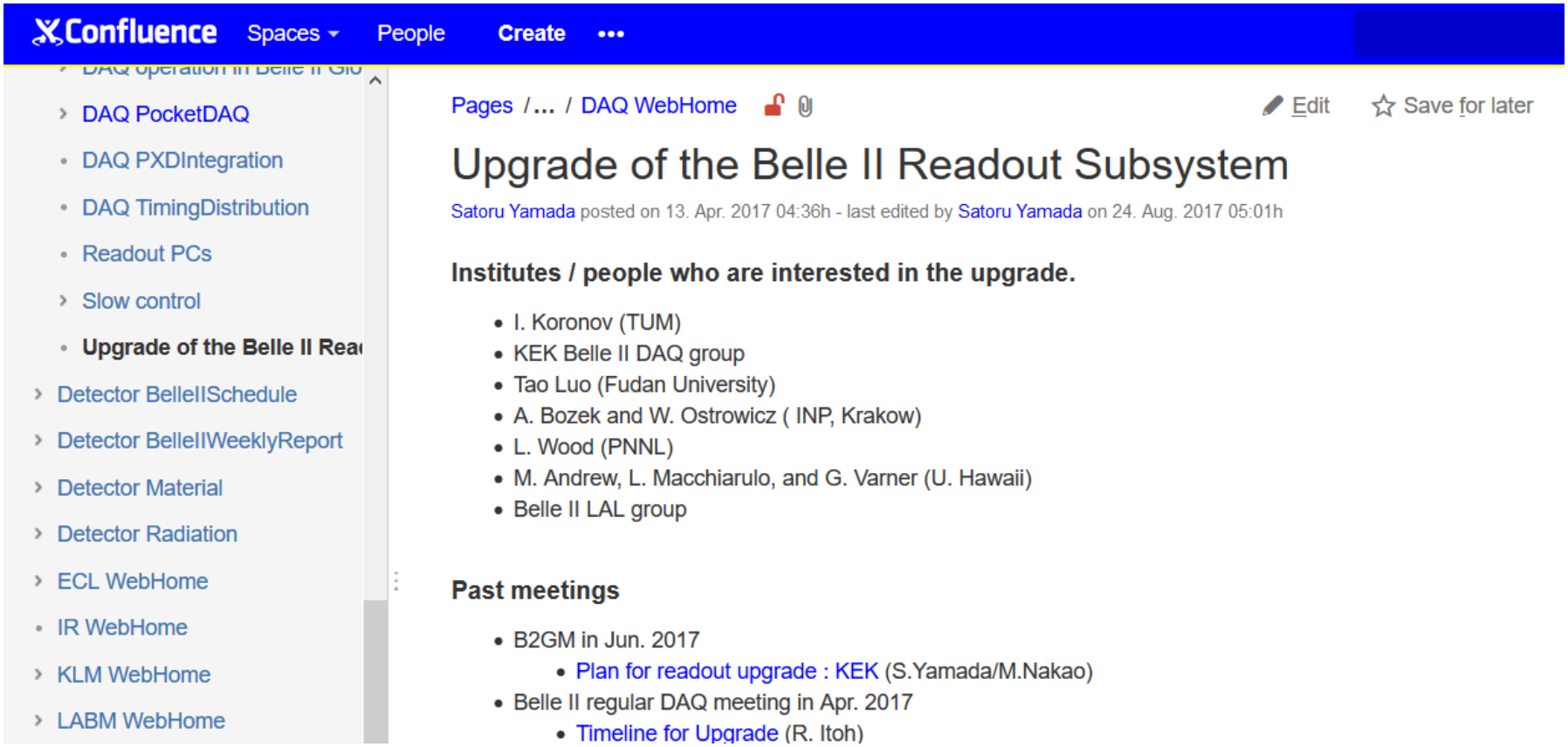
R&D prospect at KEK

S. Yamada

Web site for the upgrade

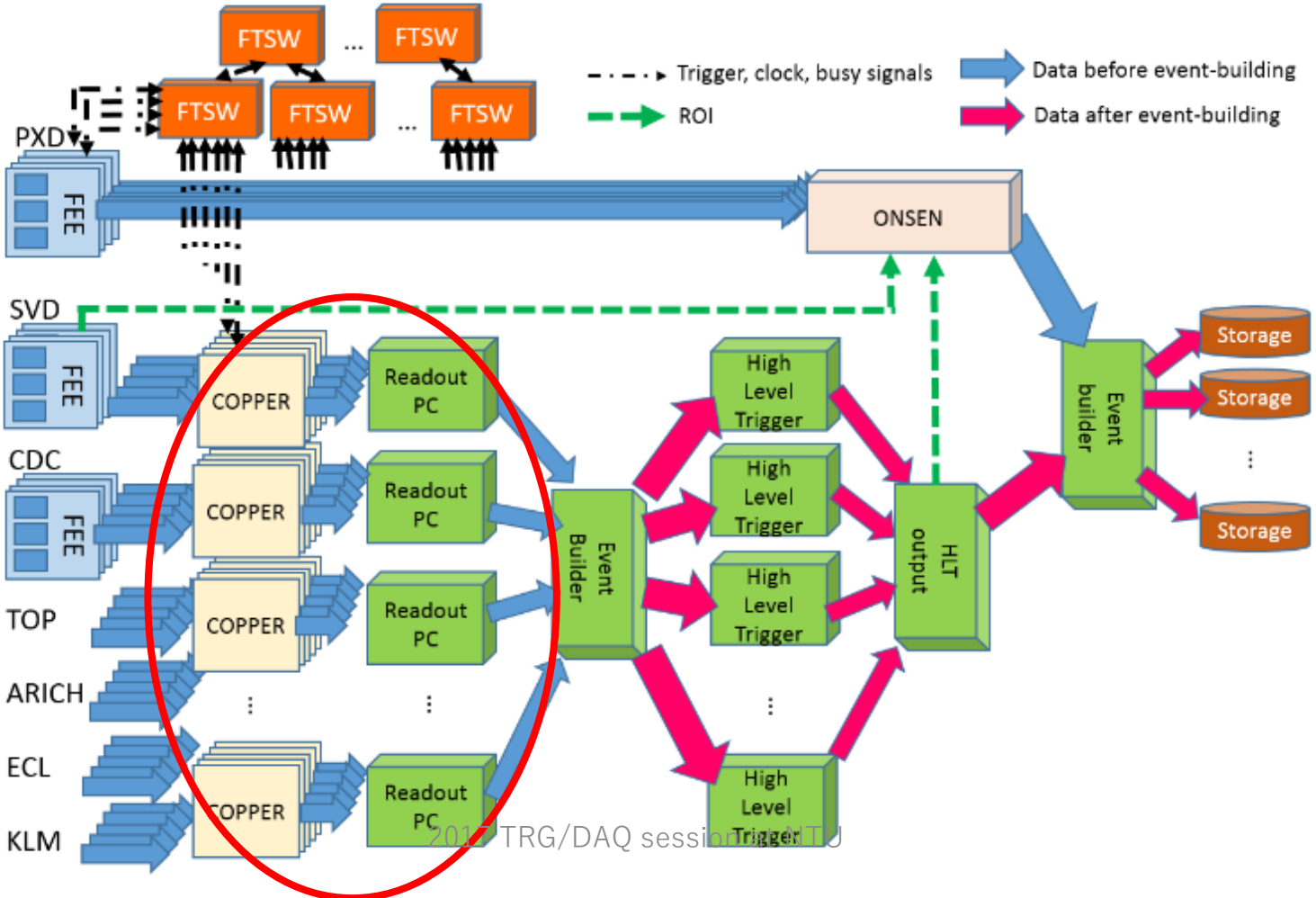
https://confluence.desy.de/display/BI/Upgrade+of+the+Belle+II+Readout+Subsystem



