Report from Grünberg Workshop

= green mountain

Sören Lange, Universität Gießen

5th International Workshop
on DEPFET Detectors and Applications
29.09.-01.10.2010, Valencia, Spain

http://panda.physik.uni-giessen.de:8080/indico/conferenceDisplay.py?confId=30

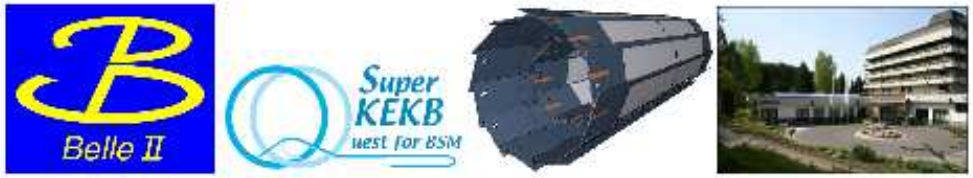
LOCAL: UTC

Belle-II PXD DAQ/Trigger Workshop

25-26 September 2010 | Sporthotel Grünberg

- Overview
- Scientific Programme
- Timetable
- Contribution List
- Author index
- Registration
  - Registration Form
- List of registrants

Home



Dates: from 25 September 2010 08:00 to 26 September 2010 18:00

Location: *Sporthotel Grünberg*
Am Tannenkopf 1
D-35305 Grünberg
Tel.: 0 64 01 / 80 20
Fax.: 0 64 01 / 80 21 66
info@sporthotel-gruenberg.de

Chairs: **Dr. Lange, Sören**
Prof. Kühn, Wolfgang

Additional info: LOCATION: **Map**

Arrival date: Friday, 24.09.2010
Departure date: Sunday, 26.09.2010 or Monday, 27.10.2010

Sessions will be Saturday 09:00-18:00 and Sunday 09:00-16:00.

last weekend

Participants

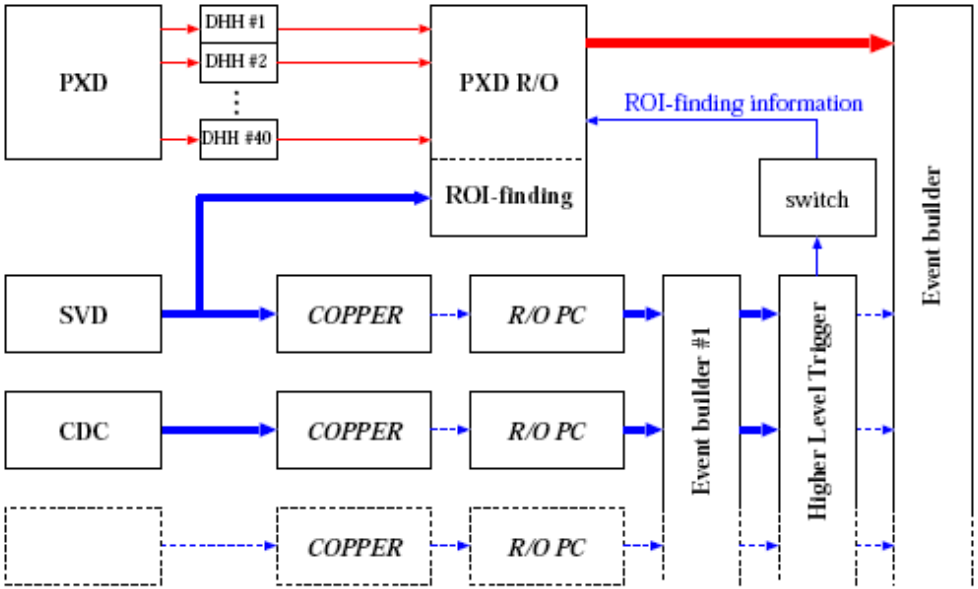
T. Higuchi-san, R. Itoh-san, N. Katayama-san, C. Kiesling-san,
P. Kodys-san, И. Коноров-san, W. Kühn-san, S.L.,
Zhen-An Liu-san, C. Heller-san, D. Münchow-san,
M. Nakao-san, S. Tanaka-san, and some more students
from Gießen (S. Fleischer, A. Kopp, M. Wagner)

Outline

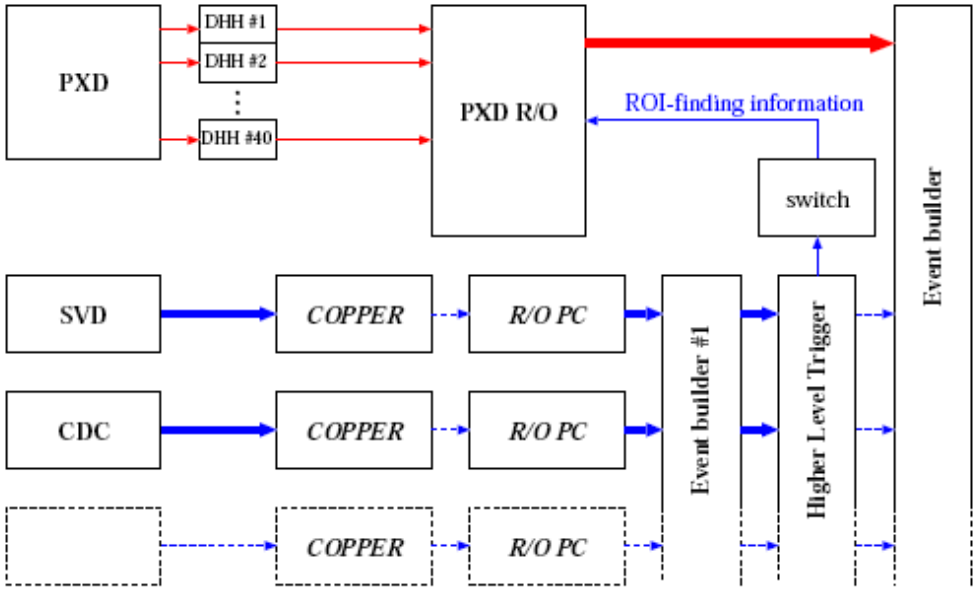
- Backend Readout System
  - ATCA based system („baseline option“)
  - PC based system („backup option“)
- DHH
- Timing and trigger distribution
- Injection veto
- HLT
- Roadmap until the decision ATCA vs. PC

ATCA based PC based

Baseline option

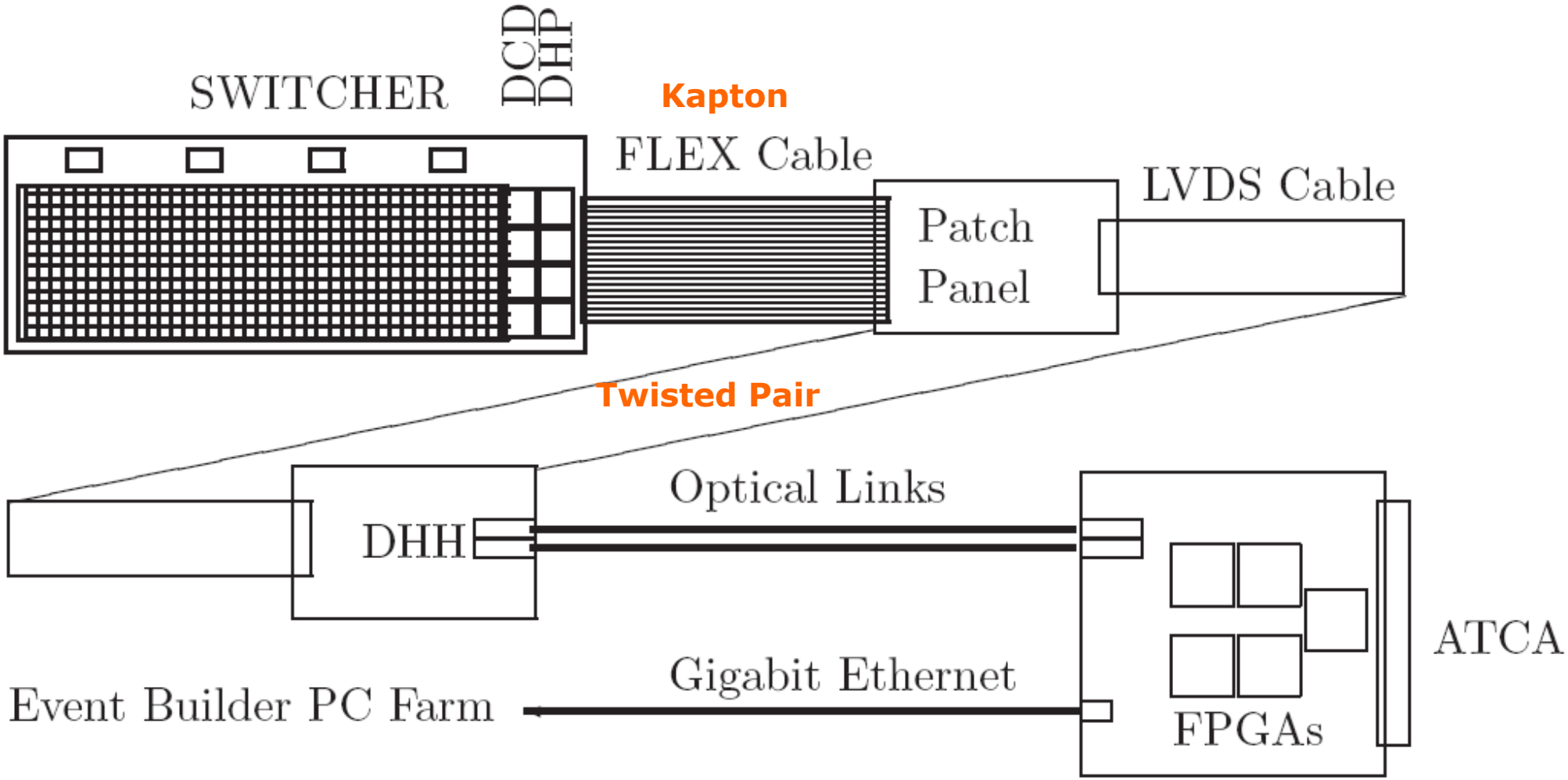


Backup option

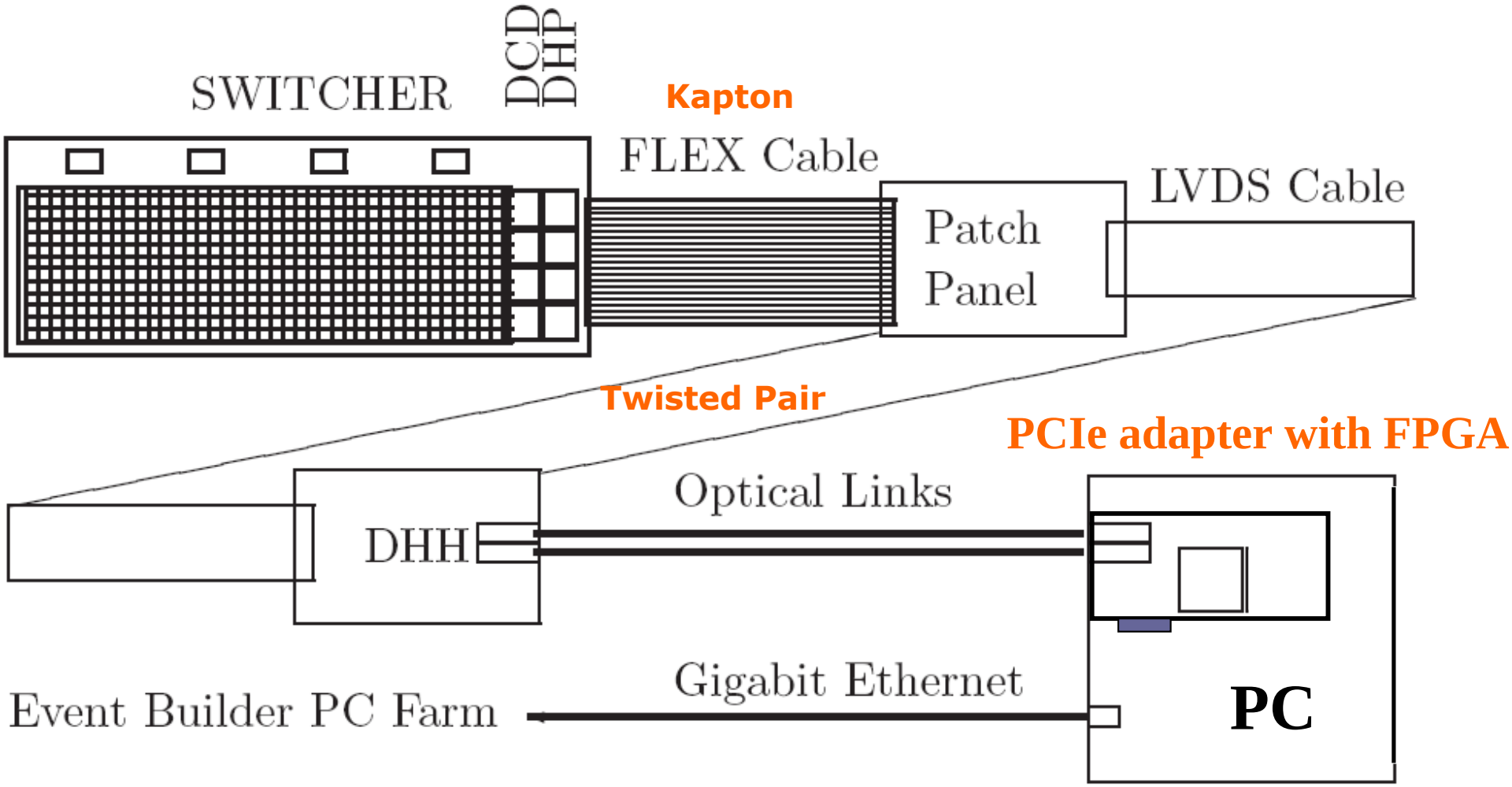


Decision: June 2011

ATCA-based System



PC-based System



What we estimated for TDR.

