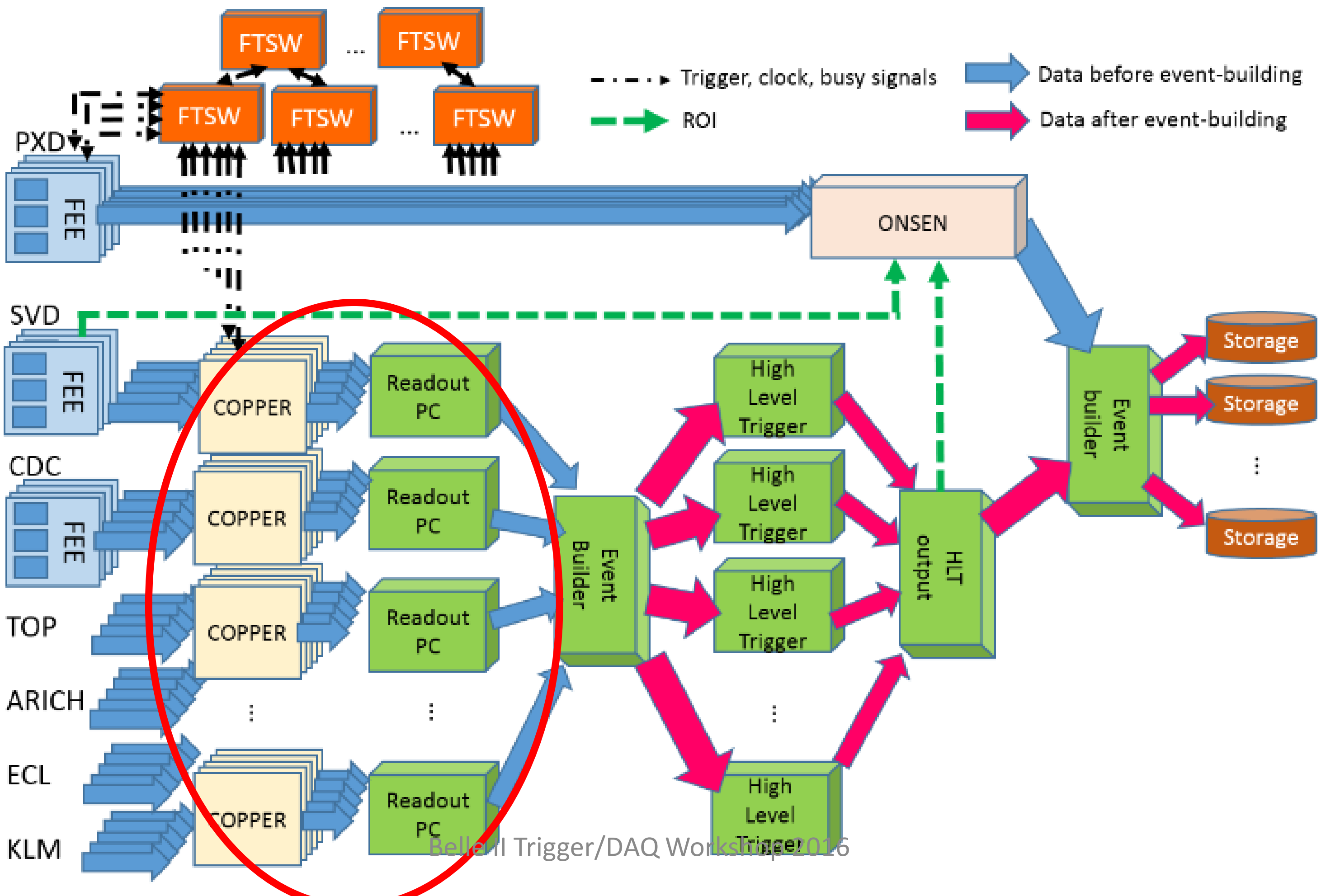
Upgrade plan 2

S. Yamada (KEK)

1. Motivation

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

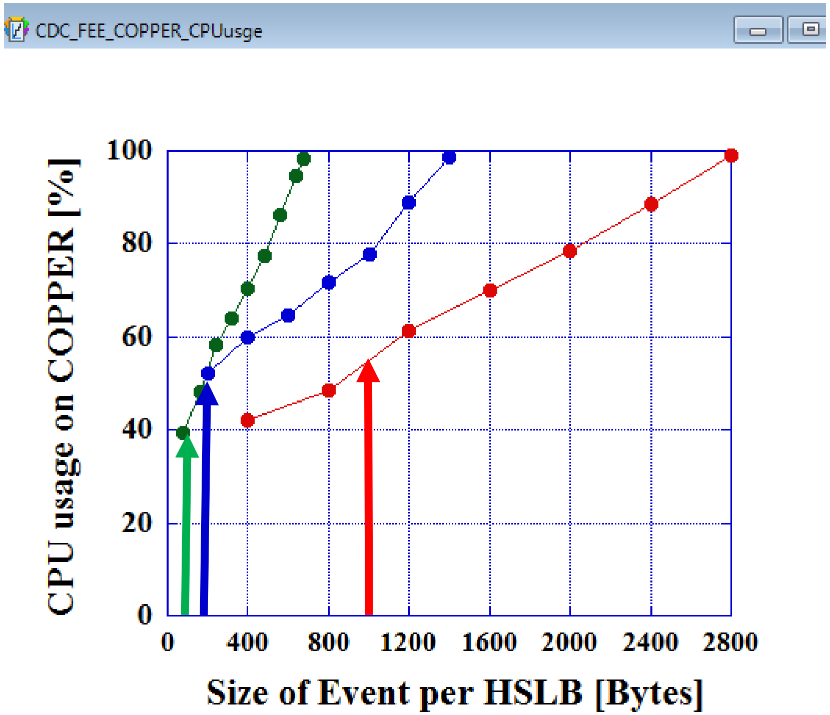
- The number of discontinued parts is increasing.
  - e.g. chipset on a PrPMC card, FIFO and LAN controller on COPPER III
