

S. Yamada (KEK)

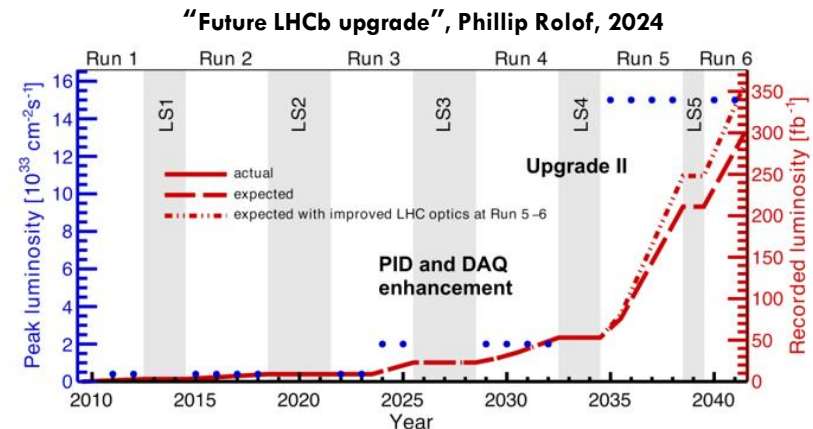
PCIE400

LHCb ONLINE LS3 ENHANCEMENTS

- LHCb Online LS3 enhancements
 - EB expansion in cavern
 - implementation of PCIe400
 - PCIe400: 47x IpGBT links, 440 Gb/s Output for RICH (and ECAL) readout.

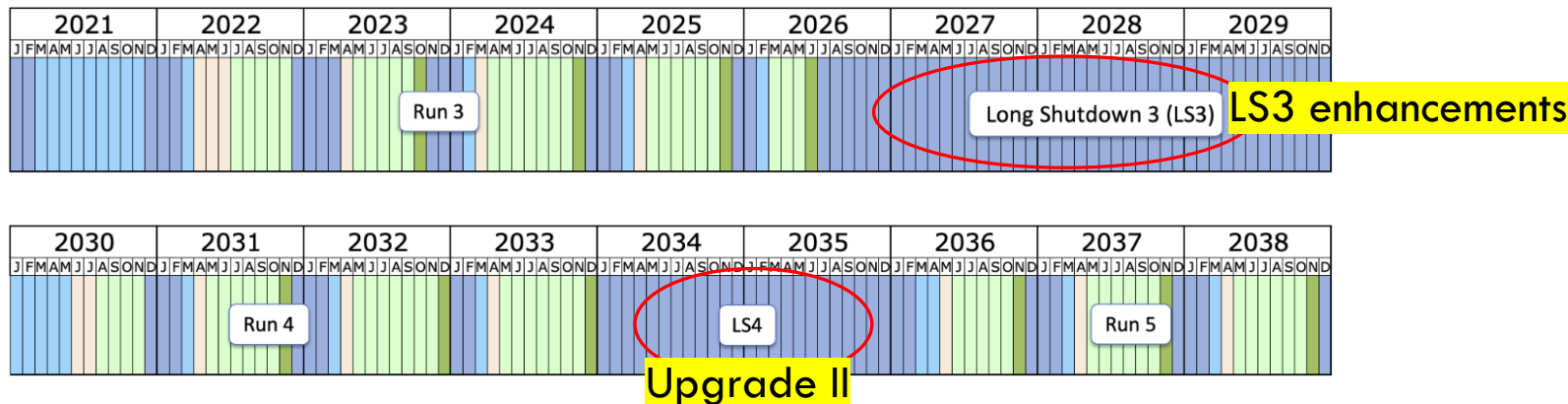
Sub-detector	DAQ links	PCIe40s	PCIe400s	EB servers	EB network links
RICH	2500	0	55	0	16
ECAL	440	20	0	7	14
Total	2940	20	55	7	30

“LHCb Data Acquisition Enhancement Technical Design Report”



https://indico.cern.ch/event/1571894/contributions/6622541/attachments/3126503/5545495/LHCC_open_session_LHCb.pdf