4.7.3 Data Rate Estimate

For the following estimate we assume a trigger rate of 30 kHz, corresponding to the highest luminosity of $\mathcal{L}$=0.8×10^{35} cm^{-2} s^{-1}. For an estimate of the required PXD readout data rate, we assume 40 PXD half ladders with 250 × 800 (or 0.2 × 10^6) pixels each and 8 × 10^6 pixels in total. The average cluster size is expected to be 2 pixels. Given a readout time of 20 μs, the average occupancy is $\simeq 1\%$, corresponding to $\simeq 8 \times 10^4$ fired pixels per frame. This occupancy still does not take into account any background. For a frame rate of 50 kHz and a trigger rate of 30 kHz, the expected reduction factor on the DHH can be calculated using Poisson statistics to a factor of 2.2. This amounts to a fired-pixel rate of $\simeq 1.8 \times 10^9$ Hz. Multiplying by a data size of 4 bytes per pixel for the encoded position and ADC charge information, we estimate a total data rate of $\simeq$ 58 GBit/s for the complete PXD. With 40 optical links, this implies $\simeq$ 1.44 GBit/s (or $\simeq$ 180 MByte/s) per optical link. However, there will be a significant contribution to the occupancy by radiative QED events of the type $e^+e^- \rightarrow e^+e^-\gamma$, beam gas events, synchrotron radiation from the upstream and downstream dipoles, and Touchek effect. For the background, we take into account a factor of three safety margin, i.e., the system will be prepared to cope with a maximum occupancy of 3%. If the background increases the occupancy to values above 3%, PXD subevents will have to be truncated. Precise track vertex reconstruction in events with such high occupancy degrades in any case because of too high combinatorics and subsequently an increased fraction of wrong PXD hit to track assignments.

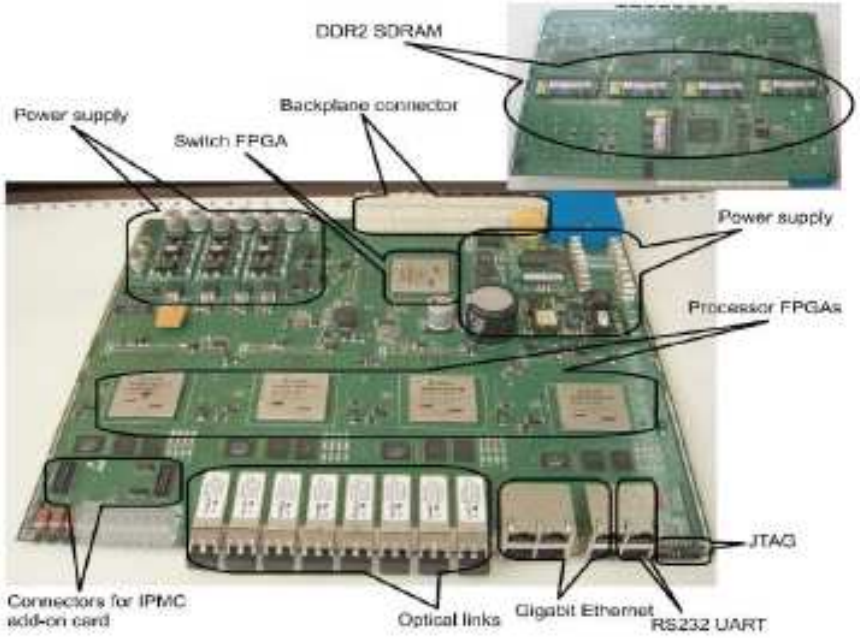
ATCA and PC Systems

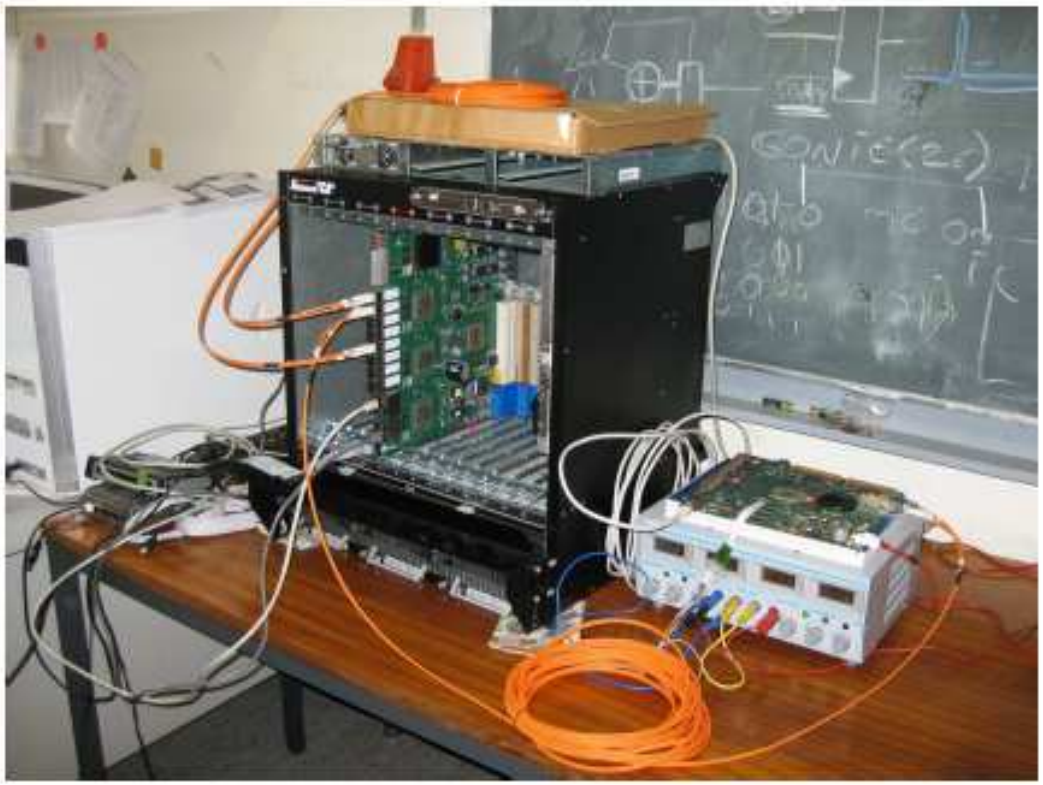
- Both systems:
  - FPGA based
  - get PXD data by optical link and RocketIO
  - buffer and wait for HLT decision (latency <5 seconds)
    → HLT sends ROI (regions-of intertest)
    → hits are deleted, if outside ROI
- Otherwise there are a few differences …

ATCA based system

(baseline option)

ATCA based system





ATCA based system

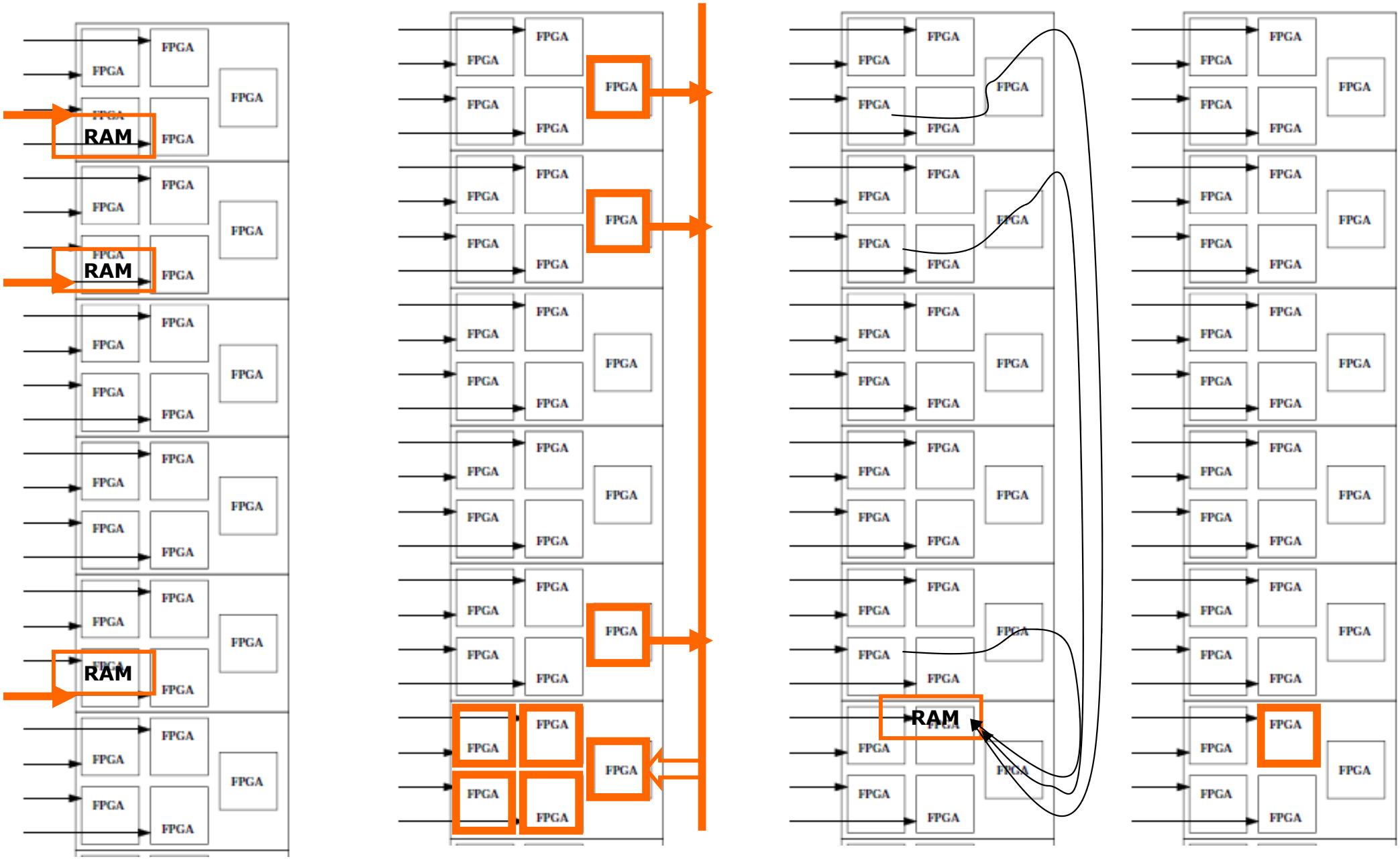
- All code in VHDL on Virtex-4
  directly accessed via PLB (FPGA peripheral bus)
  optical links, RAM, GB ethernet etc.
  (no intermediate step)
- there are RocketIO FPGA-FPGA links
  „full mesh“
  (ATCA backplane)
  > PXD subevent building
- input from SVD (80 optical links)
  FPGA algorithm SVD tracklet finding
  > stand-alone ROI selection (even w/o HLT)
- „centralized“ scheme
  there is a master FPGA
  a.) receives HLT decision and broadcasts in ATCA
  b.) will send BUSY (FIFO full) to Nakao-san

receive data
by optical links

Notify all
bookkeeper FPGAs

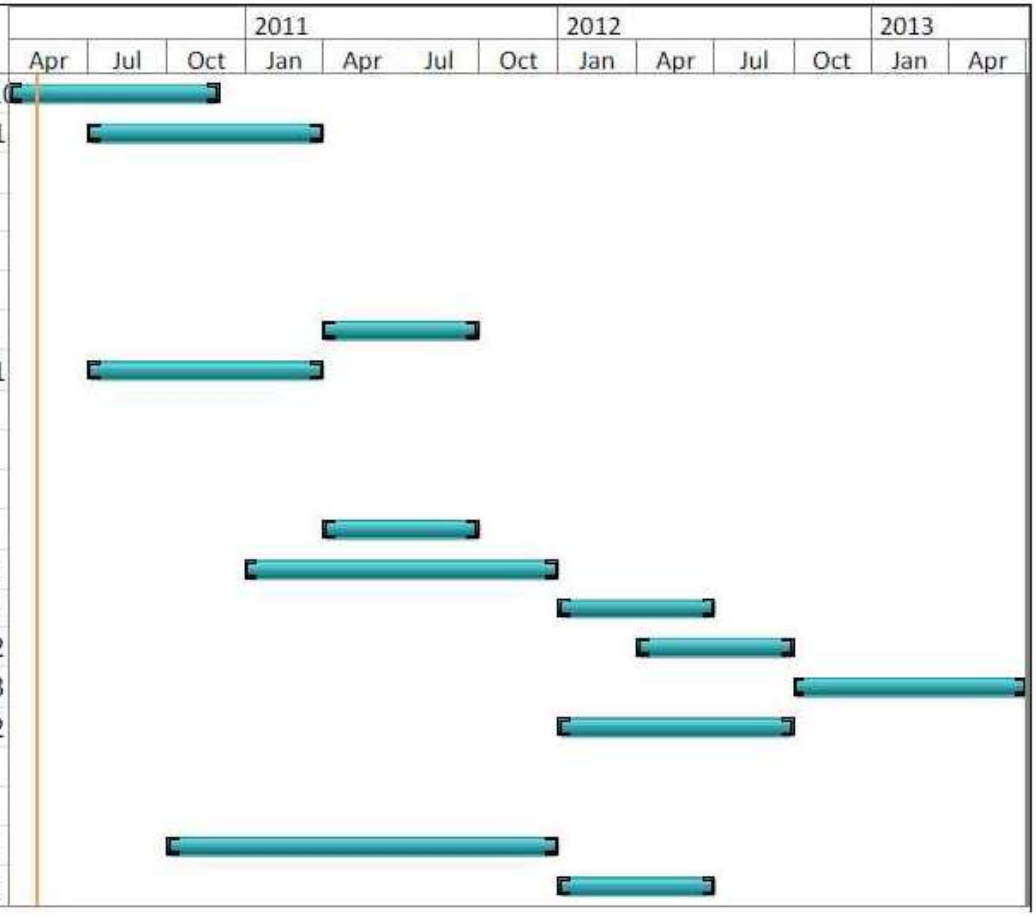
Move (copy and delete)
via ATCA backplane