  - For older COPPER II, it is basically difficult to replace parts according to manufacturer.
- Four different types of boards(COPPER, TTRX, PrPMC, HSLB) should be taken care of.

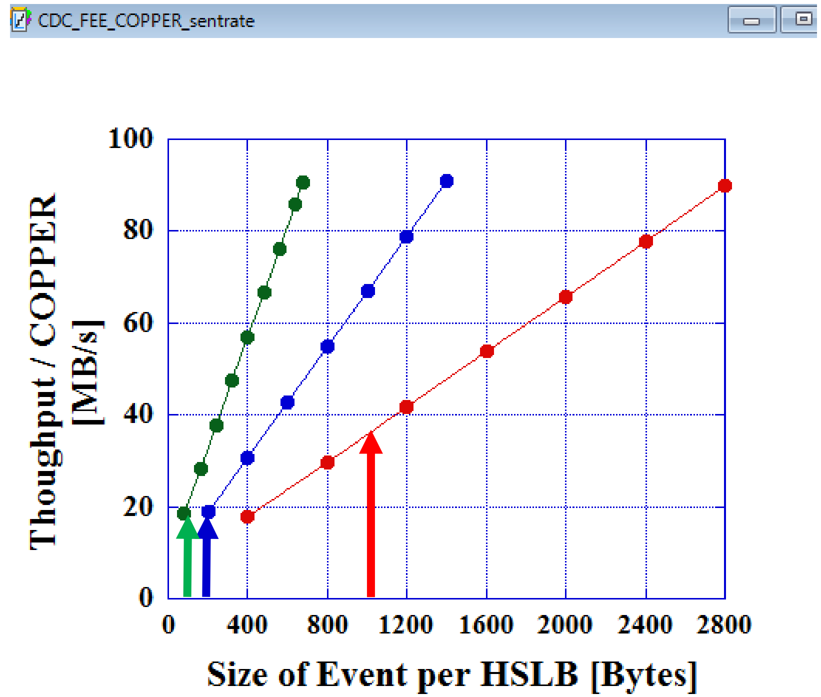
Limitation in the improvement of performance of DAQ

- 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.

