up table (LUT) technology
  - Dual 5-LUT option
  - Improved reduced-hop routing
  - 64-bit distributed RAM option
  - SRL32/Dual SRL16 option
- Powerful clock management tile (CMT) clocking
  - Digital Clock Manager (DCM) blocks for zero delay buffering, frequency synthesis, and clock phase shifting
  - PLL blocks for input jitter filtering, zero delay buffering, frequency synthesis, and phase-matched clock division
- 36-Kbit block RAM/FIFOs
  - True dual-port RAM blocks
  - Enhanced optional programmable FIFO logic
  - Programmable
    - True dual-port widths up to x36
    - Simple dual-port widths up to x72
  - Built-in optional error-correction circuitry
  - Optionally program each block as two independent 18-Kbit blocks
- High-performance parallel SelectIO technology
  - 1.2 to 3.3V I/O Operation
  - Source-synchronous interfacing using ChipSync™ technology
  - Digitally-controlled impedance (DCI) active termination
  - Flexible fine-grained I/O banking
  - High-speed memory interface support
- Advanced DSP48E slices
  - 25 x 18, two's complement, multiplication
  - Optional adder, subtracter, and accumulator
  - Optional pipelining
  - Optional bitwise logical functionality
  - Dedicated cascade connections
- Flexible configuration options
  - SPI and Parallel FLASH interface
  - Multi-bitstream support with dedicated fallback reconfiguration logic
  - Auto bus width detection capability
- System Monitoring capability on all devices
  - On-chip/Off-chip thermal monitoring
  - On-chip/Off-chip power supply monitoring
  - JTAG access to all monitored quantities
- Integrated Endpoint blocks for PCI Express Designs
  - LXT, SXT, TXT, and FXT Platforms
  - Compliant with the PCI Express Base Specification 1.1
  - x1, x4, or x8 lane support per block
  - Works in conjunction with RocketIO™ transceivers
- Tri-mode 10/100/1000 Mb/s Ethernet MACs
  - LXT, SXT, TXT, and FXT Platforms
  - RocketIO transceivers can be used as PHY or connect to external PHY using many soft MII (Media Independent Interface) options
- RocketIO GTP transceivers 100 Mb/s to 3.75 Gb/s
  - LXT and SXT Platforms
- RocketIO GTX transceivers 150 Mb/s to 6.5 Gb/s
  - TXT and FXT Platforms
- PowerPC 440 Microprocessors
  - FXT Platform only
  - RISC architecture
  - 7-stage pipeline
  - 32-Kbyte instruction and data caches included
  - Optimized processor interface structure (crossbar)
- 65-nm copper CMOS process technology
- 1.0V core voltage
- High signal-integrity flip-chip packaging available in standard or Pb-free package options

Summary of Virtex-6 FPGA Features

- Three sub-families:
  - Virtex-6 LXT FPGAs: High-performance logic with advanced serial connectivity
  - Virtex-6 SXT FPGAs: Highest signal processing capability with advanced serial connectivity
  - Virtex-6 HXT FPGAs: Highest bandwidth serial connectivity
- Compatibility across sub-families
  - LXT and SXT devices are footprint compatible in the same package
- Advanced, high-performance FPGA Logic
  - Real 6-input look-up table (LUT) technology
  - Dual LUT5 (5-input LUT) option
  - LUT/dual flip-flop pair for applications requiring rich register mix
  - Improved routing efficiency
  - 64-bit (or two 32-bit) distributed LUT RAM option per 6-input LUT
  - SRL32/dual SRL16 with registered outputs option
- Powerful mixed-mode clock managers (MMCM)
  - MMCM blocks provide zero-delay buffering, frequency synthesis, clock-phase shifting, input-jitter filtering, and phase-matched clock division
- 36-Kb block RAM/FIFOs
  - Dual-port RAM blocks
  - Programmable
    - Dual-port widths up to 36 bits
    - Simple dual-port widths up to 72 bits
  - Enhanced programmable FIFO logic
  - Built-in optional error-correction circuitry
  - Optionally use each block as two independent 18 Kb blocks
- High-performance parallel SelectIO™ technology
  - 1.2 to 2.5V I/O operation
  - Source-synchronous interfacing using ChipSync™ technology
  - Digitally controlled impedance (DCI) active termination
  - Flexible fine-grained I/O banking
  - High-speed memory interface support with integrated write-leveling capability
- Advanced DSP48E1 slices
  - 25 x 18, two's complement multiplier/accumulator
  - Optional pipelining
  - New optional pre-adder to assist filtering applications
  - Optional bitwise logic functionality
  - Dedicated cascade connections
- Flexible configuration options
  - SPI and Parallel Flash interface
  - Multi-bitstream support with dedicated fallback reconfiguration logic
  - Automatic bus width detection
- System Monitor capability on all devices
  - On-chip/off-chip thermal and supply voltage monitoring
  - JTAG access to all monitored quantities
- Integrated interface blocks for PCI Express® designs
  - Compliant to the PCI Express Base Specification 2.0
  - Gen1 (2.5 Gb/s) and Gen2 (5 Gb/s) support with GTX transceivers
  - Endpoint and Root Port capable
  - x1, x2, x4, or x8 lane support per block
- GTX transceivers: up to 6.6 Gb/s
  - Data rates below 480 Mb/s supported by oversampling in FPGA logic.
- GTH transceivers: 2.488 Gb/s to beyond 11 Gb/s
- Integrated 10/100/1000 Mb/s Ethernet MAC block
  - Supports 1000BASE-X PCS/PMA and SGMII using GTX transceivers
  - Supports MII, GMII, and RGMII using SelectIO technology resources
  - 2500Mb/s support available
- 40 nm copper CMOS process technology
- 1.0V core voltage (-1, -2, -3 speed grades only)
- Lower-power 0.9V core voltage option (-1L speed grade only)
- High signal-integrity flip-chip packaging available in standard or Pb-free package options