wait for HLT



ATCA based System – Project Plan

ID	Task Name	Start	Finish
1	Data Receive on FPGA	Thu 01.04.10	Wed 01.12.1
2	Subevent Builder on FPGA	Thu 01.07.10	Thu 31.03.11
3	1. ATCA backplane communication		
4	2. Master/Slave protocol		
5	3. implement BUSY (RAM full)		
6	4. Bookkeeping for (partially) built events		
7	Subevent Sorting for HLT	Thu 31.03.11	Fri 30.09.11
8	Data Reduction on FPGA	Thu 01.07.10	Thu 31.03.11
9	1. SVD+PXD Track finder		
10	2. Helix Extrapolation for HLT track, ROI		
11	3. Alignment/calibration parameter handling		
12	Interface to HLT farm	Fri 01.04.11	Fri 30.09.11
13	Data Output on FPGA (GB Ethernet), TCP/IP Stack	Fri 31.12.10	Sat 31.12.11
14	Prototype System (Readout of 1 Half-Module)	Sun 01.01.12	Sat 30.06.12
15	PCB Mass Production	Sun 01.04.12	Sun 30.09.12
16	1 Full System (Test at KEK)	Mon 01.10.1	Sun 30.06.13
17	MC Simulations	Sun 01.01.12	Sun 30.09.12
18	1. Efficiency Study for physics benchmark channels		
19	2. Offline Implementation of Data Reduction		
20	New PCB (Carrier-Board, AMC Card w/ Virtex 6)	Fri 01.10.10	Sat 31.12.11
21	Testing of new PCB (Softcore PowerPC)	Sun 01.01.12	Sat 30.06.12



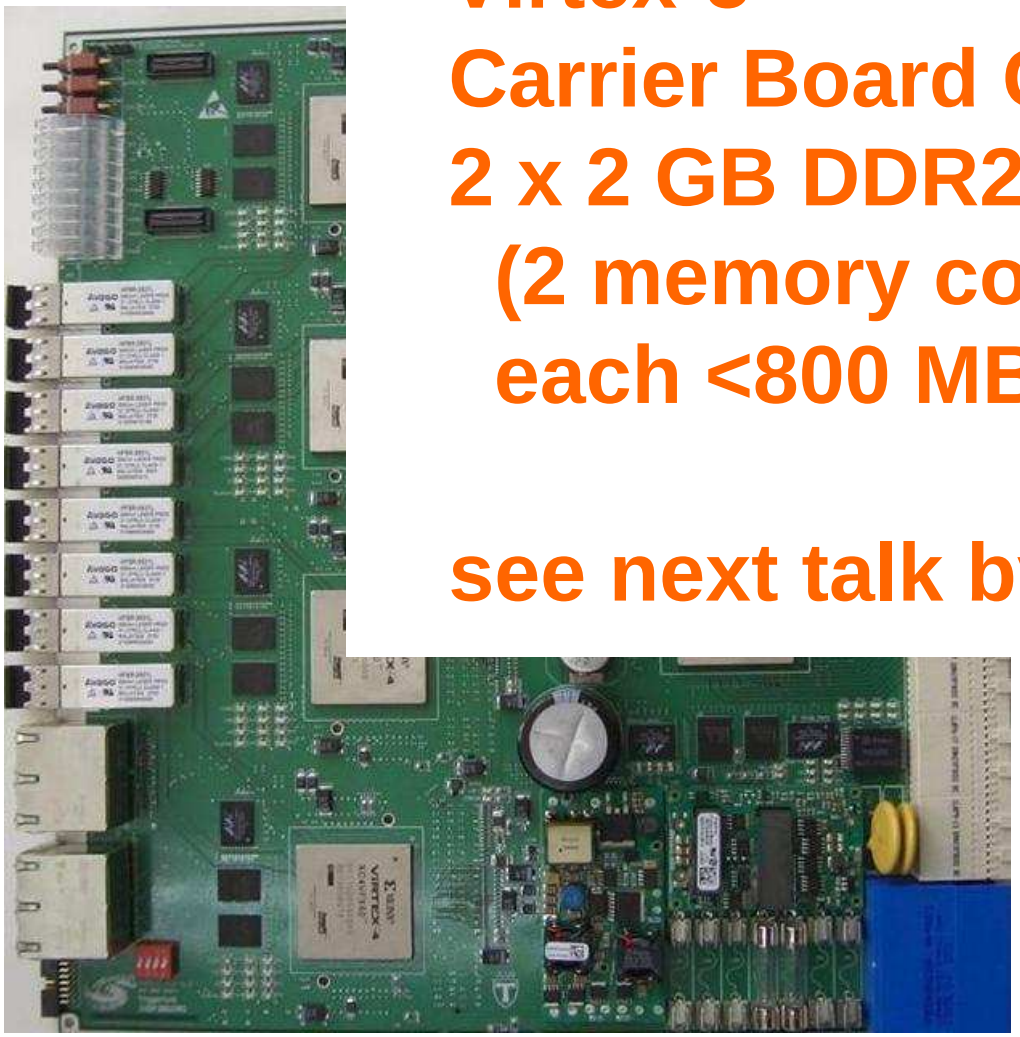
2 Ph. D. students and SL 50% of his time

Memory Issue in ATCA System

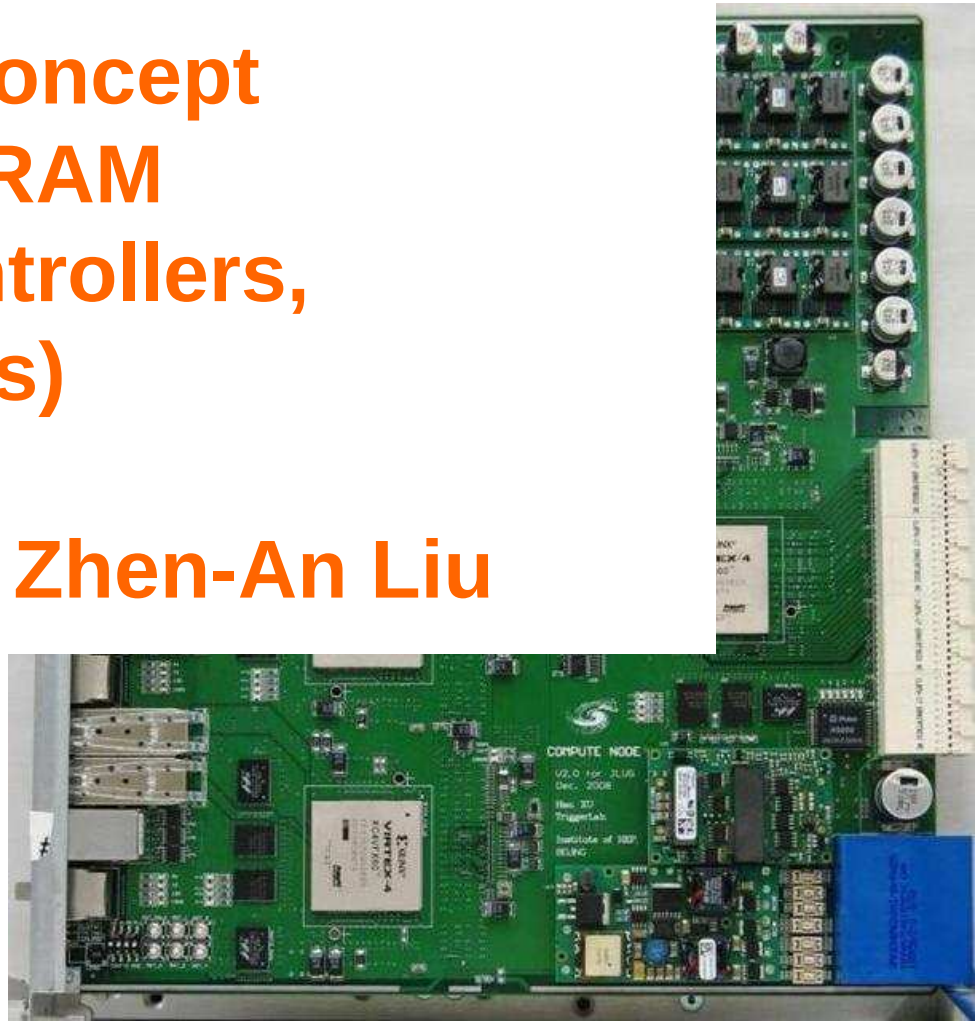
- buffering for 5 seconds until HLT decision required
- in PC based system: add more RAM (e.g. DDR3)
- at 1% occupancy = 180 MB/s per 1 optical link
- in ATCA based system:
  1 optical link = 1 FPGA = 2 GB DDR2 RAM
  so theoretically <11,1 seconds until RAM is full
  but for 3% (incl. background): 3.7 seconds only
- Approaches for improvement:
  - ATCA compute node upgrade project
    see talk by Zhen-An Liu
  - pre-clean-up
    (free memory immediately)
    → 1-pixel cluster
  - Make HLT faster?
    (e.g. can HLT treat some events with priority?, GPU?)

Compute Node Version 3
Virtex-5
Carrier Board Concept
2 x 2 GB DDR2 RAM
(2 memory controllers, each <800 MB/s)

see next talk by Zhen-An Liu



Compute Node Version #1, 2008



Compute Node Version #2, 2009

SVD Optical Concentrator

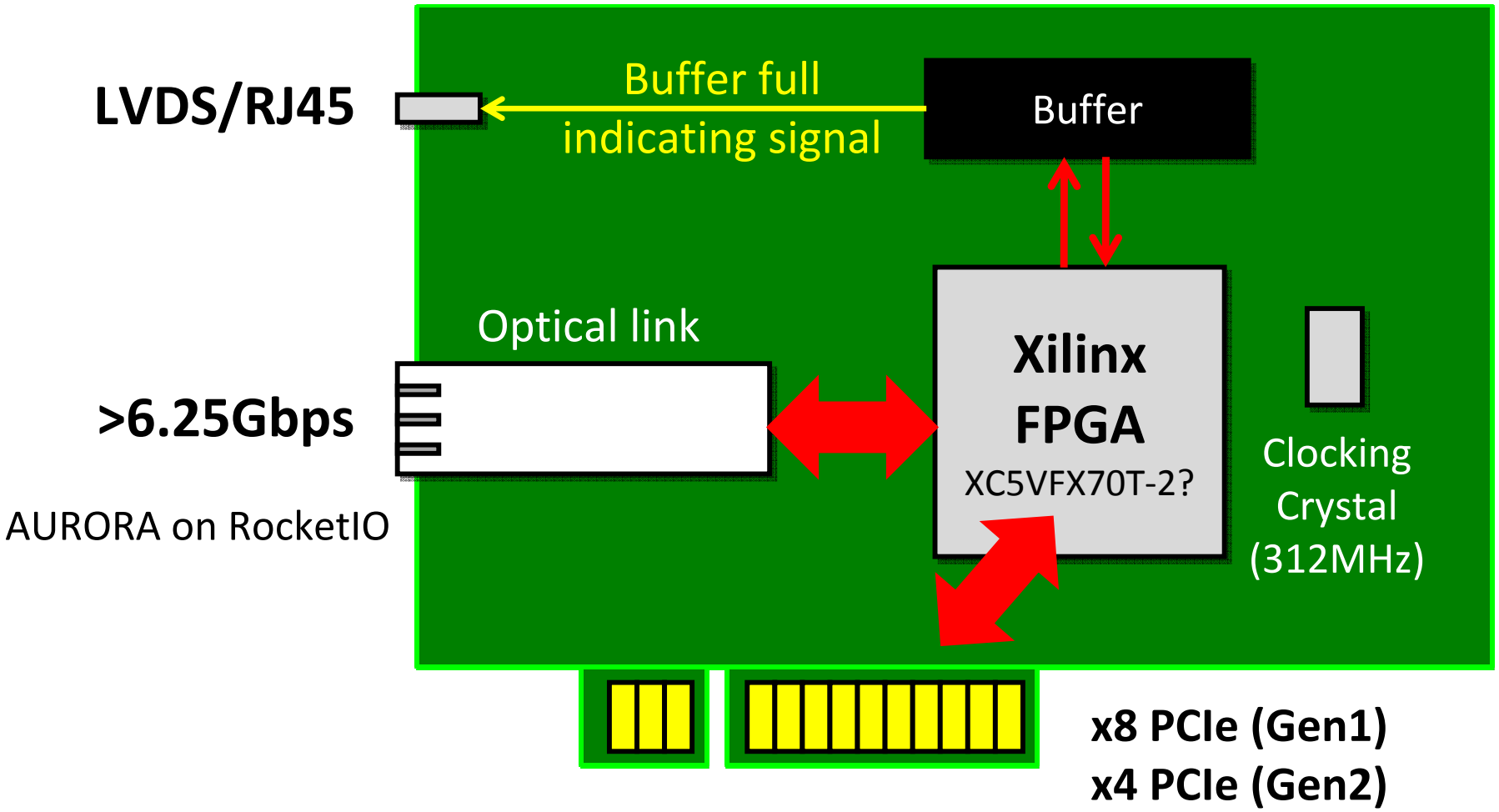
- For SVD bandwidth per optical link even in worst case factor ~9 less than PXD
- 80 links with ~small bandwidth
- project by new Bonn Group (Jochen Dingfelder)
- Plan:
  - 8 × FPGA Virtex-6 VLX240T (3.000,- EUR per FPGA) each FPGA 10 → 1 optical links
  - 2 × 12-Layer PCB (25 cm × 25 cm)