A. Bottlenecks of the current COPPER readout system

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



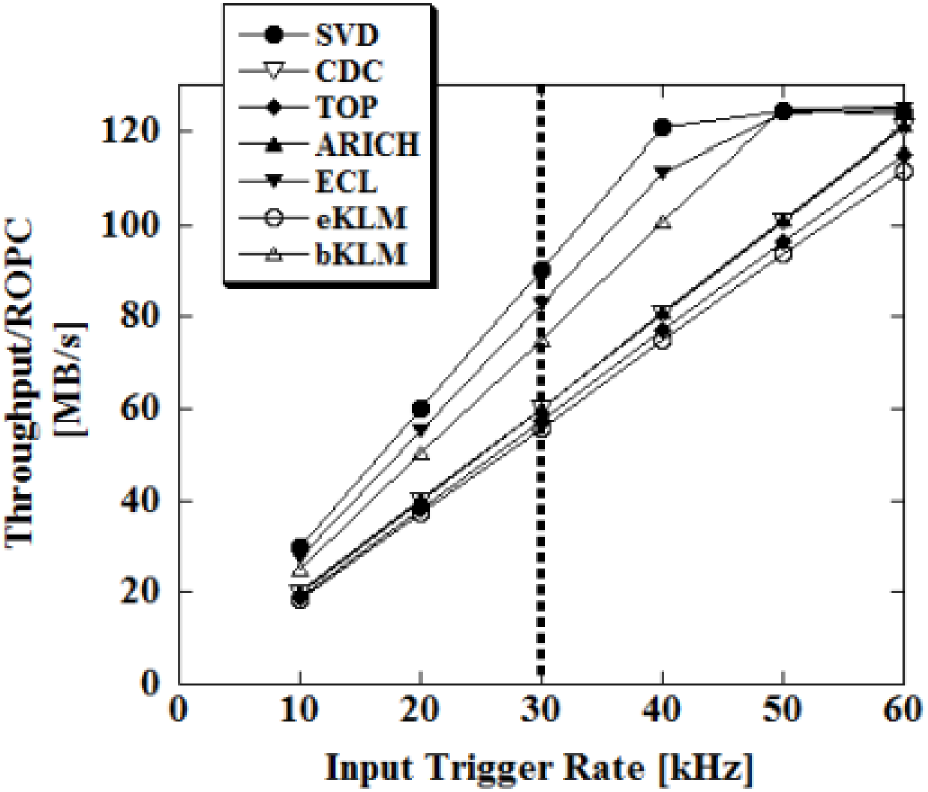


https://confluence.desy.de/display/BI/DAQ+EventSizeOfEachSubDetector

	#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	~~23~~ 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

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- -> γγ, e+e- -> μ+μ-
luminosity, calibration, QED physics topics
* single photon
- dark matter search: e+e- -> γA'(->χχ) : A'=dark photon, χ=dark matter
* Initial State Radiation(ISR) : e+e- -> γπ+π-
- important for muon g-2 measurement
* tau 1 vs 1 final states :
- each τ has one charged track
- τ->μγ etc.
* pi0 transition form factor
- two photon -> pi0 production
* Υ di-pion transition
- Υ(2,3S)-> π+π- Υ(1S) and Υ(1S) -> ν νbar or χχ
* γγ->π0π0