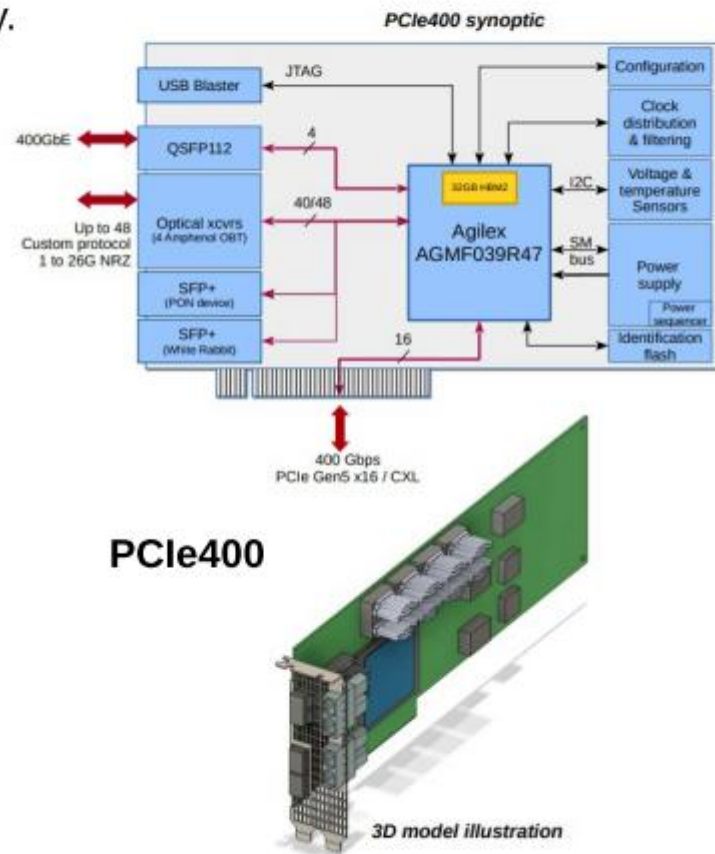
LHC long-term schedule (<http://lhc-commissioning.web.cern.ch/schedule/LHC-long-term.htm>)



- PCIe40: Has been utilized in upgrade of LHCb, ALICE, and Belle II
 - Intel Arria10
 - PCIe Gen3 x 16.
 - Belle II just completed the commissioning recently.
- PCIe400: next generation of readout board
 - Intel Agilex.
 - PCIe Gen5 x 16.
 - Planning for LHCb upgrade.
- Collaboration with CPPM Marseille:
 - Joint study on the PCIe DMA based readout firmware logic and driver software development.
 - PCIe40, PCIe400, and Versal boards.



PCIe40



source: CPPM Marseille group

PCIe400 Application to Belle II ?

COMPARISON BETWEEN PCIE40 AND PCI400

FPGA : PCIE40:ARRIA10 -> PCIE400:AGILEX 7

	GX1150	AGM 039
Logic elements (LEs)	1,150,000	3,851,520
Adaptive logic modules (ALMs)	427200	1,305,600
M20K memory blocks	2713	18,960
M20K memory size (Mb)	53	370
MLAB memory size (Mb)	12.7	40
HBM2E High Bandwidth DRAM Memory (Gbytes) Size		16 / 32

Interface with host server (PCIExpress)

- PCIe40 : 2x PCIe gen3 8lanes ~ 100Gbps
- PCIe400 : PCIe gen4 16lanes ~ 400Gbps

of XCVR

- PCIe40 : 48ch(bidirectional) up to 10Gbps
- PCIe400 : 48ch(bidirectional) up to 25Gbps

PCie400 as a spare of PCIe40

Current PCIe40 availability

Location	# of PCIE40
Used in readout PCs	21
Test benches in KEK	3
Shandong Univ.	1
VTX test in IFIC	2
Available spares	~4

There are one or two prototype(?) version boards in KEK and UH.