Case	Avgerage occupancy	Data size/ channel (B)	Event size (B)	Total data rate (B/s)	Data rate/ link (B/s)
1	1.9%	12	53.9k	1.54G	19.7M
2	1.9%	4	18.0k	527M	6.6M
3	0.4%	12	10.8k	316M	3.9M
4	0.4%	4	3.6k	105M	1.3M

PC based system

(backup option)

PCIe Card


LVDS/RJ45
Buffer full
indicating signal
Buffer
Optical link
>6.25Gbps
AURORA on RocketIO
Xilinx
FPGA
XC5VFX70T-2?
Clocking
Crystal
(312MHz)
x8 PCIe (Gen1)
x4 PCIe (Gen2)

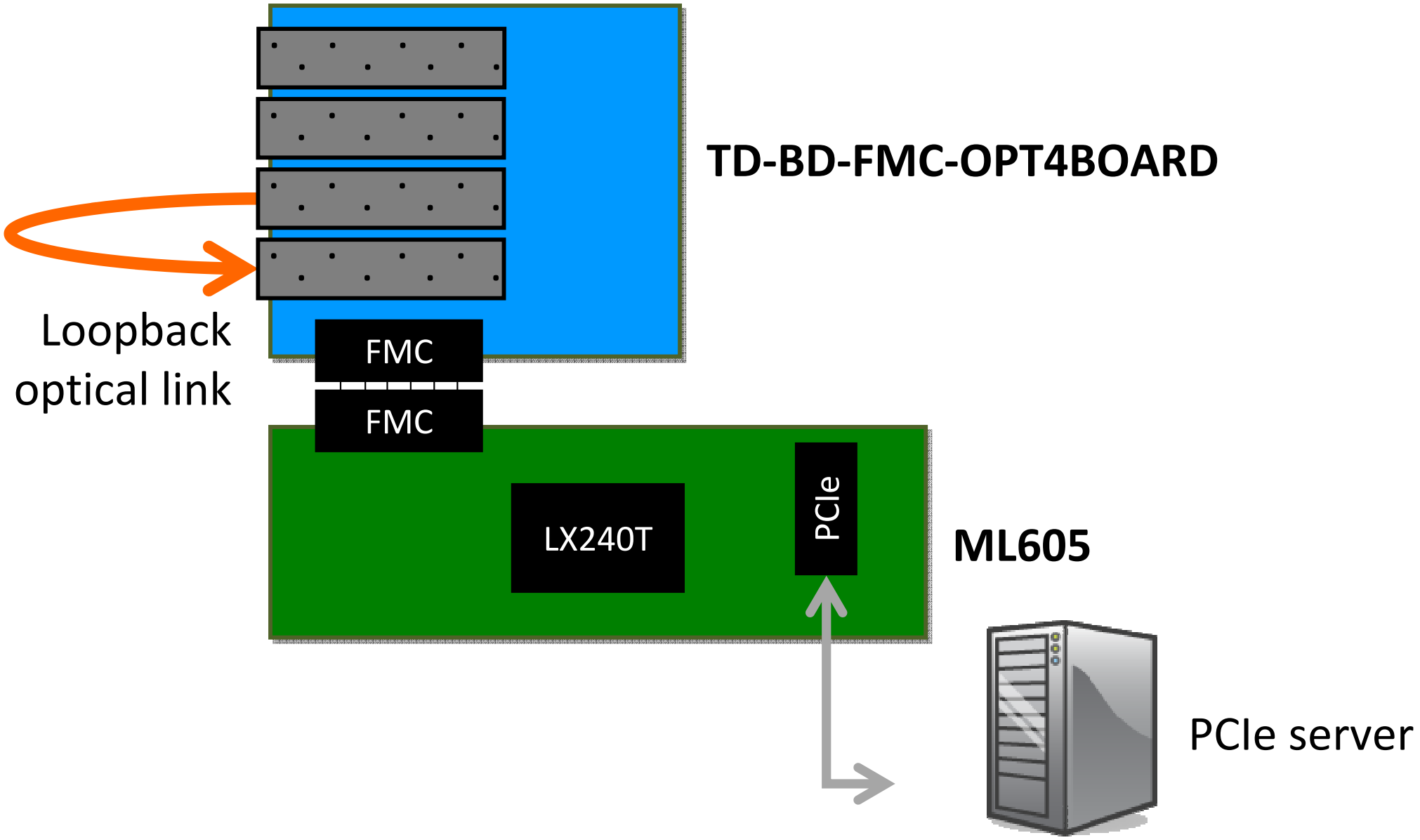
PC based system

- is not a pure PC, but has PCIe card with a FPGA
- a Linux driver has to be programmed
  for x86 < PCIe < FPGA < optical link
  → given to a company
- There are no PC-PC links
  → PXD subevent building is not possible
- No SVD input and no SVD tracklet finding
- „federal" scheme (i.e. no master PC)
  a.) HLT decision broadcasted via switch
  b.) scheme for FIFO full
  OR of all PCs?
- Pre-Study System is being set up
  - Virtex-6 XC6VLX240T
  - SFP+ (8 Gbps)
  - PCIe 2.0 x4 (2 GB/s)

  results maybe by end of this year

Schematic Drawing of the Pre-Study

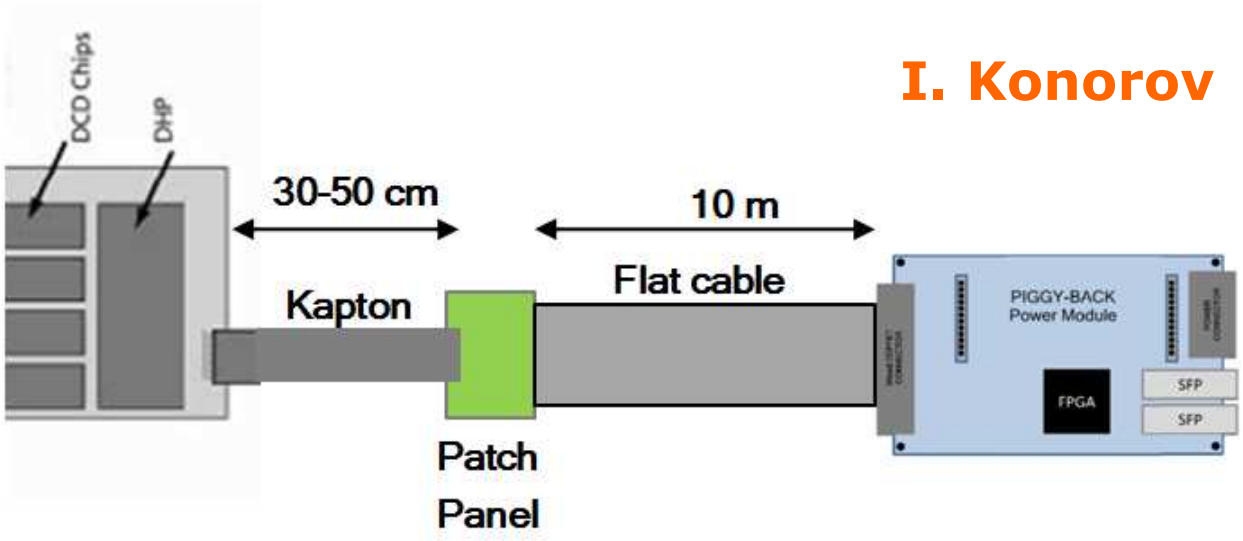
T. Higuchi



DHH

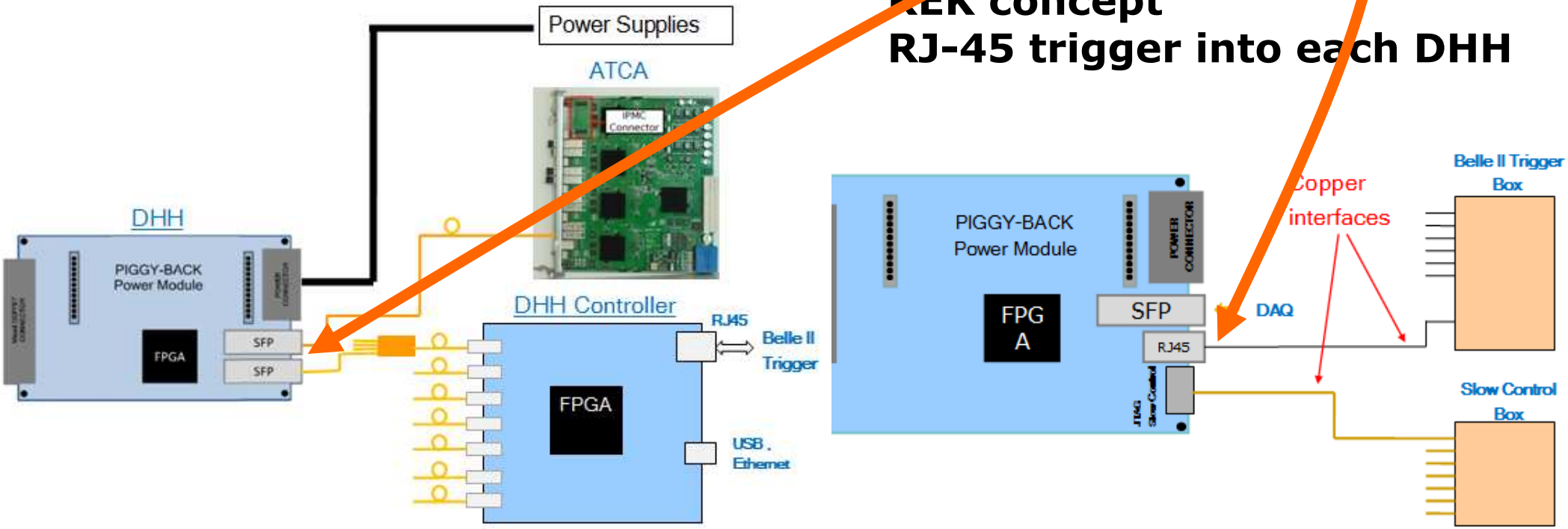
I. Konorov

DHH

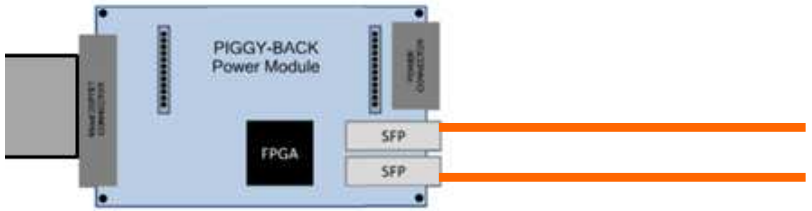


TUM Concept
RJ-45 Trigger into DHH Controller
than optical into each DHH

KEK concept
RJ-45 trigger into each DHH



Optical links from DHH to ATCA/PC



1. 1 or 2 links per halfmodule?
   between DHH and ATCA/PC
2. SFP or SFP+

   Power dissipation is

   3.6V x 240 mA ~1 W

   for both (laser)

   But price is

   $140 vs. $45

<2.5 Gbps

SFP

<8 Gbps
(but limit is
RocketIO
6.5 Gbps)

SFP+

SFP or SFP+ ?

- SFP means <2.125 Gbps
- 1% occupancy =
  1.44 Gbps per 1 optical link
  (if 40 optical links)
- background?
  For TDR we decided 3% max. occupancy
  4.32 Gbps
  needs 2 links, if only SFP
- Even w/o background,
  it contains factor 2.2 reduction
  because of triggered mode
  (50 kHz / 30 kHz, Poission statistics)
  if 1% with untriggered mode
  3.17 Gbps
- Hit paring not taken into account yet (factor 1.3?)