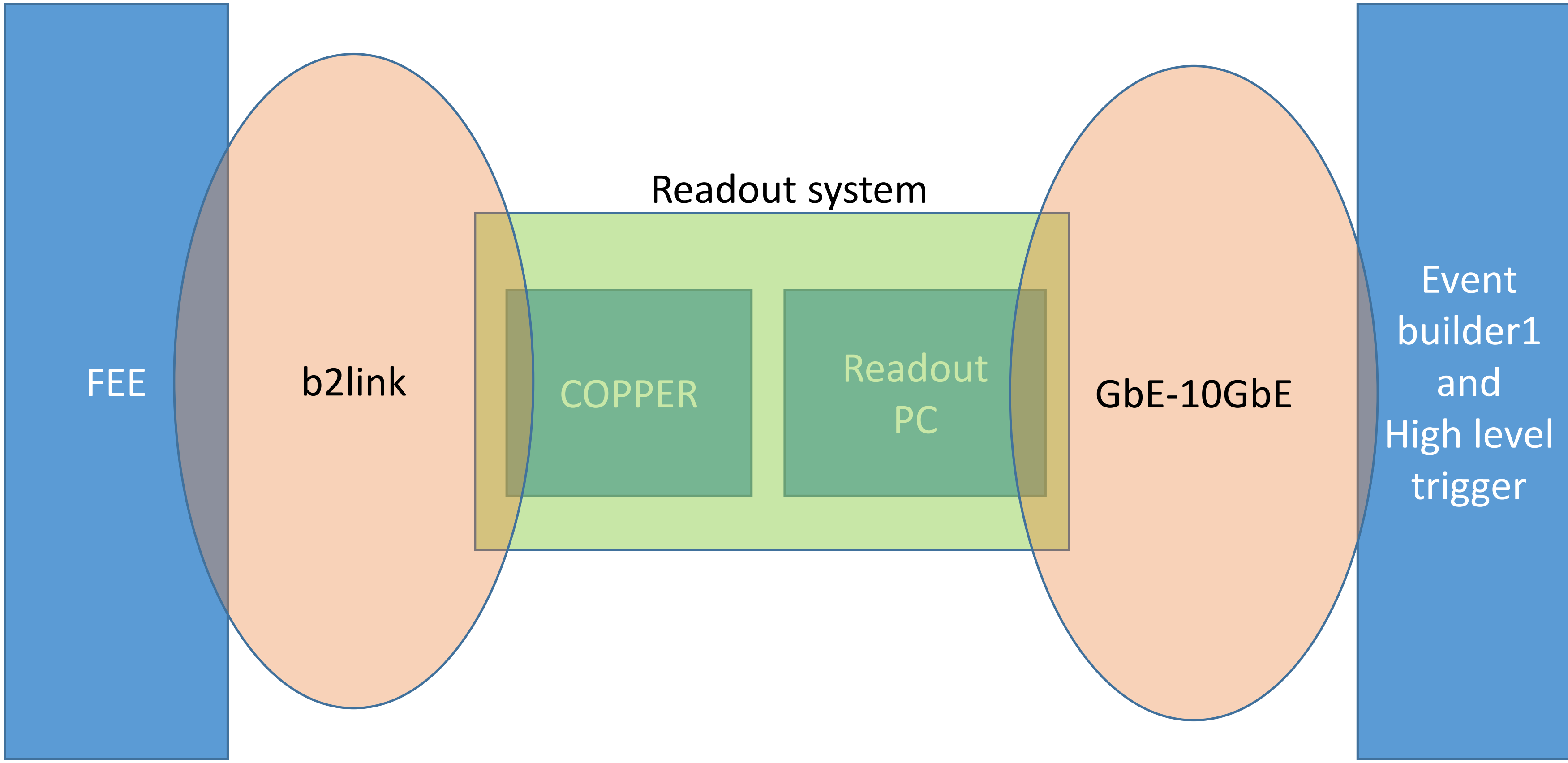
TABLE VIII: Efficiencies and Cross section after triggers

	Processes	T1:2trk	T2:1trk1mu	T3:1mu	T4:1trk1c	T1:bbc	T2:3g	T3:3t	Combine
$\epsilon(\%)$	$B^0\bar{B}^0$	-	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
	$\tau \to$generic	51.0	60.0	57.2	62.6	28.1	55.6	29.1	94.[illegible]
	$\tau\tau$(1v1)	81.0	58.1	61.8	61.3	27.9	47.4	-	97.3
	$\tau \to e\gamma$	80.0	55.1	56.0	91.7	52.3	85.7	-	99.0
	$\tau \to \mu\gamma$	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
	$B \to \pi^0\pi^0$	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.