1. Motivation of the upgrade

Role of readout system in the Belle II DAQ system

- Read data via Belle2link(from FEE) and send them over Ethernet (to Readout PCs)
- Event-building of data from 4 FEEs, which correspond to 4 FINESSE slots on a COPPER
- Data formatting (Adding header and trailer)
- Fast control (e.g. send BUSY signal to FTSW when COPPER FIFO is almost full)
- Slow control (Configure FEE though Belle2link)



Issues to be considered for the Belle II DAQ system

Difficulty in maintenance during the entire Belle-II experiment period

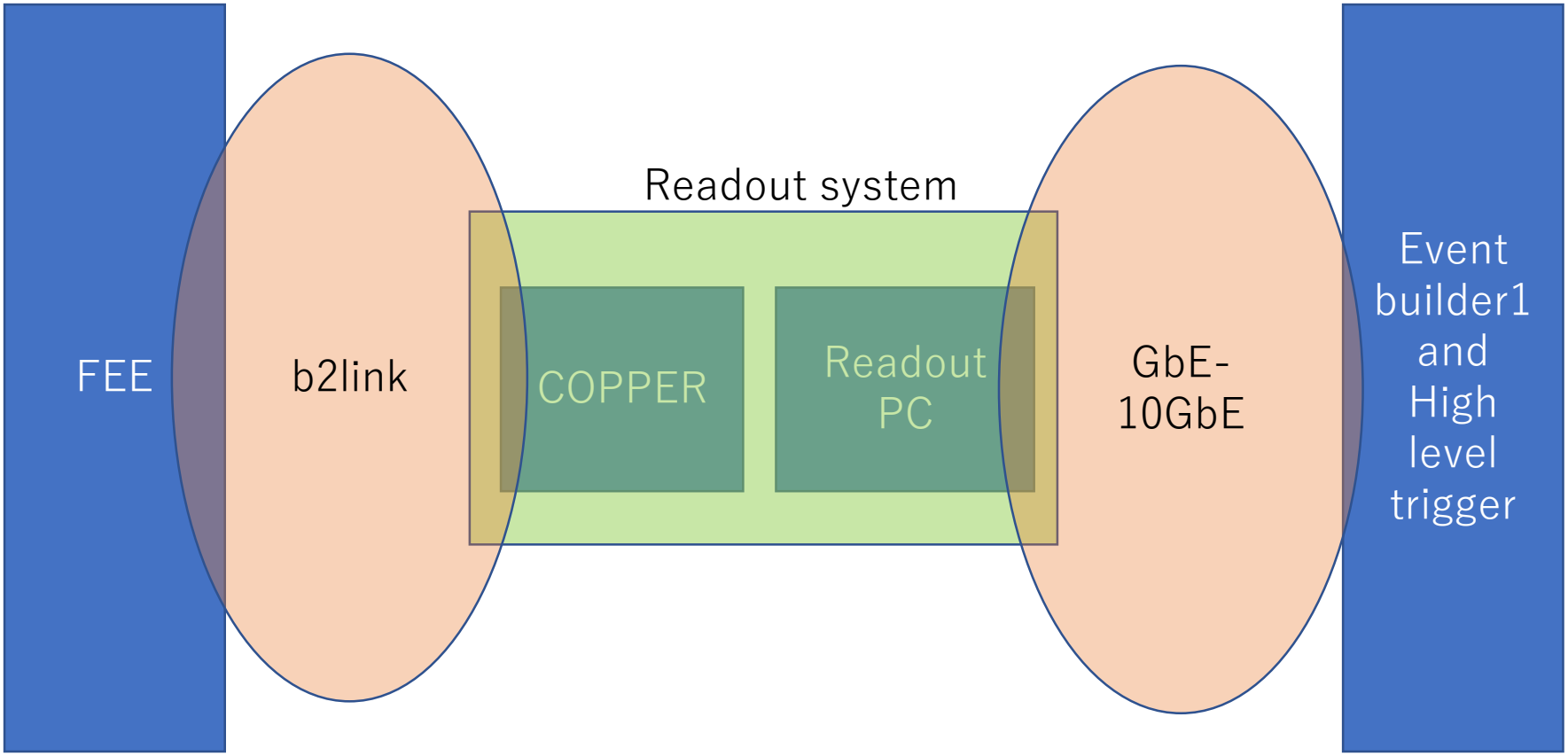
- The number of discontinued parts is increasing.
  - e.g. chipset on a PrPMC card, FIFO and LAN controller on COPPER III
  - For older COPPER II, it is basically difficult to replace parts according to manufacturer.
- Four different types of boards(COPPER, TTRX, PrPMC, HSLB) should be taken care of.

Limitation in the improvement of performance of DAQ

- A. Bottlenecks of the current COPPER readout system
  - CPU usage
    - About 60% COPPER-CPU is used at “30kHz L1 trigger rate with 1kB event size/COPPER”(=Belle II DAQ target value)
  - Data transfer speed
    - 1GbE/COPPER
- B. Bottleneck due to network output of ROPC
- We need to upgrade the readout system when
  - * luminosity of SuperKEKB exceeds expectations.
  - * Lower threshold of L1 trigger is used or trigger-less DAQ is realized.
  - Depending on throughput, network and HLT farms also need to be upgraded.

2. Possible options and firmware development

Boundary condition



Basic framework of belle2link (Rocket-IO based serial link) should be the same. Otherwise FEE's FW/HW update might be needed.

Upgrade like GbE -> 10GbE will be possible, if we upgrade switches.

Throughput

From DAQ Twiki @ 2014 (SVD : 3samples/hit) : (maybe obsolete)

	occupancy	# of link	flow/link	daq ovh	detector buffer total flow [MB/s]	# of inputs/board = 4		# of inputs/board = 10		# of inputs/board = 20		# of inputs/board = 30		# of inputs/board = 40	
			[MB/s]			data flow /boards	# of RO boards	data flow /boards	# of RO boards	data flow /boards	# of RO boards	data flow /boards	# of RO boards	data flow /boards	# of RO boards
SVD	1.7	48	8.9		428	35.7	12	85.6	5	142.7	3	214.0	2	214.0	2
CDC	10	302	0.6		175	2.3	76	5.6	31	10.9	16	15.9	11	21.9	8
TOP	2.5	64	1.5		96	6.0	16	13.7	7	24.0	4	32.0	3	48.0	2
ARICH	1.5	90	1.1		84	3.7	23	9.3	9	16.8	5	28.0	3	28.0	3
ECL	33	52	7.7		360	27.7	13	60.0	6	120.0	3	180.0	2	180.0	2
BKLM	1	24	9.7		60	10.0	6	20.0	3	30.0	2	60.0	1	60.0	1
EKLM	2	36	15.9		42	4.7	9	10.5	4	21.0	2	21.0	2	42.0	1
sum							155		65		35		24		19