- SFP means <2.125 Gbps
- 1% occupancy =
  1.44 Gbps per 1 optical link
  (if 40 optical links)
- background?
  For TDR we decided 3% max. oc
  4.32 Gbps
  needs 2 links, if only SFP
- Even w/o background,
  it contains factor 2.2 reduction
  because of triggered mode
  (50 kHz / 30 kHz, Poission statistics)
  if 1% with untriggered mode
  3.17 Gbps
- Hit paring not taken into account yet (factor 1.3?)

We decided that we must assume 3%
(we still keep TDR value)

- SFP means <2.125 Gbps
- 1% occupancy =
  1.44 Gbps per 1 optical link
  (if 40 optical links)
- background?
  For TDR we decided 3% max. occupancy
  4.32 Gbps
  needs 2 links, if only SFP
- Even w/o background,
  it contains factor 2.2 red
  because of triggered mode
  (50 kHz / 30 kHz, Poission statis
  if 1% with untriggered mode
  3.17 Gbps
- Hit paring not taken into account yet (factor

here DHH can provide a trigger and reduce factor 2

Timing Distribution

Timing distribution

M. Nakao

- Timing signal RJ-45, cat7 lan cable
  but thin (cheaper, but performance ~15 m acceptable,
  jitter ~30 ps)
  = 4 pairs of LVDS + 1 clock line 127 MHz

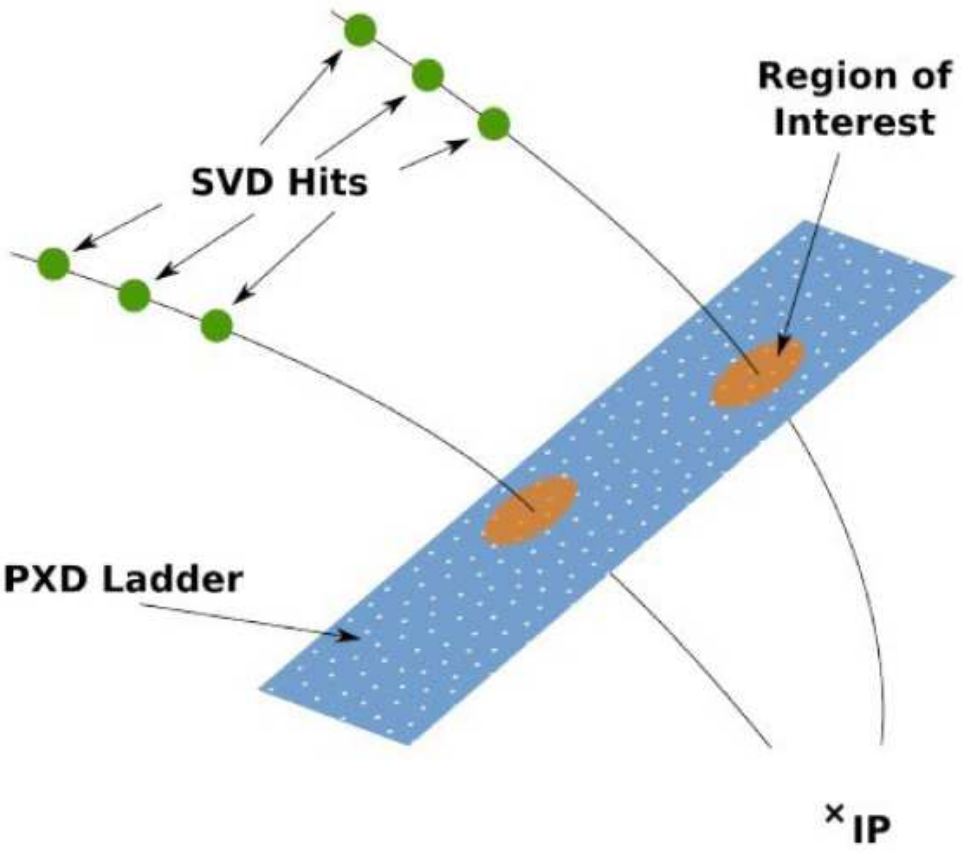
Timing Distribution Board

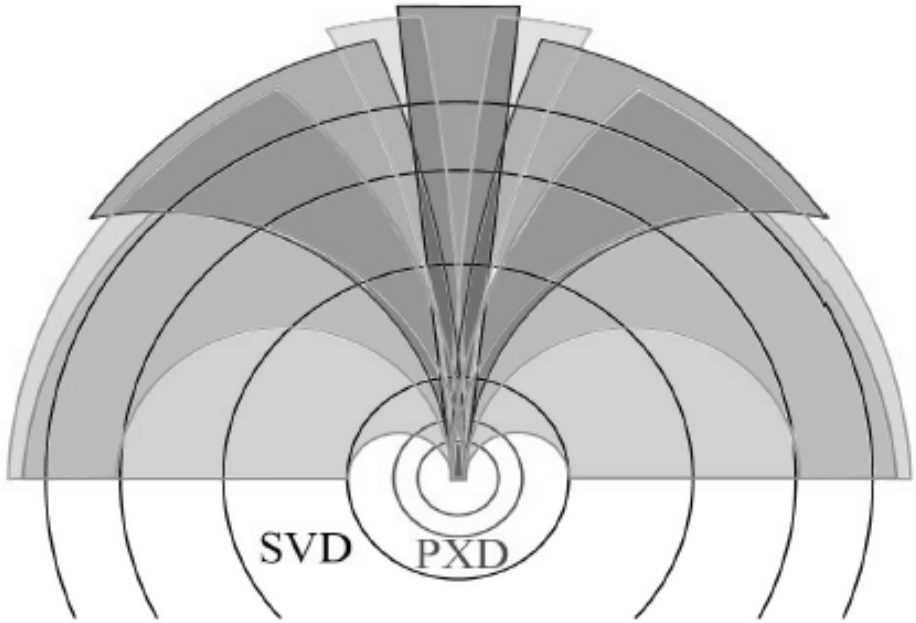
M. Nakao

- FTSW (Frontend Timing Switch Module), VME (6U)
- This board will be connected to DHH
- 1-to-20 LVDS or 1-to-(12 LVDS + 4 optical)
- Virtex-5 FPGA
- first 2 boards arrived at KEK

Algorithms

Claudio Heller Algorithm for ROI Selection





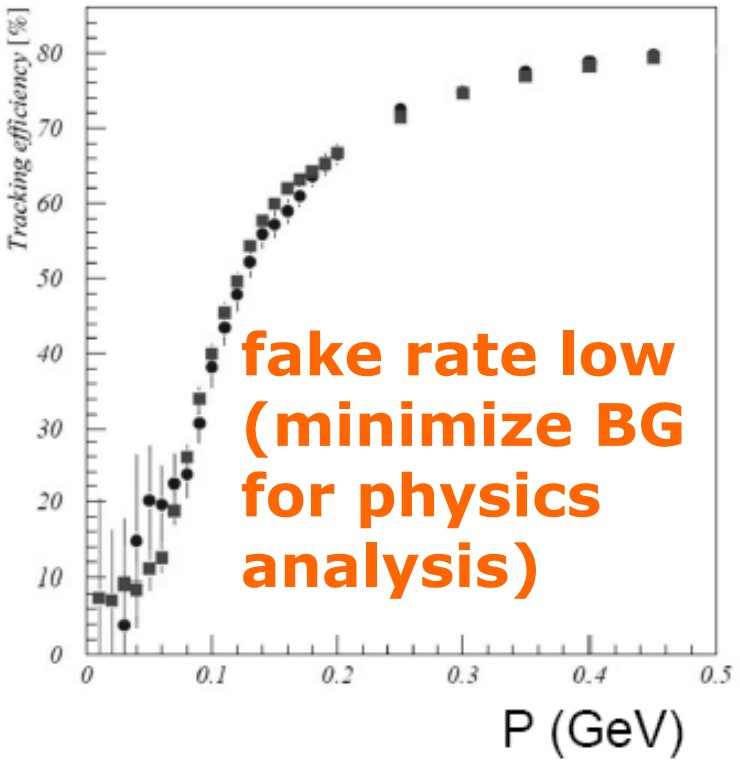
Total # of sectors and „wedges“ for Hough transform
520
(parallisable!)

- 40 straight (for high pT)
- 8 x 60 curved (for low pT)

Track Efficiencies

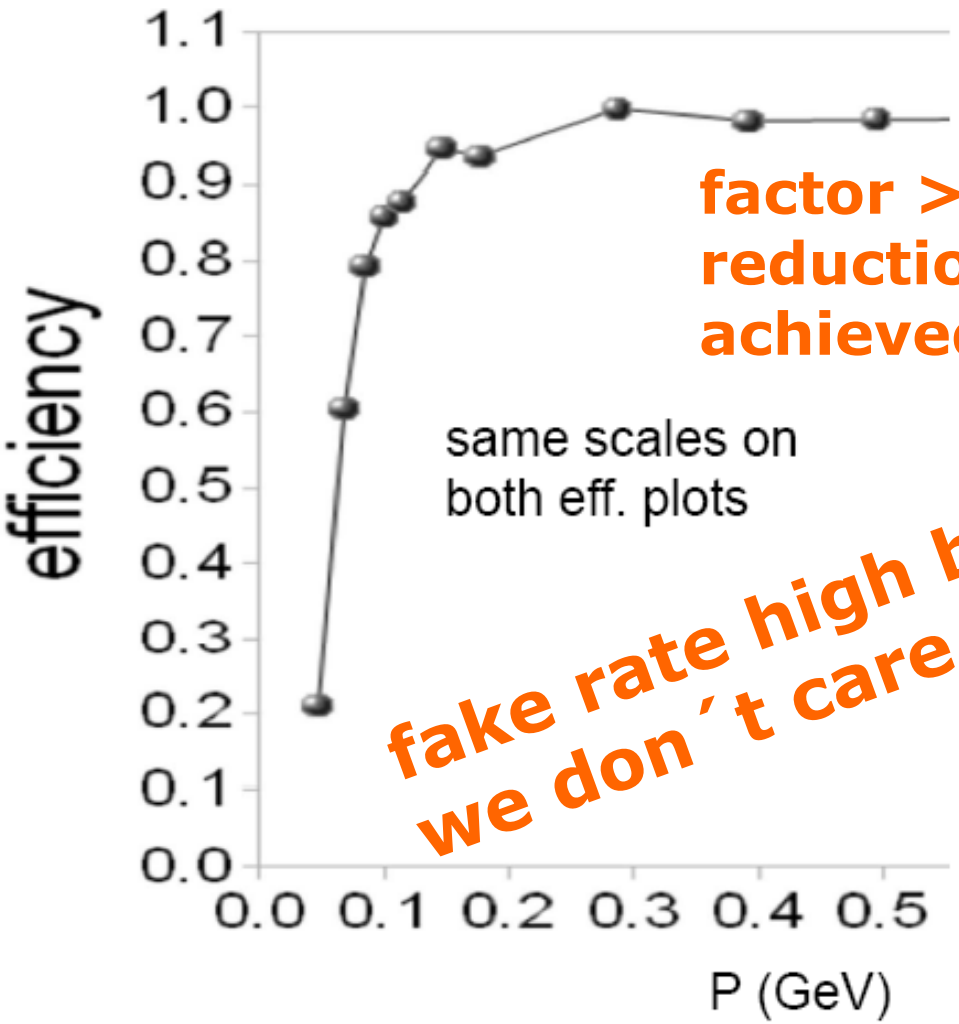
single muons

Full reconstruction, Belle
SVD + CDC



fake rate low (minimize BG for physics analysis)