2. Requirement

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.

What is required for the Belle II readout system

- Functionality
  - Interface with FEE and HLT
    FEE : B2link
    HLT : Ethernet
  - Partial event-building
  - Data-formatting and reduction

- Performance
  - Accumulation of inputs
  - Processing data with Belle2link line rate
  - Large data output rate

Key factors

- Data flow : Gigabit, 10GbE, Rocket I/O …
- Data processing : CPU/FPGA processing power

Dataflow

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

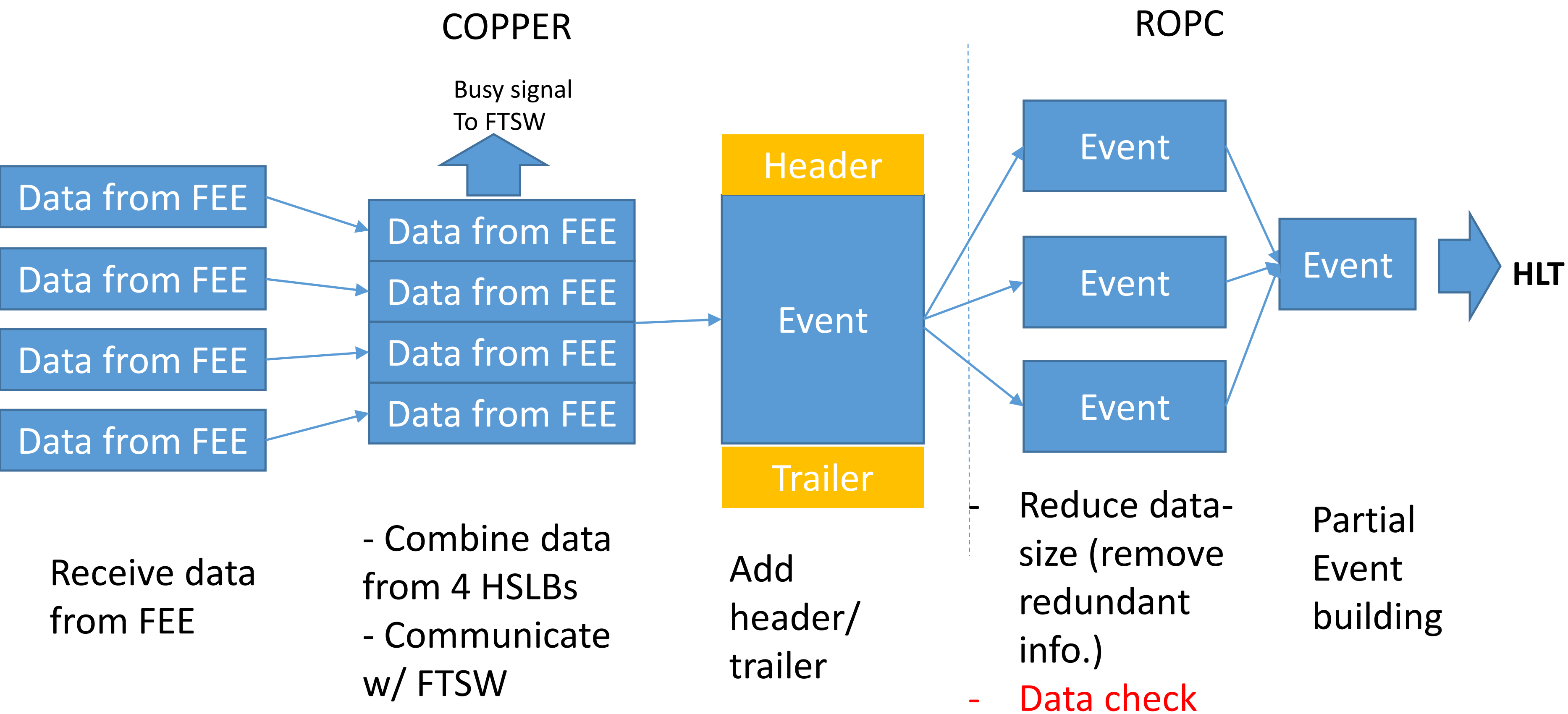
- Data flow per b2link is not so large.