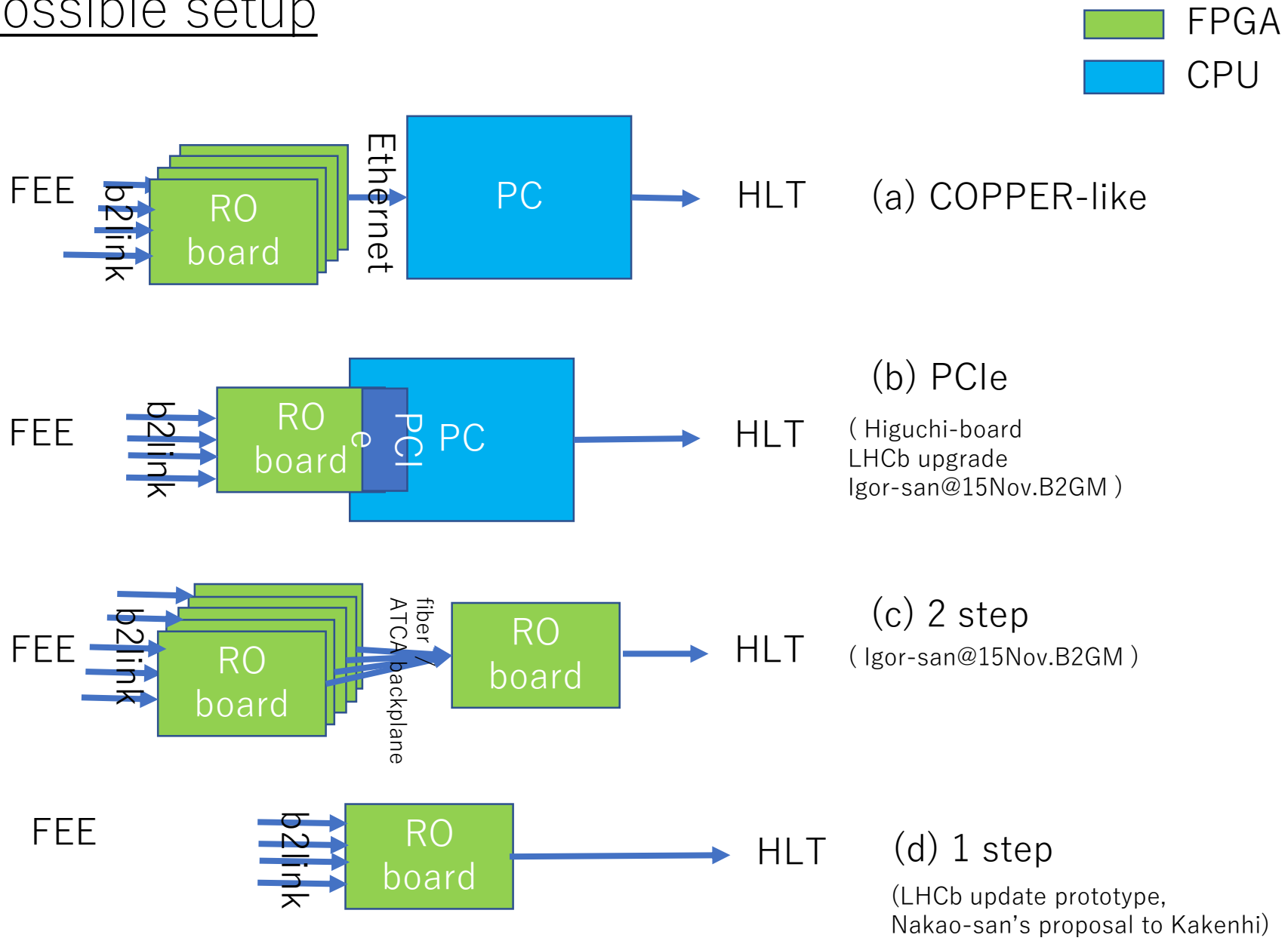
- Data flow per b2link is not so large.

-> if the inputs per board is increased from current 4HSLB/COPPER, we can largely reduce # of RO boards.

-> In that case, some of outputs will become larger than the GbE limit. We need to use 10GbE or reduce # of inputs per RO board for some sub-detectors.

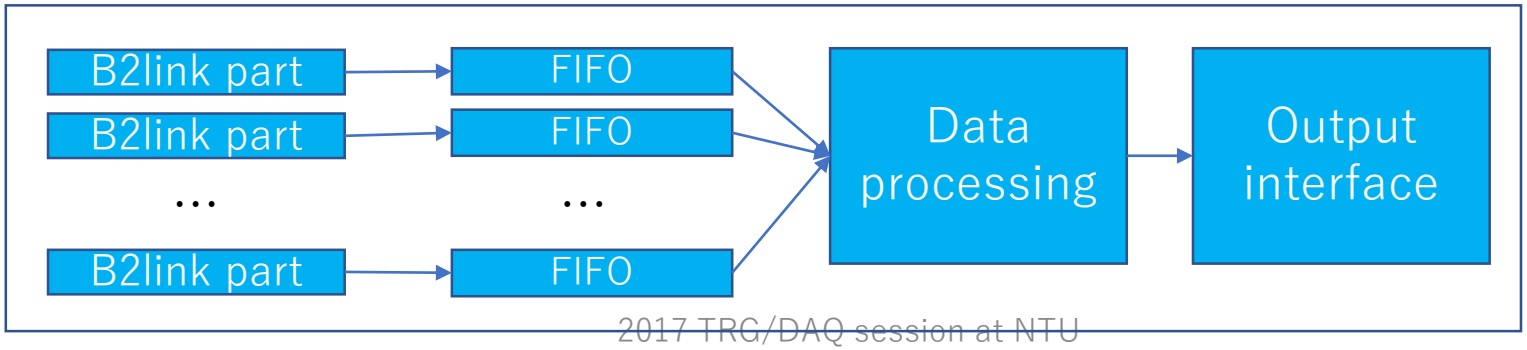
- # of inputs ch affect of the selection of FPGA

Possible setup



Some of Key factors lies in FPGA firmware development

- # of inputs
  - ～10 B2links(GPT) on one board
- Even-building and formatting
- Output protocol
  - Ethernet : 1GbE or 10GbE
- 10GbE output by FPGA((u)ATCA) or PC (PCIe option)
  - FPGA : which IP core will be used ? How to deal with the network congestion
- Long term support for maintenance
  - Board/Firmware development team needs to closely watch the system for years after it starts working in the Belle II.