- So far, there have been no significant board failures.
 - One rklm1 PCIe40 board was replaced to check the frequent b2lllost issue between ch0 of rklm1 and dc38 although it is unclear if this issue came from the PCIe40.
- However, once these boards start to fail, we can easily expect a desperate need to increase the number of spare boards, even by one or two.
- It would be meaningful to get a PCIe400 card as a spare card

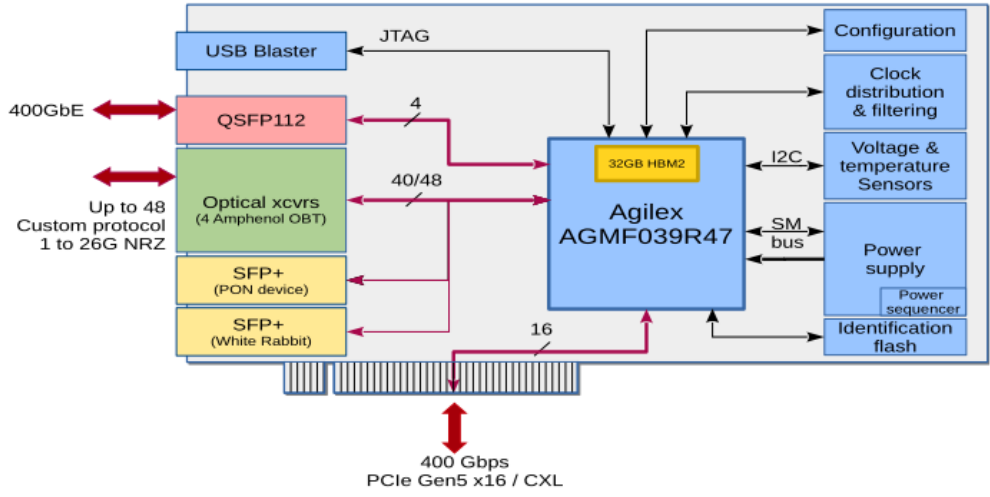
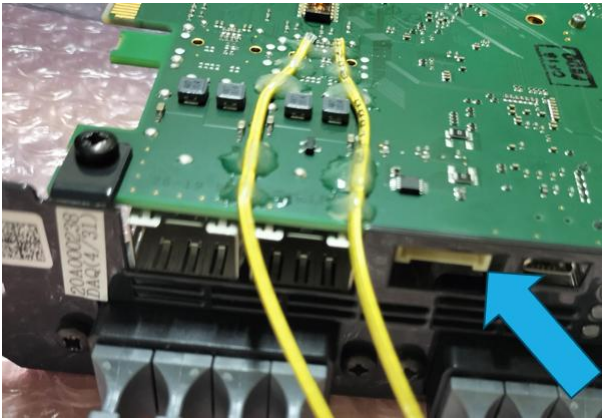
LHCb Data Acquisition enhancement TDR CERN-LHCC-2024-001

¹It should be stressed, that so far there is no indication whatsoever of any ageing effect in the PCIe40s, but by the end of Run 4 they will have been in continuous operation for more than 10 years.

B2TT INTERFACE ..