Summary of Spartan-6 FPGA Features

- Spartan-6 Family:
  - Spartan-6 LX FPGA: Logic optimized
  - Spartan-6 LXT FPGA: High-speed serial connectivity
- Designed for low cost
  - Multiple efficient integrated blocks
  - Optimized selection of I/O standards
  - Staggered pads
  - High-volume plastic wire-bonded packages
- Low static and dynamic power
  - 45 nm process optimized for cost and low power
  - Hibernate power-down mode for zero power
  - Suspend mode maintains state and configuration with multi-pin wake-up, control enhancement
  - Lower-power 1.0V core voltage (LX FPGAs, -1L only)
  - High performance 1.2V core voltage (LX and LXT FPGAs, -2, -3, and -4 speed grades)
- Multi-voltage, multi-standard SelectIO™ interface banks
  - Up to 1,050 Mb/s data transfer rate per differential I/O
  - Selectable output drive, up to 24 mA per pin
  - 3.3V to 1.2V I/O standards and protocols
  - Low-cost HSTL and SSTL memory interfaces
  - Hot swap compliance
  - Adjustable I/O slew rates to improve signal integrity
- High-speed GTP serial transceivers in the LXT FPGAs
  - Up to 3.125 Gb/s
  - High-speed interfaces including: Serial ATA, Aurora, 1G Ethernet, PCI Express, OBSAI, CPRI, EPON, GPON, DisplayPort, and XAUI
- Integrated Endpoint block for PCI Express designs (LXT)
- Low-cost PCI® technology support compatible with the 33 MHz, 32- and 64-bit specification.
- Efficient DSP48A1 slices
  - High-performance arithmetic and signal processing
  - Fast 18 x 18 multiplier and 48-bit accumulator
  - Pipelining and cascading capability
  - Pre-adder to assist filter applications
- Integrated Memory Controller blocks
  - DDR, DDR2, DDR3, and LPDDR support
  - Data rates up to 800 Mb/s (12.8 Gb/s peak bandwidth)
  - Multi-port bus structure with independent FIFO to reduce design timing issues
- Abundant logic resources with increased logic capacity
  - Optional shift register or distributed RAM support
  - Efficient 6-input LUTs improve performance and minimize power
  - LUT with dual flip-flops for pipeline centric applications
- Block RAM with a wide range of granularity
  - Fast block RAM with byte write enable
  - 18 Kb blocks that can be optionally programmed as two independent 9 Kb block RAMs
- Clock Management Tile (CMT) for enhanced performance
  - Low noise, flexible clocking
  - Digital Clock Managers (DCMs) eliminate clock skew and duty cycle distortion
  - Phase-Locked Loops (PLLs) for low-jitter clocking
  - Frequency synthesis with simultaneous multiplication, division, and phase shifting
  - Sixteen low-skew global clock networks
- Simplified configuration, supports low-cost standards
  - 2-pin auto-detect configuration
  - Broad third-party SPI (up to x4) and NOR flash support
  - Feature rich Xilinx Platform Flash with JTAG
  - MultiBoot support for remote upgrade with multiple bitstreams, using watchdog protection
- Enhanced security for design protection
  - Unique Device DNA identifier for design authentication
  - AES bitstream encryption in the larger devices
- Faster embedded processing with enhanced, low cost, MicroBlaze™ soft processor
- Industry-leading IP and reference designs