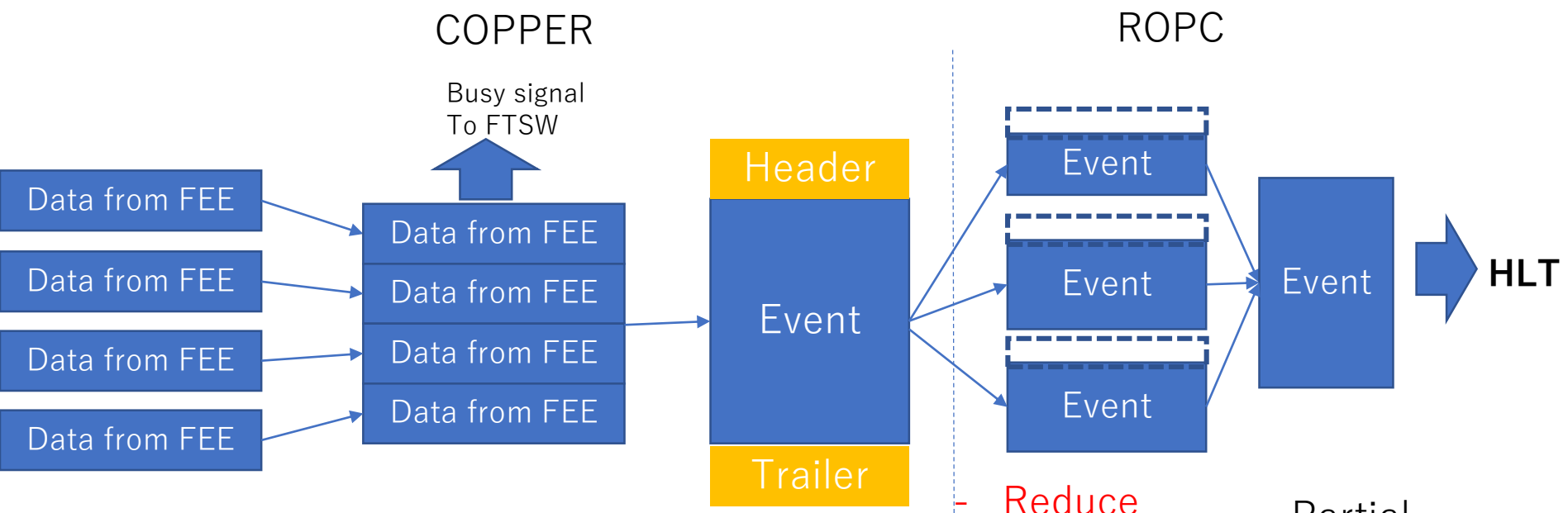
To push things forward, experience of firmware work for more concrete discussion is needed

Available hardware for test/firmware development was discussed at the last B2GM

- Evaluation board (no budget at KEK this fiscal year ?)
- DHPCIe
  - Prof. Kuhn would ask Igor-san to provide one board to KEK.
- Higuchi-board
  - I borrowed one board used at DEPFET project from Konno-san
  - Konno-san is working for providing resource (DEPFET firmware/driver) to DAQ group.

  -> Those boards are not available at KEK now.

Current Data processing by COPPER and readout PC to be covered by new RO board



Receive data from FEE

- Combine data from 4 HSLBs
- Communicate w/ FTSW

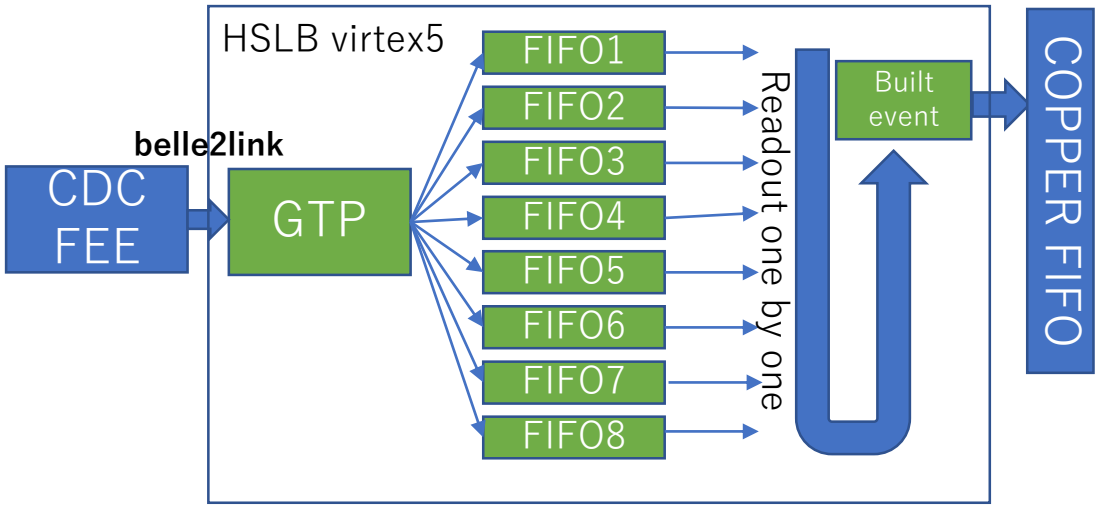
Add header /trailer

- Reduce data-size (remove redundant info.)
- Data check

Partial Event building

- Not so complicated operation, which could be done by firmware.
- But some data-check and error handling needs to be done by software
  - Keep readout PCs or HLT may be able to do those detailed check

Start to play with a HSLB board for firmware development



- Modify the current HSLB firmware
- Copy data from one FEE to 8 FIFOs, build an event from the 8 fragments and send it to COPPER CPU.
- Memory in FPGA is used for the FIFOs.
- Header reduction/attachment and data-check function are not implemented yet.

Event # = 0x200b

00000000 7fff0008 00000000 00000000 00000000 00000000 00000000 00000000 fffffafa 00000dff 00000df8
00000010 00000000 00000000 00000000 ffaa**200b** 49a32737 0000200b 599e1522 00001d00 49a33340 2000002a
...
00000450 00000000 00000000 00000000 00000000 00000000 00000000 00000000 49a32737 200bc2fb ff550000
00000460 ffaa**200b** 49a32737 0000200b 599e1522 00001d00 49a33340 2000002a 4cc30000 0000200b 2200002a
...
00000900 00000000 00000000 00000000 00000000 49a32737 200bc2fb ff550000 ffaa200b 49a32737 0000200b
...
00001350 00000000 49a32737 200bc2fb ff550000 ffaa**200b** 49a32737 0000200b 599e1522 00001d00 49a33340
...
00001800 ff550000 ffaa**200b** 49a32737 0000200b 599e1522 00001d00 49a33340 2000002a 4cc30000 0000200b
...
00002240 00000000 00000000 00000000 00000000 00000000 49a32737 200bc2fb ff550000 ffaa**200b** 49a32737
...
00002690 00000000 00000000 49a32737 200bc2fb ff550000 ffaa**200b** 49a32737 0000200b 599e1522 00001d00
...
00003140 200bc2fb ff550000 ffaa**200b** 49a32737 0000200b 599e1522 00001d00 49a33340 2000002a 4cc30000
...
00003580 00000000 00000000 00000000 00000000 00000000 00000000 49a32737 200bc2fb ff550000 fffff5f5
00003590 7fff0f00 7fff0009 00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000

FIFO1
FIFO2
FIFO3
FIFO4
FIFO5
FIFO6
FIFO7
FIFO8

Hexdump of data read out by COPPER CPU (rate : 10Hz)

Event # = 0x200c

00000000 7fff0008 00000001 00000000 00000000 00000000 00000000 00000000 fffffafa 00000dff 00000df8
00000010 00000000 00000000 00000000 ffaa**200c** 55c44187 0000200c 599e1522 00001d00 55c44d80 2000002a

Itoh-san's previous talk about discussion items :

Example of task sharing (discussion item)

1. Detector interface : Belle2link (and FTSW?)
   * Need to implement HSLB firmware in new readout card
   * Revisit to sender firmware might also be necessary
   * Update in FTSW related firmware together?
   * High-density implementation (>20 optical inputs/board)

2. Porting of COPPER data processing software in FPGA firm
   * Data formatting
   * Event building
   * Data reduction

3. Output implementation
   * Possibly PCI-e interface to be connected to readout PC
   * Option : direct 10GbE output with some ethernet core.
   * Readout PC software is a part of coverage.