-> 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

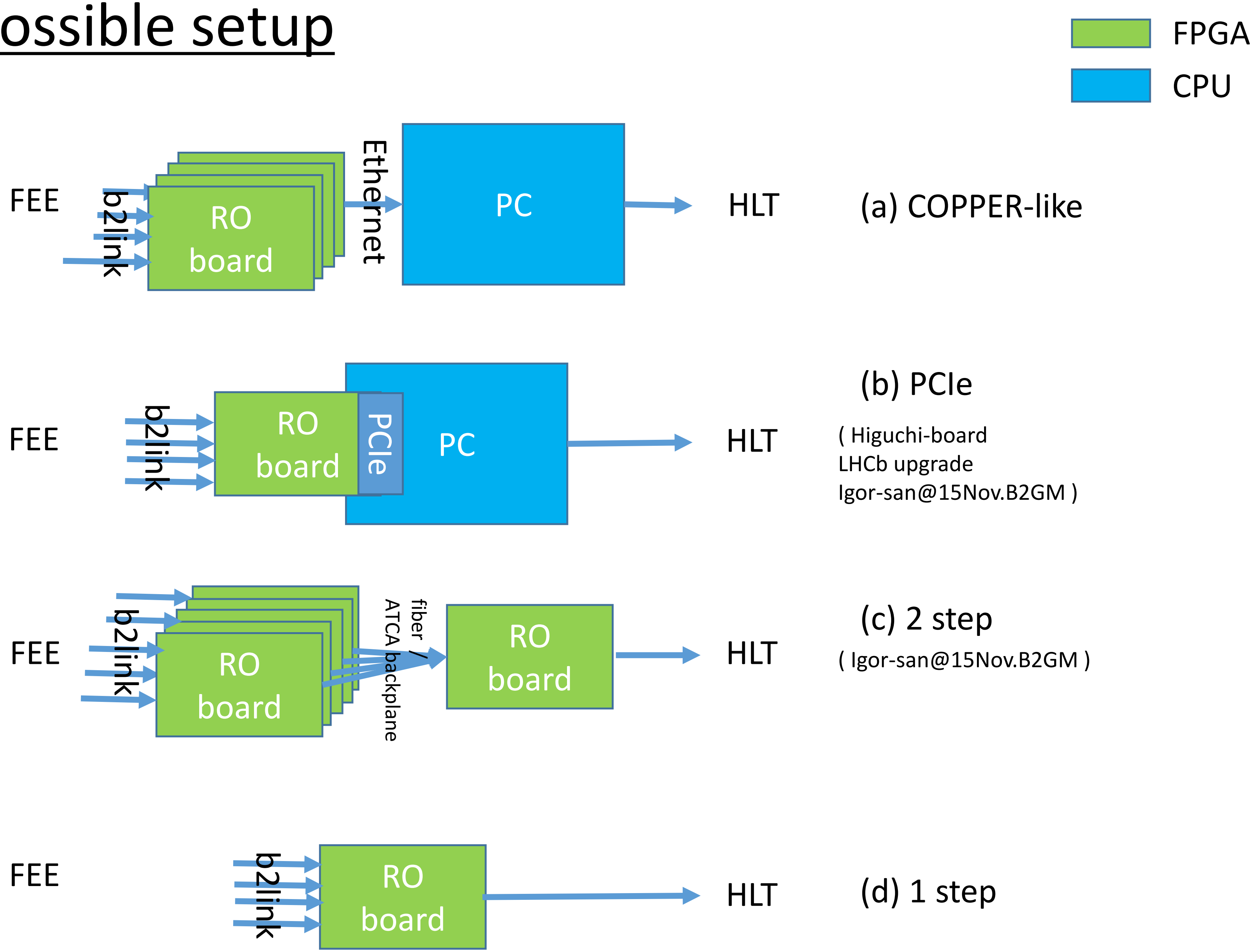
Data processing :