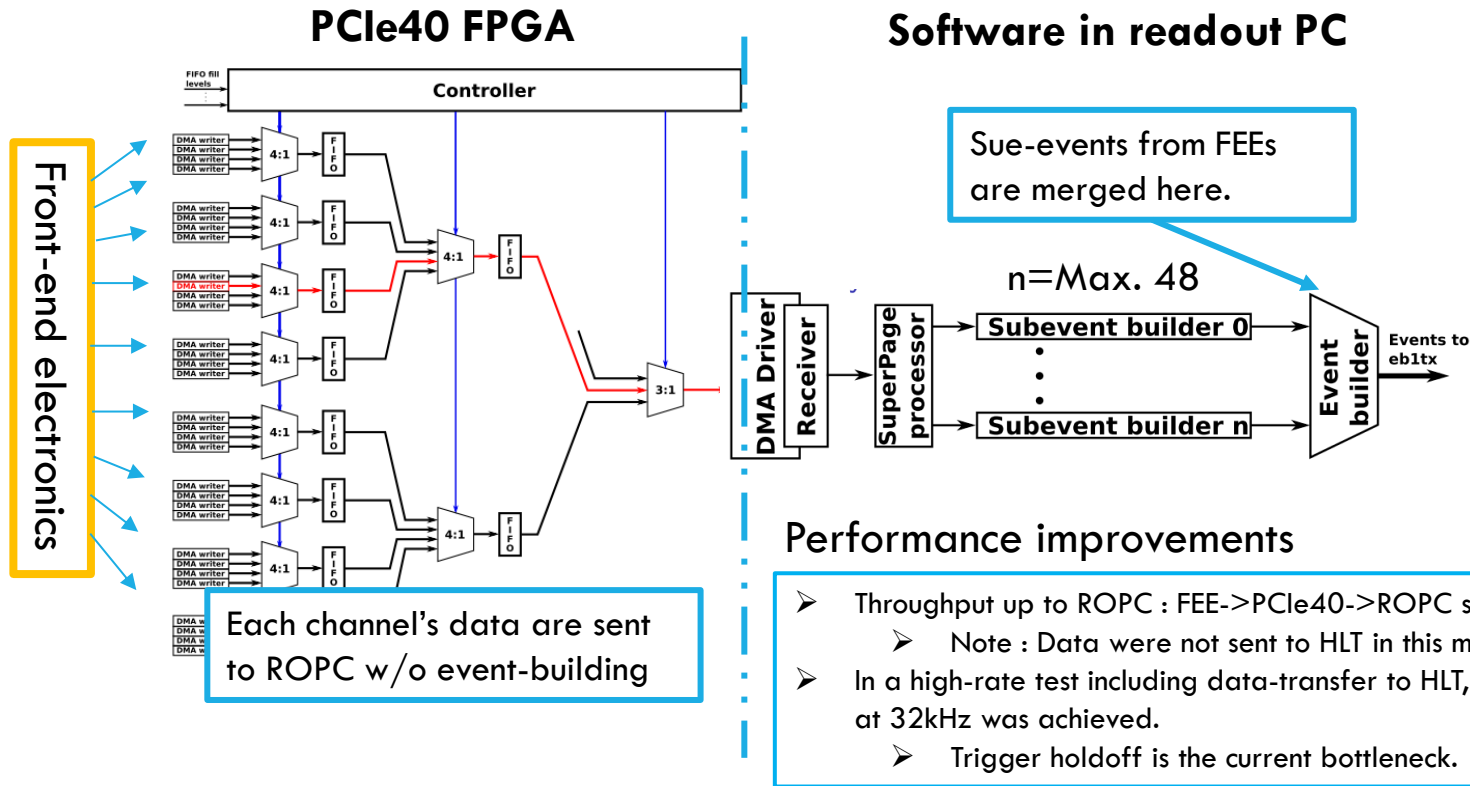
- As far as I know, unlike the PCIe40, the PCIe400 has no connector for accessing the FPGA GPIO.
- Therefore, data transmission and reception must be done through the FPGA transceivers.
- We need to send data to the FTSW synchronously with the original FTSW clock, but that clock is not available.
- Same problem with the current optical b2tt project for PCIe40.
- If the FTSW (Vertex-5 FPGA) can receive data via its transceiver and perform CDR, could that solve the issue?

Connector for b2tt interface on PCIe40



READOUT : VARIOUS IMPROVEMENTS IN PCIe40 FIRMWARE AND SOFTWARE

(D. Levit)