Hough with SVD, Belle-II



factor >10 reduction achieved

same scales on
both eff. plots

fake rate high but we don´t care

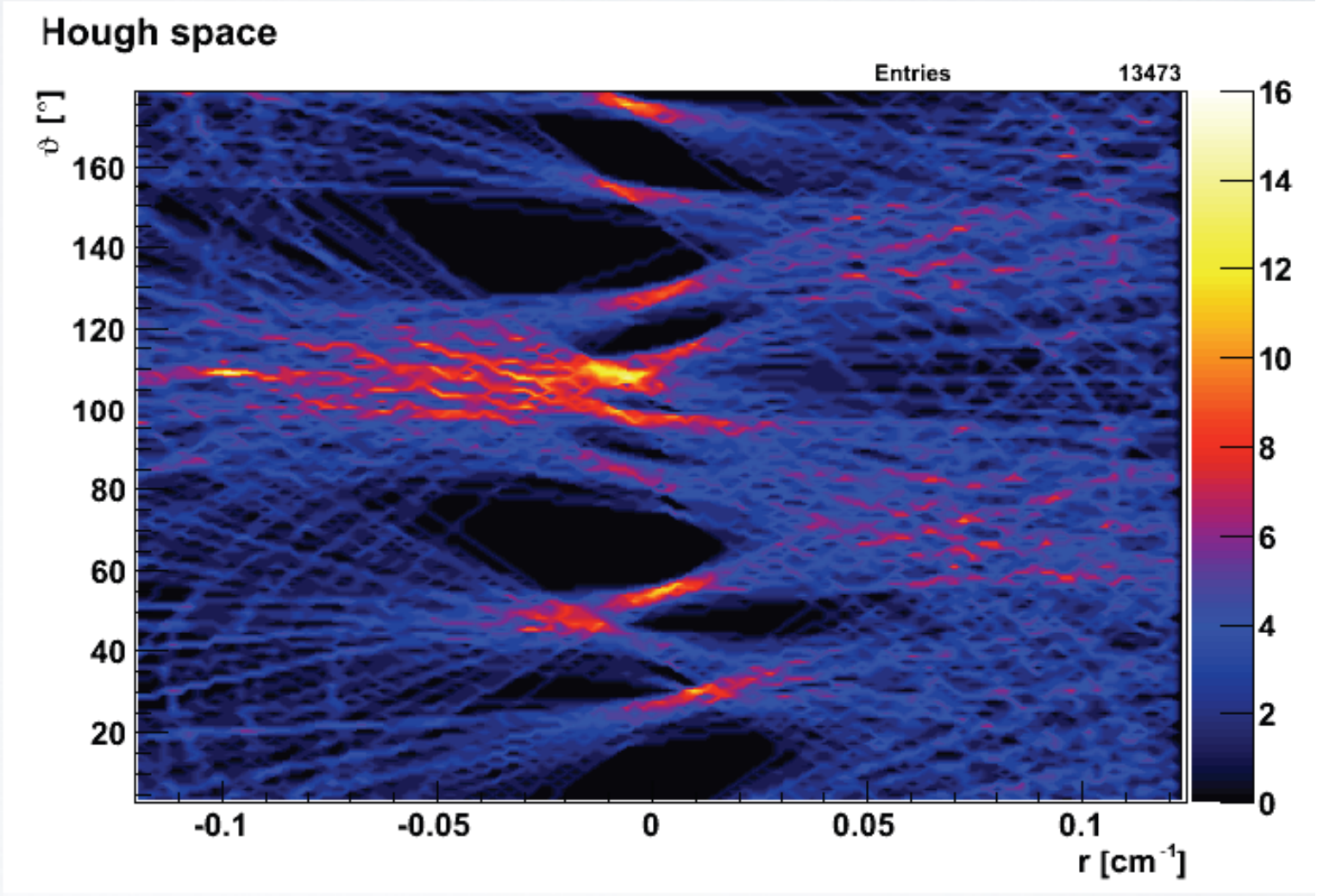
C. Kiesling, 2nd PXD-DAQ-Meeting, Grünberg, Sep 25-26, 2010

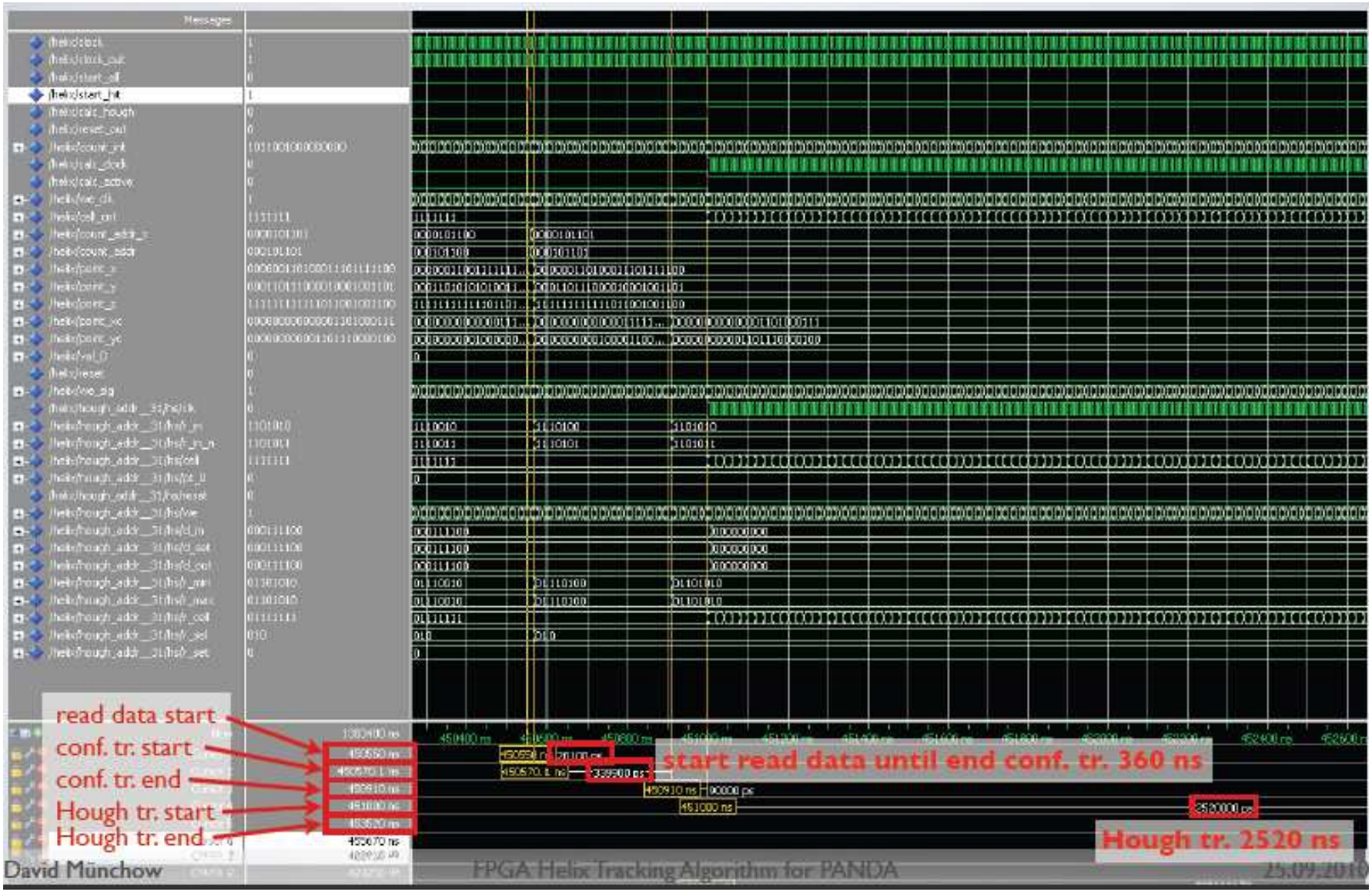
11

Itoh-san proposed to run PXD algorithms even on HLT (w/o CDC, so even low pT)

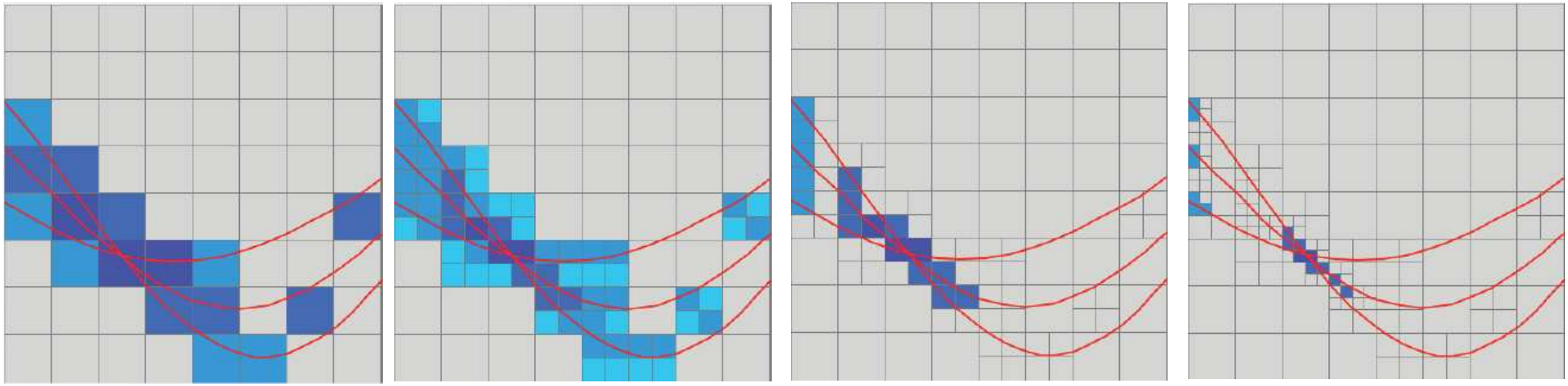
we have to be careful, as GDL trigger bits
for CDC for non-PXD data
imply: pT cut >300 MeV/c

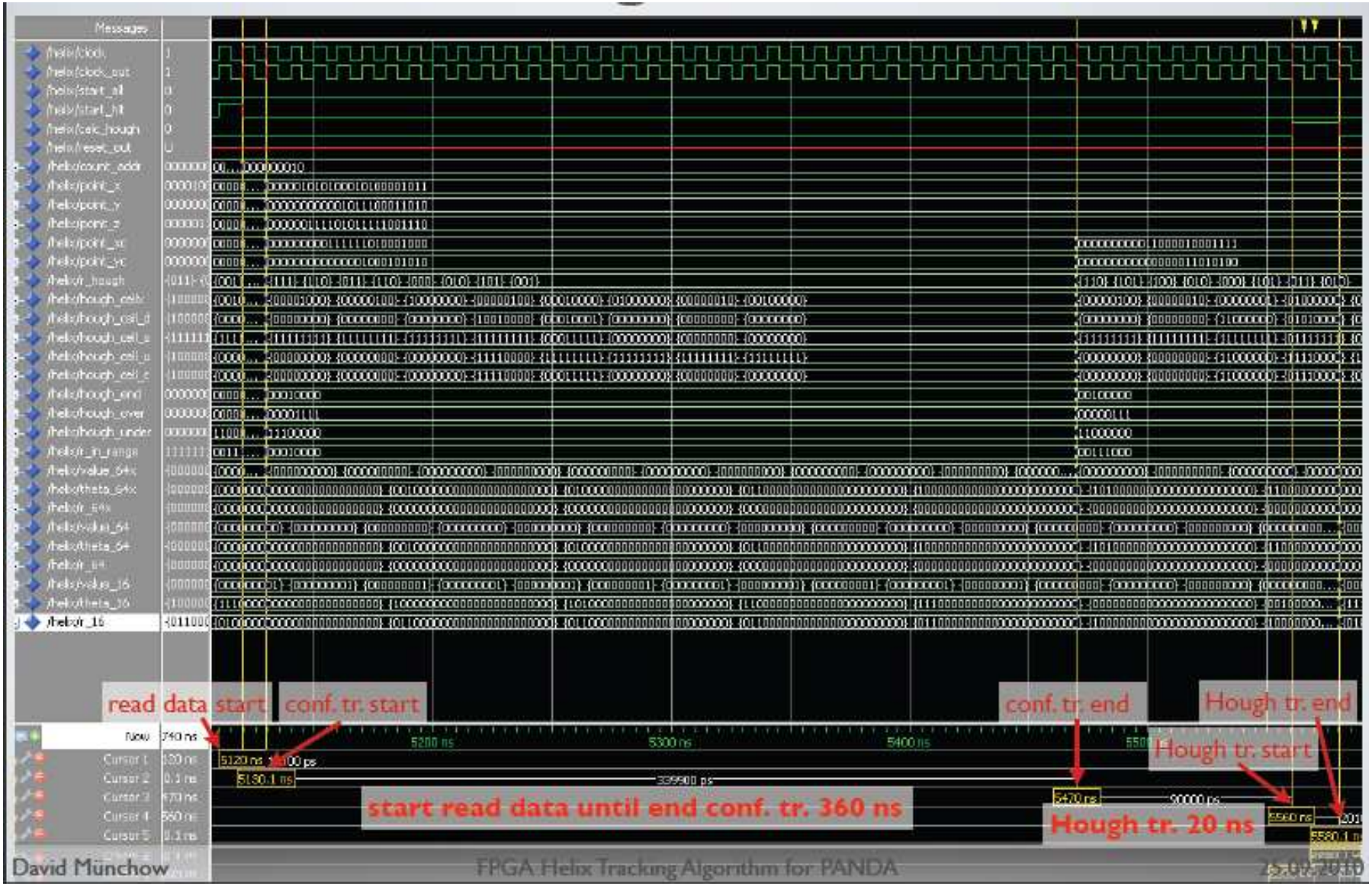
David Münchow (Ph. D. student, Giessen)
Implementation of Panda track finder algorithm conformal map + Hough transform succeeded on Virtex-4 (ML403)




read data start
conf. tr. start
conf. tr. end
Hough tr. start
Hough tr. end
start read data until end conf. tr. 360 ns
Hough tr. 2520 ns
David Münchow
FPGA Helix Tracking Algorithm for PANDA

Fast Hough Transform:
= two tasks at same time:
1. iterative Hough transform („zooming“)
2. peak finder