4. Hardware development for board mass-production
   * Evalulation of latest FPGA
   * High-density implementation of optical fiber receiver
   * PCI-e interface

My personal opinion/feeling about Itoh-san's slide :

Example of task sharing (discussion item)

1. Detector interface : Belle2link (and FTSW?)
   * Need to implement HSLB firmware in new readout card
   * to sender firmware might also be necessary
   * FTSW re… are together?
   * …ty imp… optical inputs/board)

2. …of COPPE… processing software in FPGA firm
   * Data …ormatting
   * Event building
   * Data reduction

3. Output implementation
   * Possibly PCI-e interface to be connecte…
   * Option : direct 10GbE output with se…
   * Readout PC software is a part of cover…

4. Hardware development for board mass-production
   * Evalulation of latest FPGA
   * High-density implementation of optical fiber receiver
   * PCI-e interface

High density belle2link by IHEP/Fudan people ?

Igor-san already has event – building from memory

Porting our Belle II DAQ program to firmware

Each group will make/bring a proto-type board for possible options and compare cost and performance ?

Summary

- Even though we have not started the Belle II experiment, it is useful to start thinking possible option of future upgrade of Belle II readout system, because
  - It will become difficult to repair of broken COPPER boards
  - We need to handle the unexpected increase of event-rate or event size.
- Hardware spec. is still open.
  - 'Input = belle2link' and 'output = Ethernet or PC server' will be the boundary condition.
- Firmware in the new RO board should do the data-processing currently done by COPPER and readout PC.
- Start playing with available hardware for the firmware development.

Backup

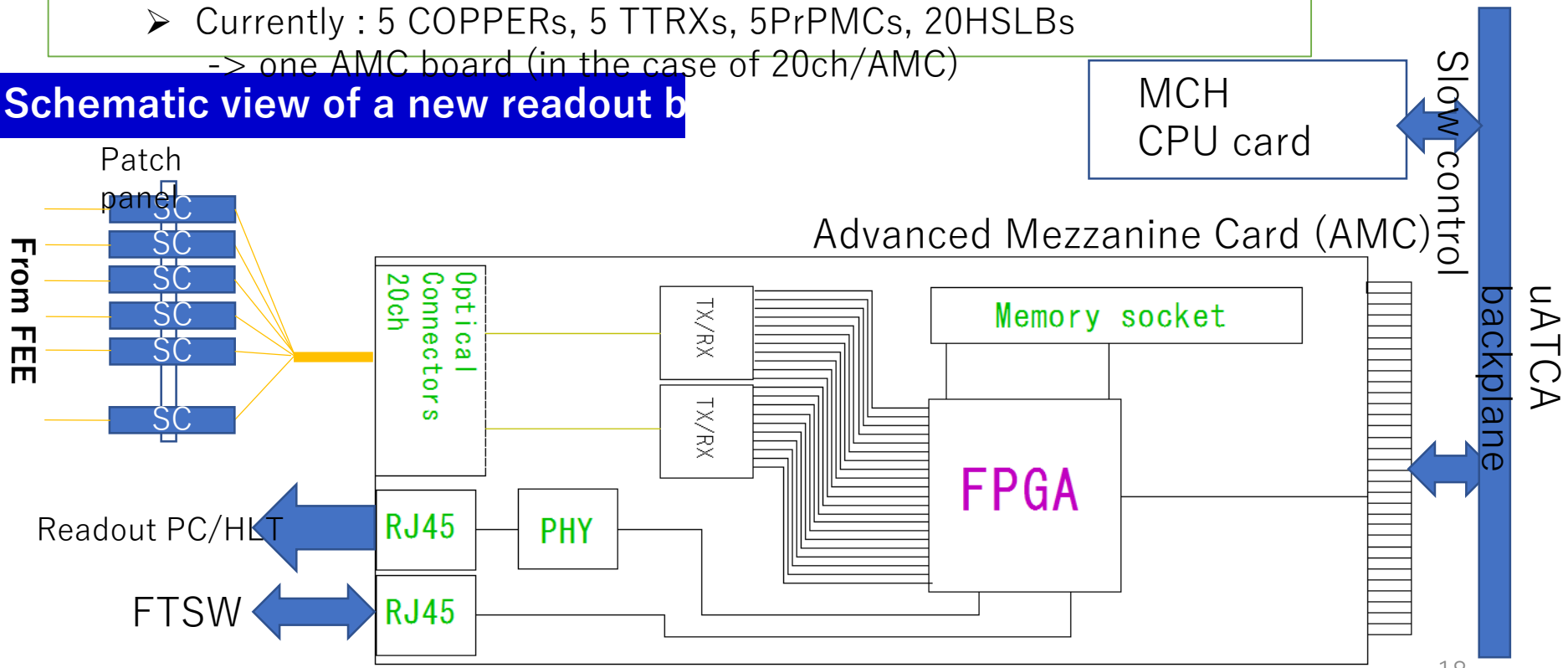
Example of a board sketch

(2015. 8.6 MOF meeting

New readout system = High-density FGPA-based system u

- Data processing speed
  - Fast FPGA-based data processing
- Data transfer speed
  - 10GbE (directory connected to a HLT unit) or 1GbE (keep readout PCs)
- Compact and high-density system
  - high density connector and higher throughput
- Easier maintenance
  - Currently : 5 COPPERs, 5 TTRXs, 5PrPMCs, 20HSLBs -> one AMC board (in the case of 20ch/AMC)