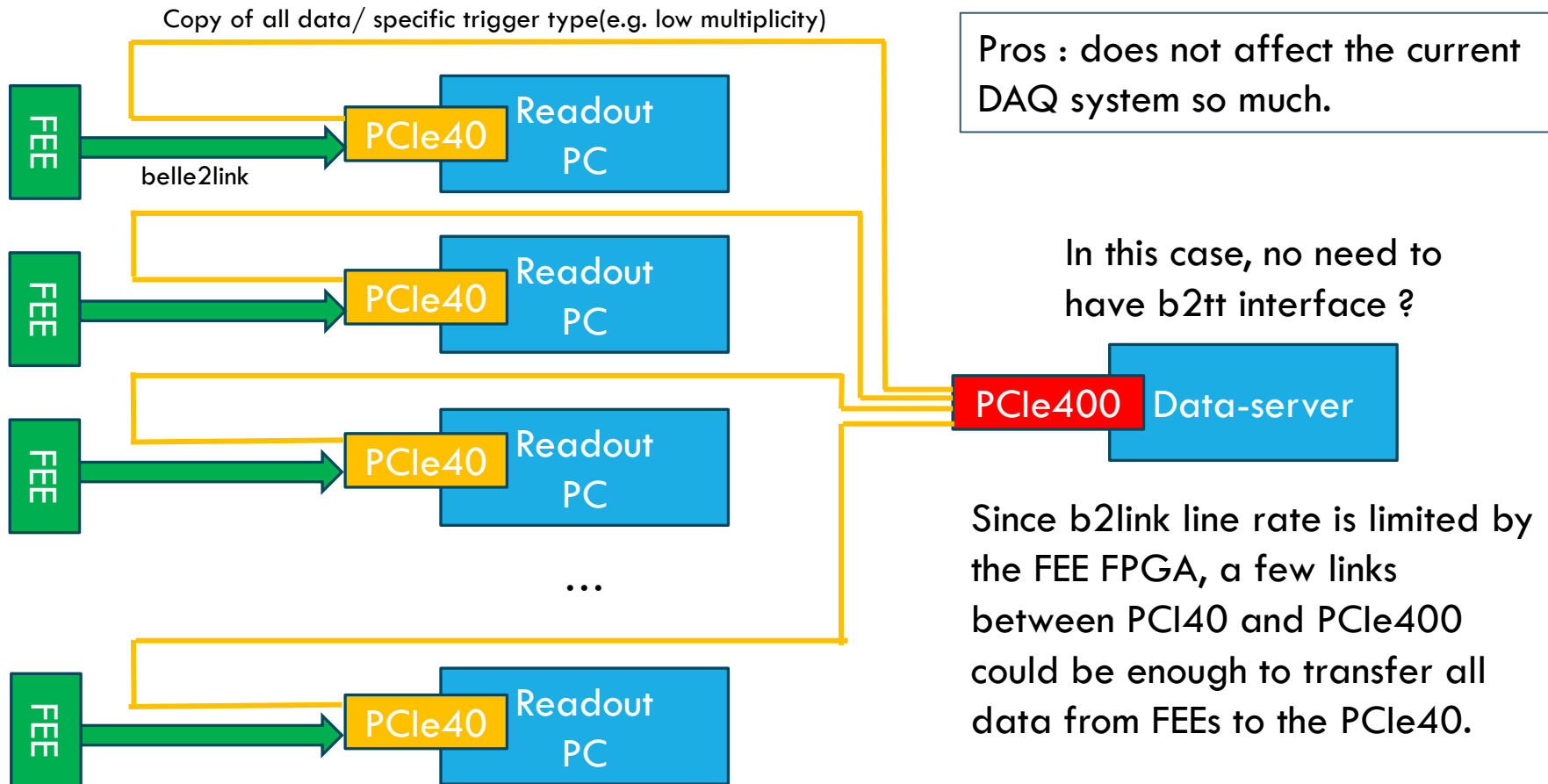
- ▶ move CRC calculation from the subevent class to the event class
 - ▶ subevent checks data from different pages independently
 - ▶ event class copies all data to contiguous memory space
 - ▶ more straight forward CRC check
 - ▶ Performance comparison at B4 test bench with 3 event builder threads:
 - ▶ throughput 3.5 GB/s
 - ▶ throughput per read-out thread: 1.3 GB/s (short run time before FEE BUSY)
 - ▶ CRC in the subevens class: **760 % CPU**
 - ▶ CRC in the event class: **550 % CPU**
- ⇒ processing larger contiguous data blocks more efficient!

3.5GB/s < 40Gbps=5GB/s
 The performance of the current readout system is limited by CPU usage.
 So, replacing PCIe40 with PCIe400(=400Gbps) is too much ?