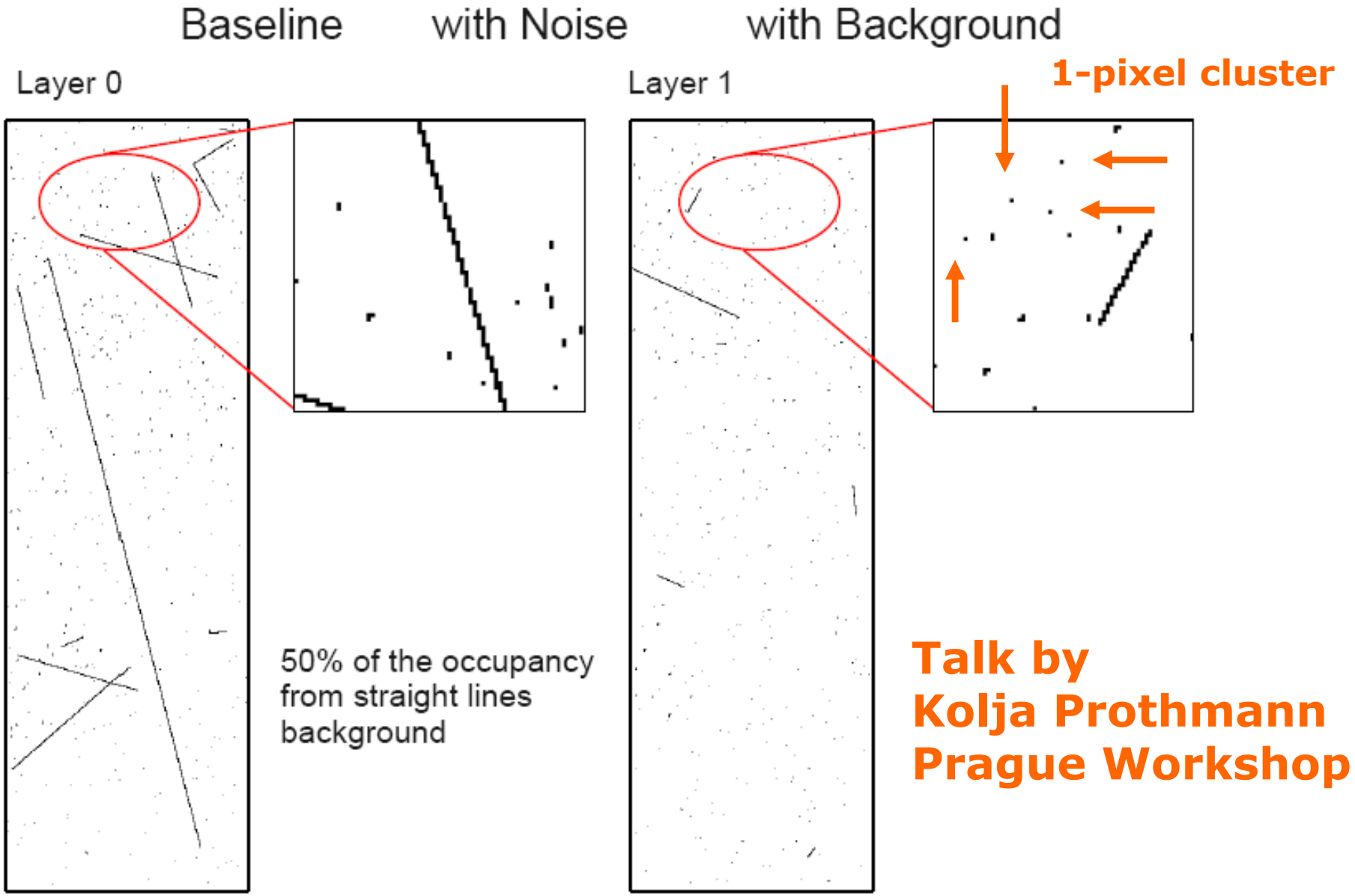



read data
start
conf. tr. start
conf. tr. end
Hough tr. end
Hough tr. start
start read data until end conf. tr. 360 ns
Hough tr. 20 ns
David Münchow
FPGA Helix Tracking Algorithm for PANDA

FPGA Algorithm Timing results

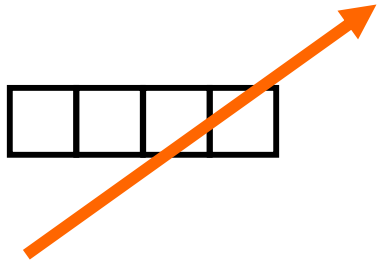
- Read data + conformal transformation
  360 ns per 1 hit
- Hough transform → Fast Hough transform
  2520 ns → 20 ns per 1 hit
  (64 cells parallel)
- Fast Hough transform requires hit sorting
  sorting algorithm was implemented
  unclocked (!)
  sorting is included in 20 ns
- Comparison between PC and FPGA:
  scaled to 800 x 800
  (instead of 128 x 128 for fast Hough, 5 steps)
  2nd z Hough transform
  max. 512 hits, 10 muon mit 2 GeV (~300 hits)
  2.5×10^3
- large factor because of parallelization
  - Hough space in ϑ is parallized (but r serial)
  - per step 64 cells parallel (in fast Hough transform)
  - divider is not parallized yet

1-PIXEL CLUSTER

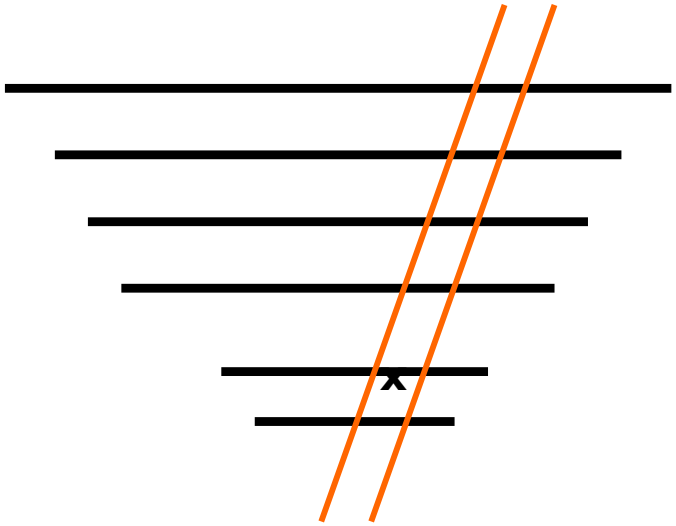


Talk by
Kolja Prothmann
Prague Workshop

- it might be that low energy photons generate only a signal in 1 pixel
- this could help pre-cleanup on FPGA (free RAM immediately)



Charged particles
under polar angles
Must generate 2-pixel clusters

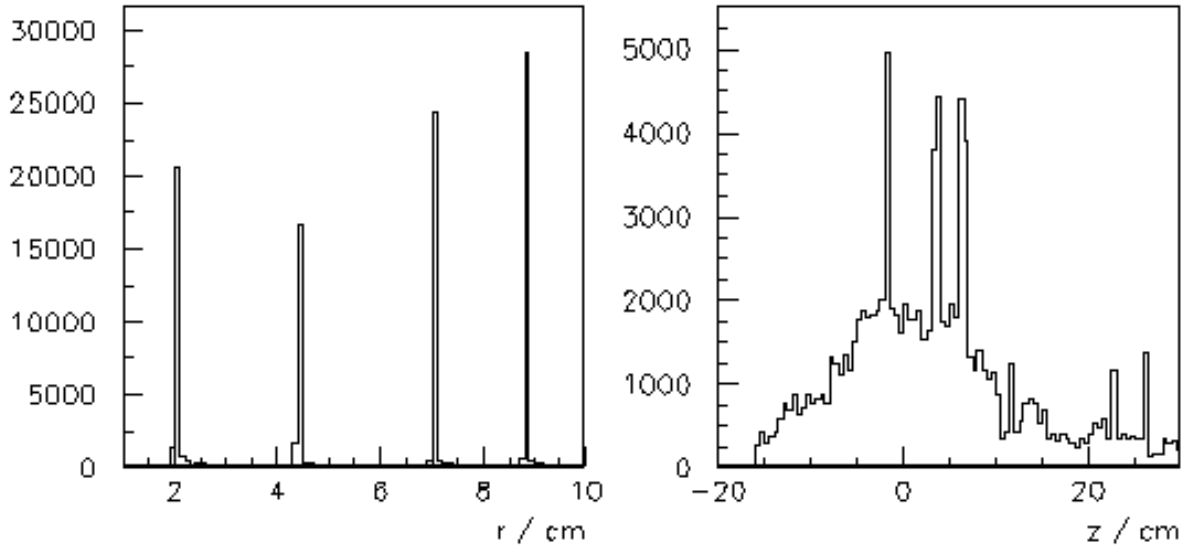


SVD 1-strip hits from real Belle-I data

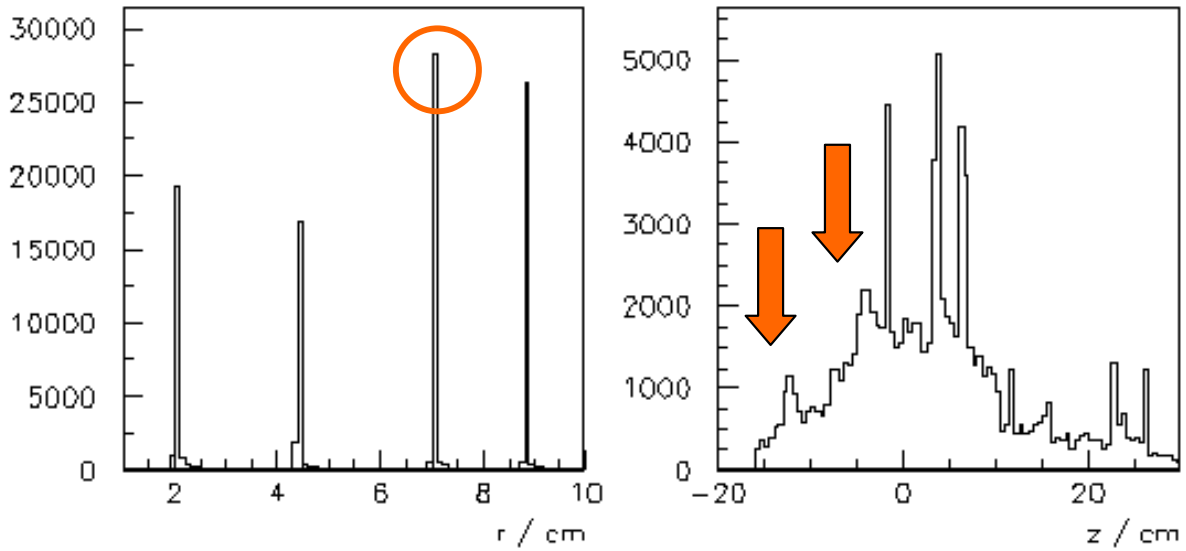
- Processing:
  - Needs SVD hits (pre-clustering)
    needs stand-alone unpack/decode SVD data
  - full production raw data → DST → MDST
    but switch L4 off
- exp. 73, run 419, 28. May 2010, 21:49-22:01
  PXD and Belle II background study
  - CDC background rate 3.7kHz
  - SVD pin diode 2.25 mrad
    (both factor ~1.5-2.0 higher than all other BG runs)
  L4 removes factor ~20
- exp. 69, run 1203, 17. Jun 2009, 14:20-17:25
  $L = 21.083 \times 10^{33}$
  - highest Belle peak luminosity
  L4 removes ~10%