Schematic view of a new readout b



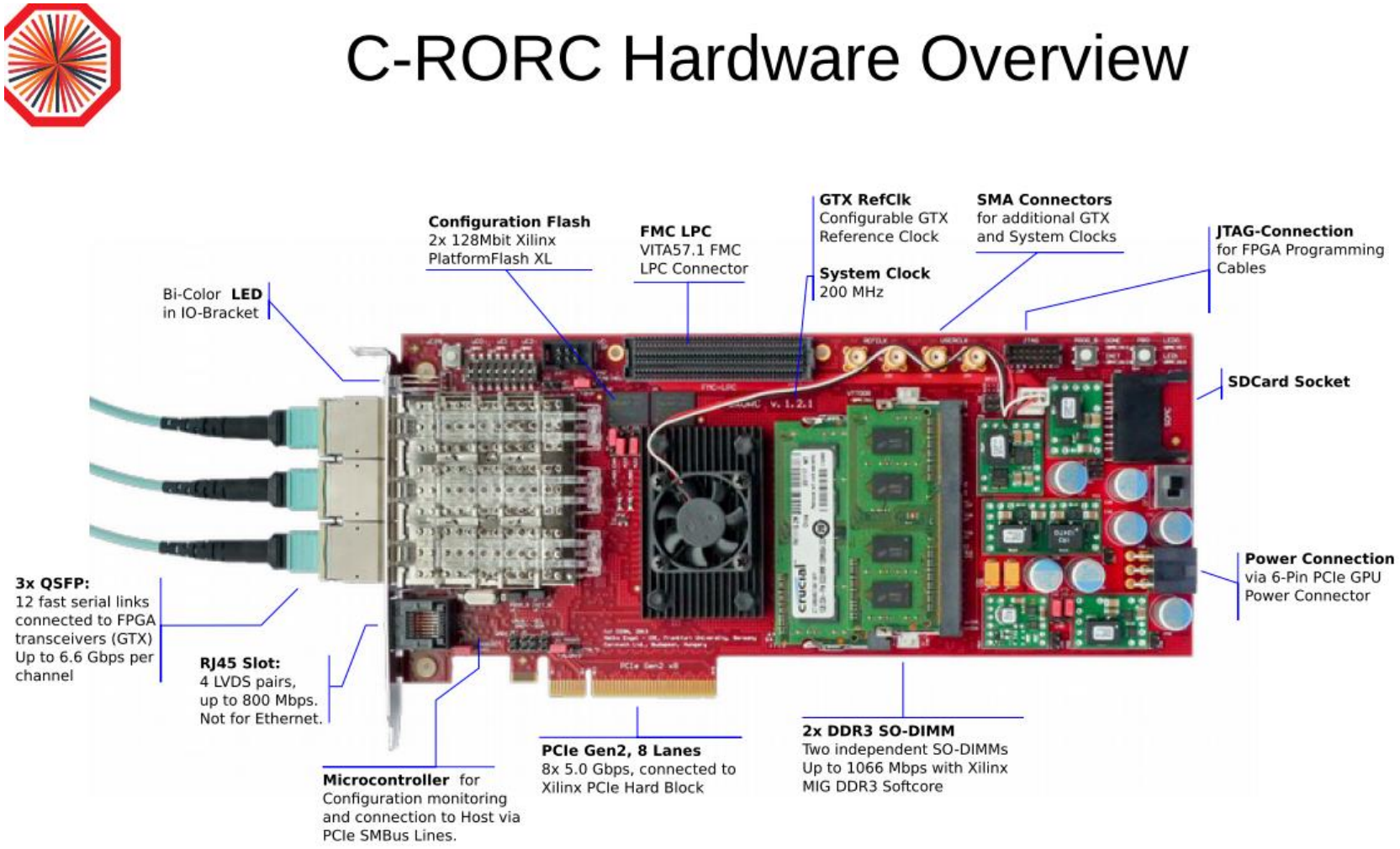
Comparison of setups

	RO boards	# of PCs	Output to HLT	Data-handling
COPPER-like	20-50 [1)]	20-50	1GbE [2)]	Software ☺
PCIe	20-50	20-50	1GbE	Software ☺☹
2 step	20-50	0	10GbE [3)]	firmware ☹
1 step	20-50	0	1GbE	firmware

We still have time to decide what to choose.

- Information of event size in actual data-taking will be obtained in the phase II run.
- Estimating processing and I/O ability(implementing many b2link cores and data processing function) by using a test board will be very useful in R & D phase.
- Hopefully, better/cheaper ‘commercial off-the-*shelf’ products will come.*
  - FPGA
  - Servers, NIC, switch, PCIe

ALICE RUN2 readout board



Former candidate for LHCb Run3

MiniDAQ1 hardware

CENTRE DE PHYSIQUE PARTICULES DE MARSE

AMC40 mezzanine + AMCTP carrier

- **AMC40**
  - Stratix5 FPGA
  - 3 MiniPOD AFBR-811VxyZ (Tx)
  - 3 MiniPOD AFBR-821VxyZ (Rx)
  - Up to 24 GBT/WB/GWT
  - Up to 12 10GBASE-R Ethernet
- **AMCTP**
  - Local 40/80 MHz oscillator
  - External clock input
  - COM Express Module
  - PCI Express x1 to FPGA
  - GbE to LAN

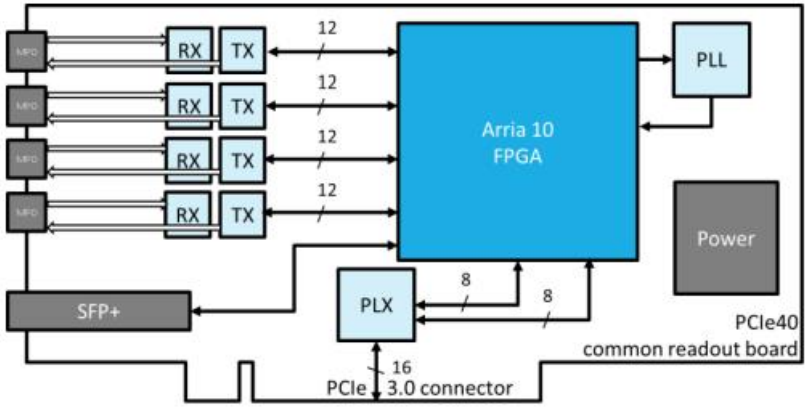
22/05/2017 TIPP 2017 - PAOLO DURANTE - MINIDAQ1 6

Current candidate for LHCb Run3

MiniDAQ2 hardware (PCIe40)

CENTRE DE PHYSIQUE PARTICULES DE MARSE

- PCI express add-in card
  - Full-length, full-height
- Arria10 FPGA
  - 2x resources as Stratix5
  - 24 links: 85% on S5 to 46% on A10
- High-density optical IO
  - Up to 48 bidirectional links
- PCIe Gen3.0 interface to Event Builder
  - Custom 100 Gb/s DMA engine
- Design has been validated
  - Full board self-test
- Initial production started
- Collaboration institutes have started to receive first devices



PXD DHH/ COMPASS

- Probably, details in Igor-san's talk

UT3/4

Universal Trigger module 3 (UT3)

- **FPGA : Virtex-6 HXT**
  - **FF1923 package : 3 FPGA choices**
  - **VHX380T ... 14 modules**
  - **VHX565T ... 14 modules**
    - **GDL : 4 (2 spares)**
    - **CDC : 18 (2 spares)**
    - **KLM : 2 (1 spares)**
    - **For test bench : 4**
- **IO**
  - **Main board**
    - **Clock : 1 in, 1 out**
    - **NIM : 2 in, 2 out**
    - **24 GTH (11 Gbps x 24)**
    - **LVDS : 64(32x2) in/out**
  - **GTX daughter board (optional)**
    - **40 GTX (6.25 Gbps x 40)**
  - **General IO board (optional)**
    - **Clock : 2 out**
    - **NIM : 8 in, 8 out**
    - **RJ-45 for Belle2Link : 4**

Other motivation for faster readout system ?