Recent improvements in CPU usage in a readout PC

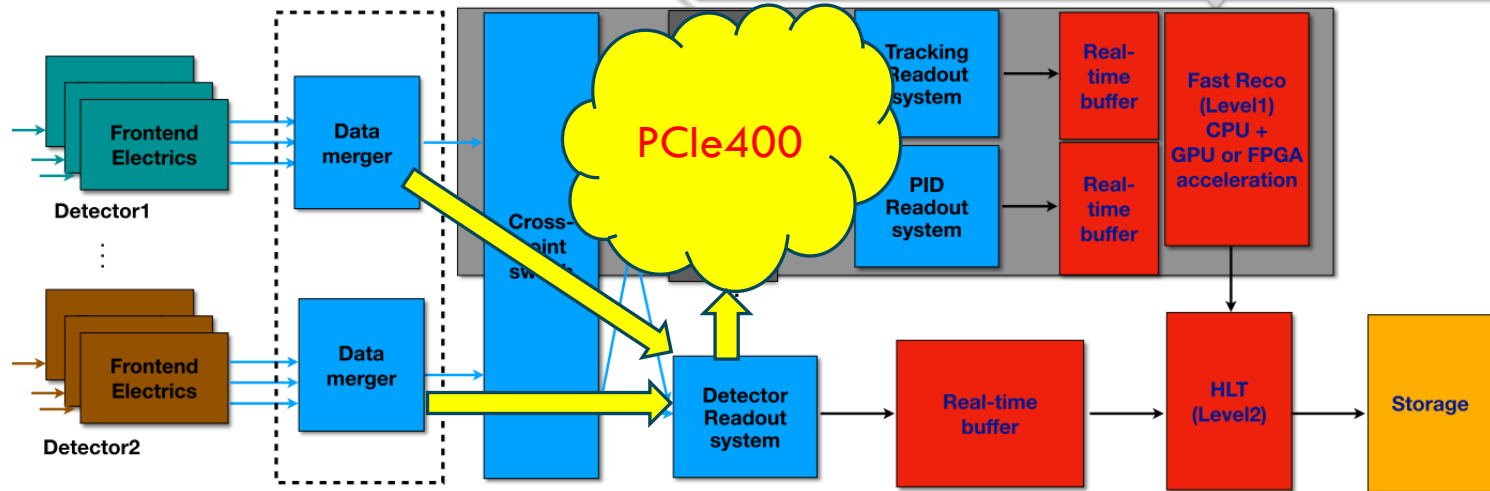
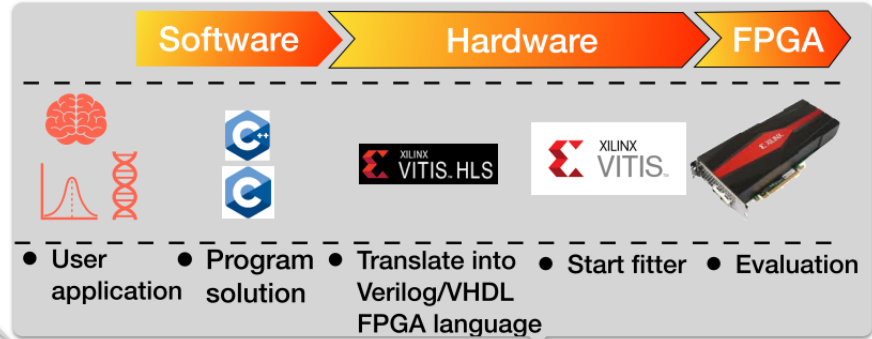
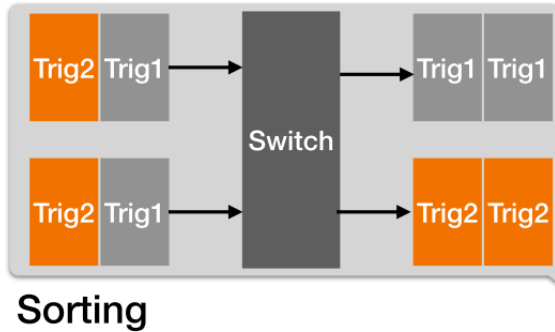
Other usage ?