1-strip hits analysis, preliminary results

Exp. 69
0.38%
1-strip hits



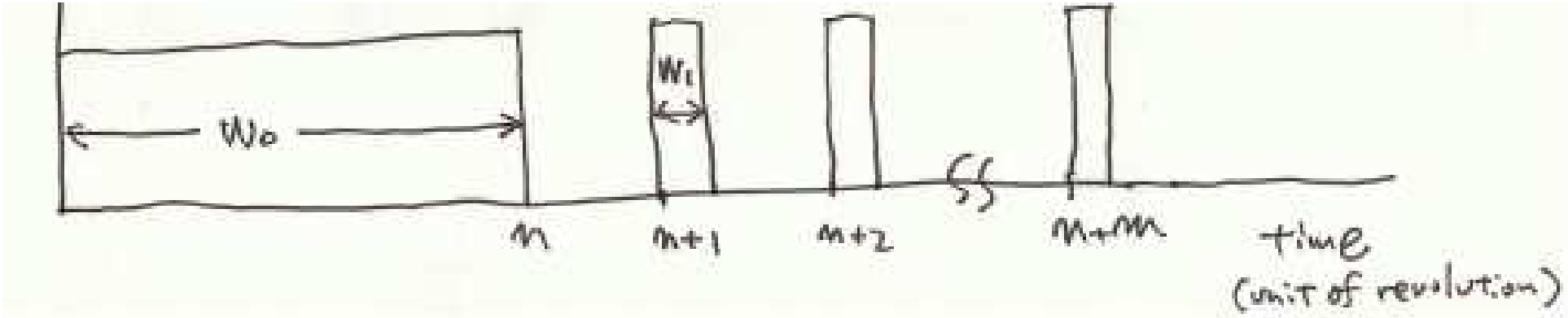
Exp. 73
0.39%
1-strip hits



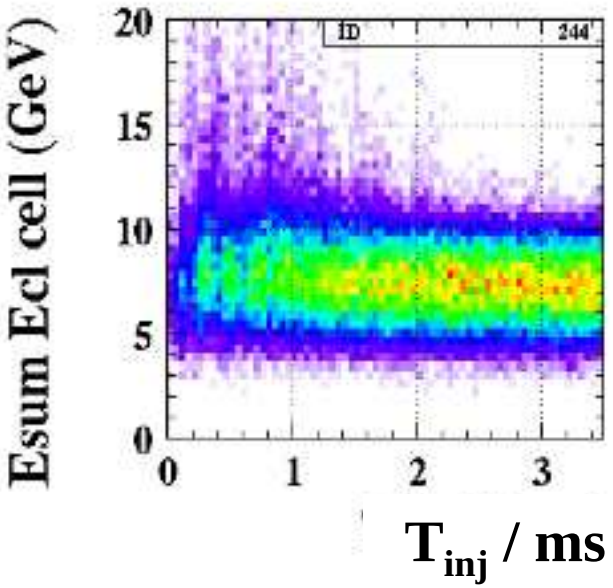
Injection Veto

S. Tanaka, (S. Uno, Y. Iwasaki)

Injection veto

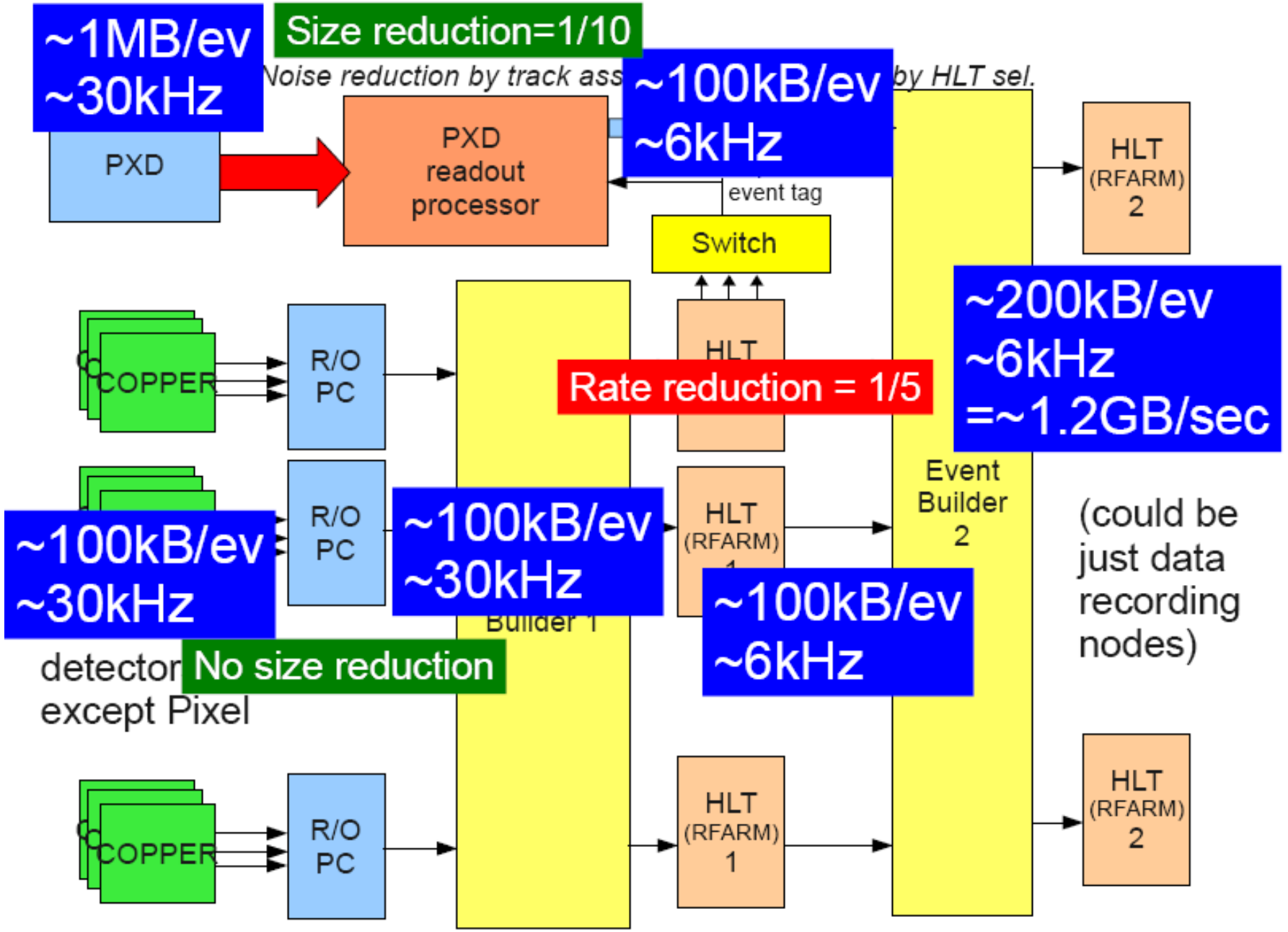


- 2 phases
  - first 10-100 turns w/ 10 usec: veto completely
  - then ~300 turns veto spikes of ~1 usec
  - Veto signal is distributed by trigger (GDL)
- For PXD DAQ it means
  - if PXD uses the veto signal, then: 20% dead time
  - if PXD ignores the veto signal, then: occupancy ~30-40% (?)

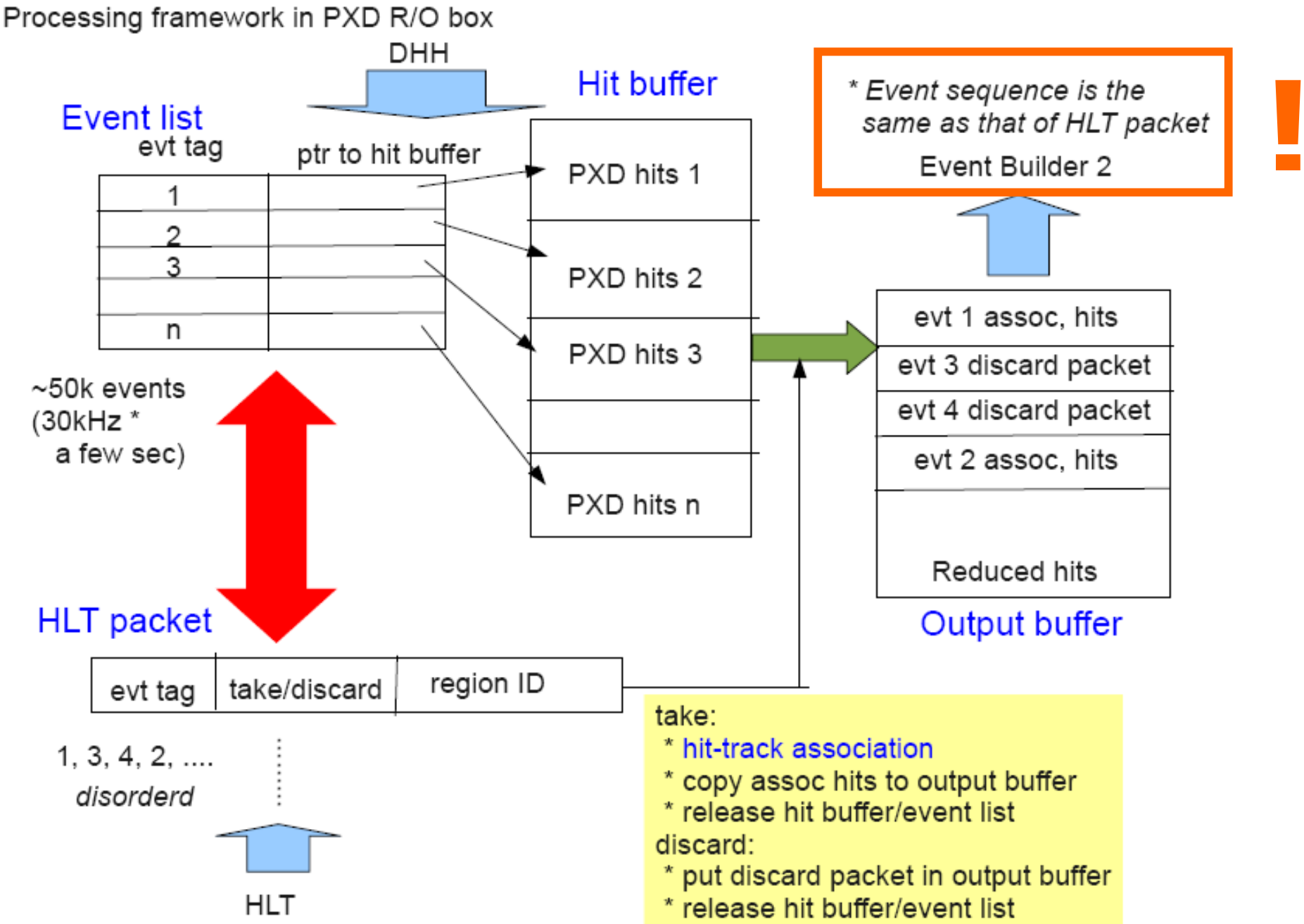


HLT

(High Level Trigger)


~1MB/ev
~30kHz
Size reduction=1/10
Noise reduction by track ass
by HLT sel.
PXD
PXD readout processor
~100kB/ev
~6kHz
event tag
Switch
HLT (RFARM) 2
~200kB/ev
~6kHz
=~1.2GB/sec
COPPER
R/O PC
Rate reduction = 1/5
HLT
~100kB/ev
~30kHz
R/O PC
~100kB/ev
~30kHz
HLT (RFARM) 1
Event Builder 2
(could be just data recording nodes)
~100kB/ev
~6kHz
detector
No size reduction
except Pixel
COPPER
R/O PC
HLT (RFARM) 1
HLT (RFARM) 2

SORTING required !! R. Itoh



Do we really need 4GB/DHH for data buffering?

1. - Average HLT latency for an event is expected to be less than 1 sec as shown in previous meeting.

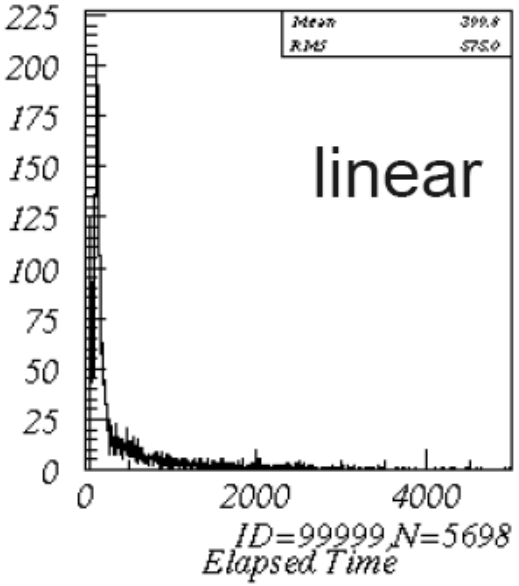
2. - Thanks to parallel processing, the HLT packets for events with shorter processing latency come faster. Not necessary to wait for the 5 sec latency for one particular problematic event.

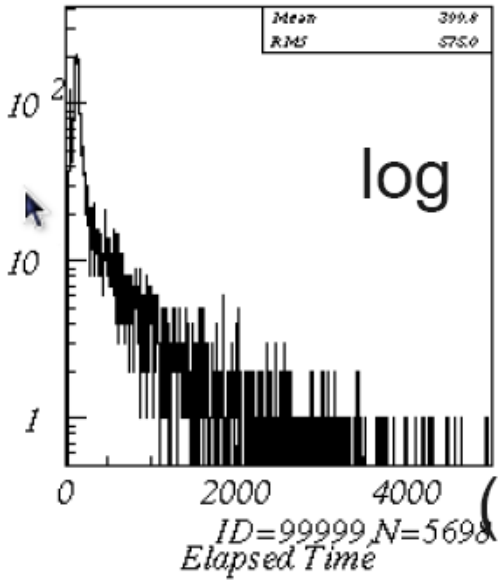
3. - The PXD hit buffer is supposed to be released immediately after the hit-track association (if the algorithm shown in previous slide is used).

We don't have to buffer every event for 5sec.

* It seems the required buffer size is much smaller than 30kHz*5sec*600MB/sec.
  <- This size is necessary only when 5 sec latency is consumed by all of ~2000 cores in HLT at one time. ==> very rare!
* We need a simple MC study to estimate the actual size.
* Sophisticated memory management scheme is required on FPGA in case of Option 12, anyway.

- The processing time for full event reconstruction (incl. both full tracking + energy clustering) is measured for "L4 passed" events. (Exp.57, ~5000 events)





3.2GHz
Intel Xeon,
e-time / core.

(msec)

However, we agreed that we need safety margin.
(4 GB per 1 FPGA)

See the next talk of
Zhen-An Liu
about the Compute Node Upgrade.

Roadmap until the decision ATCA vs. PC

- ~~Funding decision is most important input.~~
- So far, only ATCA prototypes exist and were tested.
  PC based system is preparing a prototype,
  to see if performance meets the requirements.
- Next PXD DAQ meeting:
  when?
  → the week before Ringberg
  (= the week before Golden Week)
  where?
  → Ringberg or some castle
  I already checked and we can get
  this one (Burg Greifenstein, near Gießen)

- From the Grünberg memo:
  „Both systems need to demonstrate the ***baseline option*** from Itoh-san's talk.
  - **One ATCA can receive the ROI from the HLT, and**
  - **ATCA can produce an output based on that.**
  - ***No technical problem, then, take ATCA."***

Please apologize if I missed any important result or statement.

Summary

- Virtex-4,5,6 FPGAs everywhere
  ATCA, PC, DHH, Timing distribution, SVD concentrator board (CDC trigger, …)
- RJ-45 for copper connections
- SFP for optical connections
- System prototypes are progressing.
- First Hough transform algorithms implemented on Virtex-4 FPGA.
  Timing results.
- DHH is taking shape
  (trigger and timing interfaces are being defined)
- discussing interfaces/protocols in both directions
  (sorting, back pressure BUSY signals etc.)
- SVD-only algorithms
  (optimize efficiency, not minimize fake rate)
  maybe even on HLT.