- Not so complicated operation, which should be done by software.
- 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

3. Possible setups

Possible setup



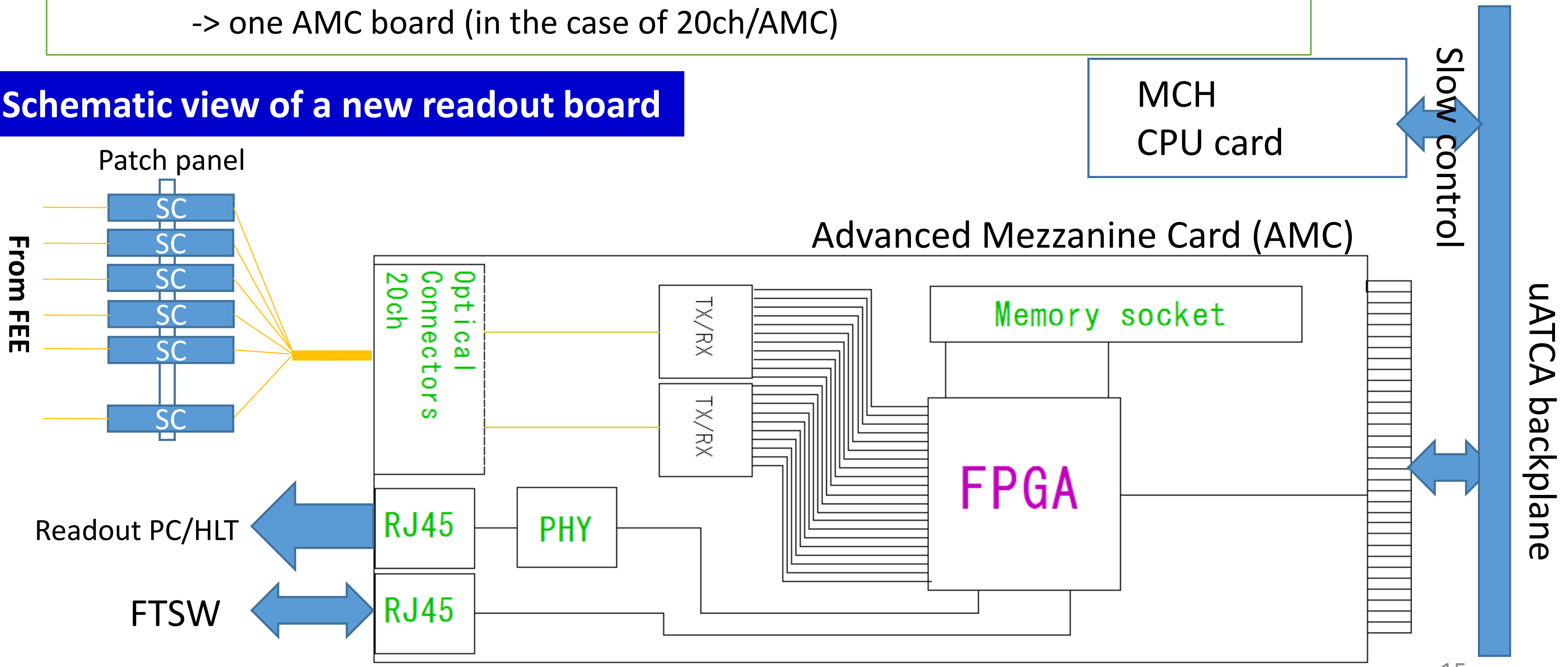
Example of a board sketch

(2015. 8.6 MOF meeting)

New readout system = High-density FGPA-based system using uTCA

- 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 board



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

Summary

- Even though we have not started the Belle II experiment, it is useful to start thinking possible option of future uprade 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.

- ‘Input : belle2link’ and ‘output:Ethernet’ will be the boundary condition.
- Compared with the current system, it is likely to reduce # of boards drastically.

- Hopefully, more information about the throughput to extrapolate will be obtained in phase-II run.