From b2note : “L1 Trigger Menu for Low Multiplicity Physics”
https://d2comp.kek.jp/search?ln=en&cc=Belle+II+Notes+%3A+Physics&sc=1&p=&f=&action_search=Search

Physics related with low multiplicity event
* Bhabhas, e+e- -> $\gamma\gamma$, e+e- -> $\mu+\mu$-
luminosity, calibration, QED physics topics
* single photon
- dark matter search: e+e- -> γ A'(-> $\chi\chi$) : A'=dark photon, χ=dark matter
* Initial State Radiation(ISR) : e+e- -> $\gamma\pi+\pi$-
- important for muon g-2 measurement
* tau 1 vs 1 final states :
- each τ has one charged track
- τ -> $\mu\gamma$ etc.
* pi0 transition form factor
- two photon -> pi0 production
* Y di-pion transition
- Y (2,3S)-> $\pi+\pi$- Y (1S) and Y (1S) -> ν ν bar or $\chi\chi$
* ~~$\gamma\gamma$ -> π 0 π 0~~

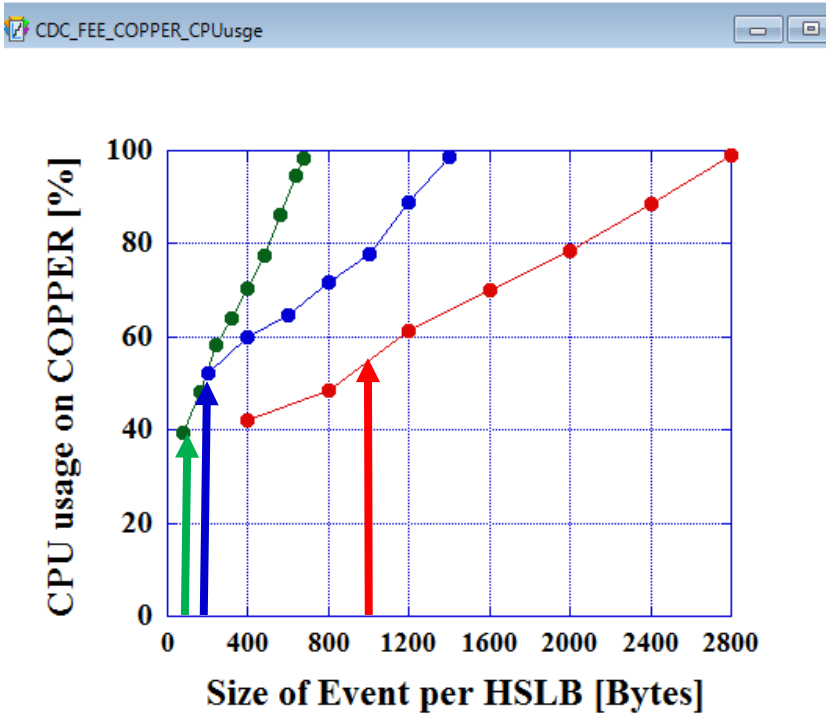
TABLE VIII: Efficiencies and Cross section after triggers

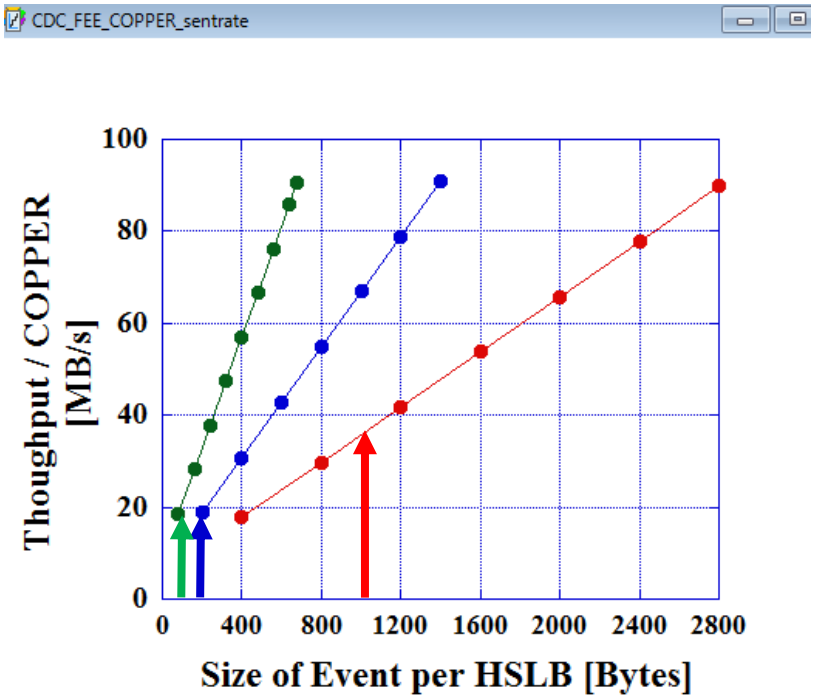
	Processes	T1:2trk	T2:1trk1mu	T3:1mu	T4:1trk1c	T1:bbc	T2:3g	T3:3t	Combine
ϵ(%)	$B^0\bar{B}^0$	-	96.5	50.0	82.9	44.8	93.4	99.4	> 99.9
	B^+B^-	-	96.5	51.7	84.1	46.2	92.6	99.5	> 99.9
	ccbar	-	96.8	65.9	89.4	52.1	84.8	98.0	> 99.9
	uds	-	96.5	68.0	89.1	50.0	81.1	97.2	> 99.9
	$\tau\rightarrow$generic	51.0	60.0	57.2	62.6	28.1	55.6	29.1	94.2
	$\tau\tau$(1v1)	81.0	58.1	61.8	61.3	27.9	47.4	-	97.3
	$\tau\rightarrow e\gamma$	80.0	55.1	56.0	91.7	52.3	85.7	-	99.0
	$\tau\rightarrow\mu\gamma$	76.1	48.1	46.2	87.7	57.9	82.2	-	97.1
	$\pi\pi(\gamma)$	67.9	51.9	67.4	80.0	43.4	42.5	-	97.4
	$\pi\pi(\gamma)$[0,1]	66.7	49.4	66.3	79.1	43.0	38.6	-	97.2
	$B\rightarrow\pi^0\pi^0$	11.1	83.4	35.4	96.3	92.4	17.0	81.7	> 99.9
	$\mu\mu$	98.9	94.5	99.7	-	-	-	-	> 99.9
σ(nb)	eeee	2.2	0.1	0.1	1.1	0.8	0.9	0.1	3.4
	$ee\mu\mu$	2.6	0.8	0.7	0.1	0.1	0.5	0.1	3.3
	$ee(\gamma)$	7.2	7.3	10.5	11.1	13.1	2.9	0.6	32.2

- If there are some trigger modes with low efficiency, lowering threshold with reinforced RO system may contribute the improvement of the efficiency.
- But, it is not straightforward for the Belle II experiment, where trigger efficiency is already high.

A. Bottlenecks of the current COPPER readout system

RED: 1HSLBs/COPPER (SVD)
BLUE: 2HSLBs/COPPER (ECL)
GREEN: 4HSLBs/COPPER (CDC,TOP,ARICH,KLM)



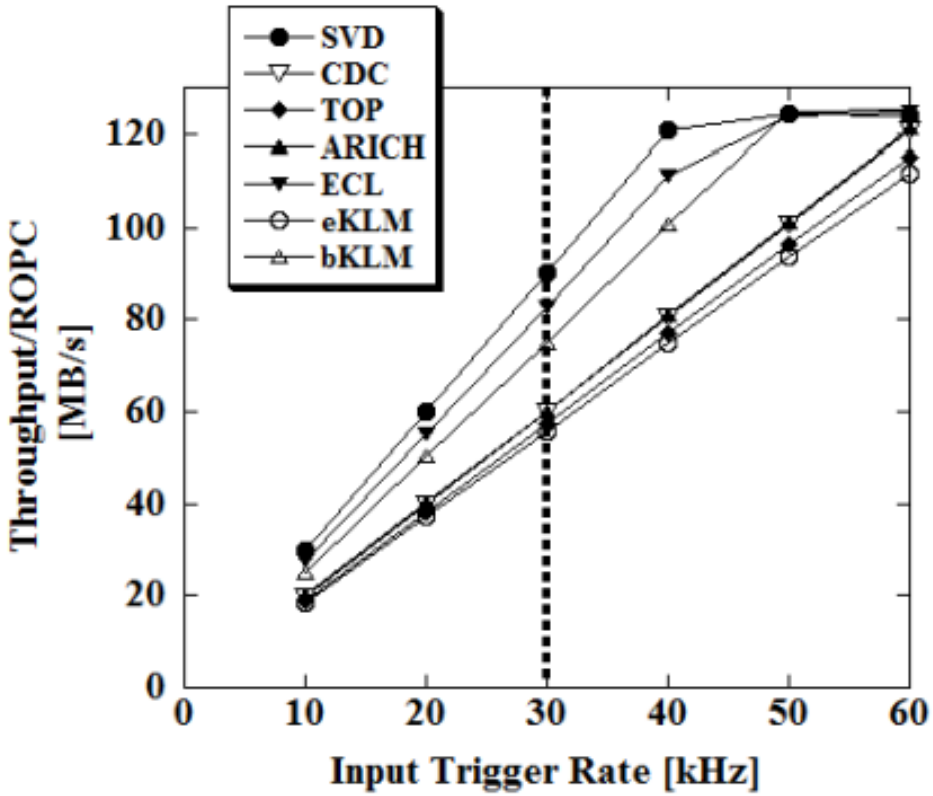


https://confluence.desy.de/display/BI/DAQ+EventSizeOfEachSubDetector

	#ch	occ [%]	#link	/link [MB/s]	#CPR	ev sz [kB]	total [MB/s]	/CPR [MB/s]
PXD	8	2	40	455	—	800	1820	—
SVD	223744	1.7(5.5)	48	8.9(33.8)	48	14.9	428	8.9(33.8)
CDC	14336	10	302	0.6	76	6	175	2.3
BPID	8192	2.5	64	1.5	16	3.2	96	8
EPID	65664	1.5	~~90~~ 72	1.1	~~25~~ 18	2.8	84	4.2
ECL	8736	33	52	7.7	26	12	360	15
BKLM	19008	1	24	9.7	6	2	60	10
EKLM	16800	2	16	35.8	~~9~~ 4	1.4	42	4.7
TRG			19		10			

- COPPER CPU usage will be the bottleneck.

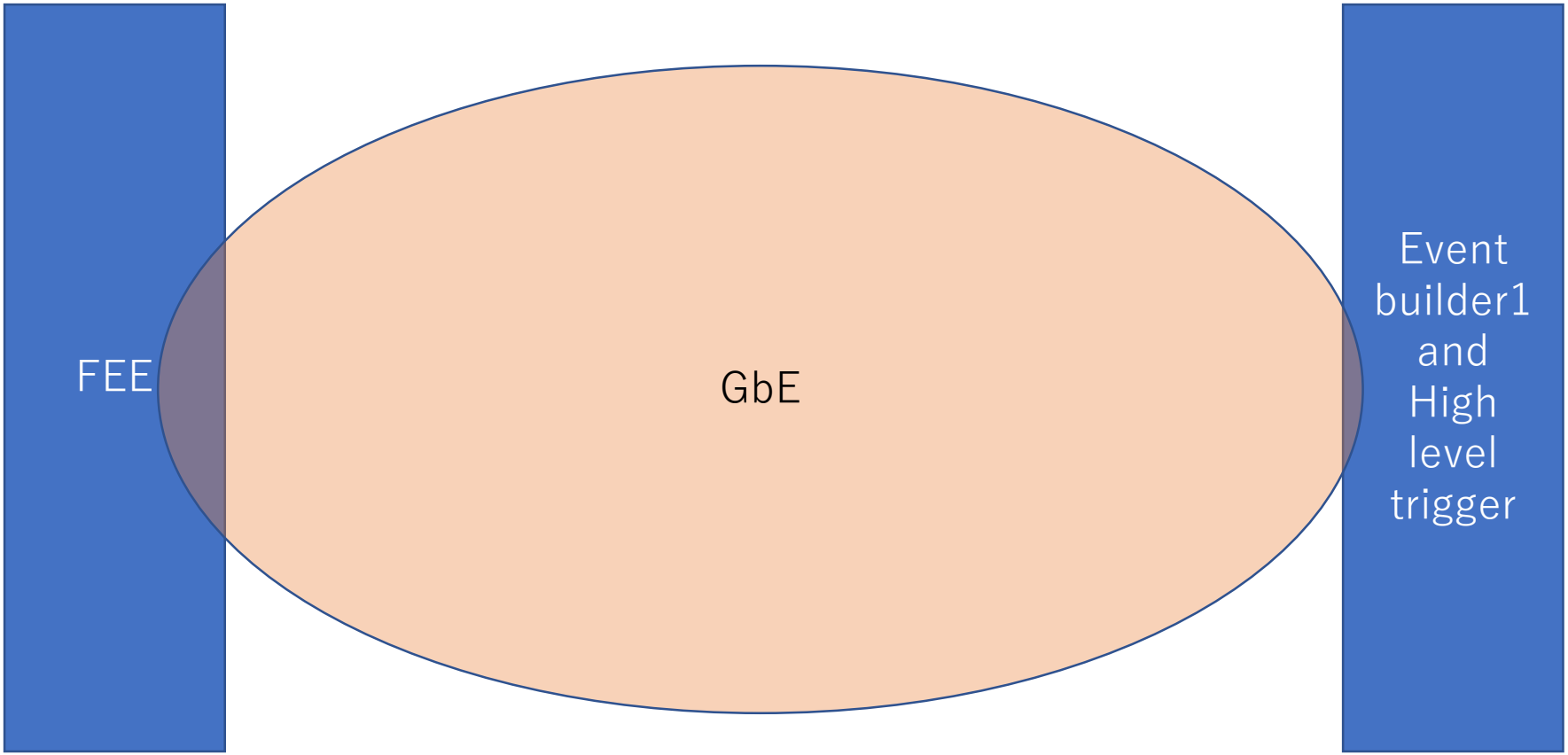
B. Bottlenecks of the readout PC



- Throughput is saturated due to the limit of output GbE bandwidth.

* Bottlenecks
- COPPER -> CPU
- ROPC -> network output

If the update of FEE sides is allowed …



Better Debugging/Maintenance ? :

- Firmware-debugging seems to take more than x10 long time than software-debugging for non-experts like me …

But

- A lot of firmware update in FEE sides
- Probably, don't have enough buffer on FEE boards
  - Busy signal to FTSW like COPPER ?

→ Difficult