SPLITTING DATA FOR INDEPENDENT ANALYSIS : ONE -> HLT, THE OTHER TO PCIE400



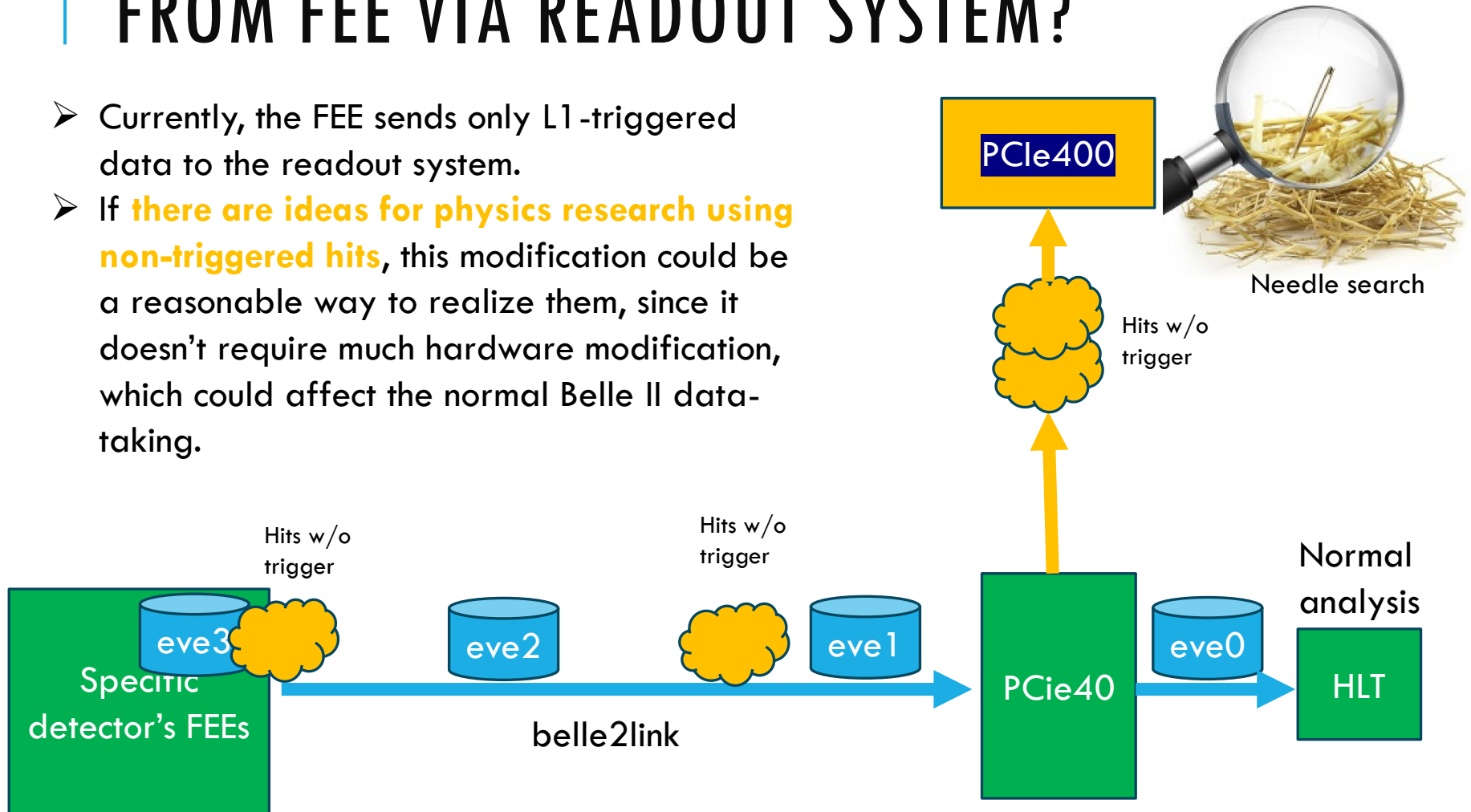
DUPLICATING DATA : HELP OF QIDONG-SAN'S HARDWARE HLT ?

Partial software trigger



EXTRACT MORE THAN L1-TRIGGERED DATA FROM FEE VIA READOUT SYSTEM?

- Currently, the FEE sends only L1-triggered data to the readout system.
- If **there are ideas for physics research using non-triggered hits**, this modification could be a reasonable way to realize them, since it doesn't require much hardware modification, which could affect the normal Belle II data-taking.



SUMMARY

- LHCb will produce PCIe400 boards in the coming years for the LS3 enhancements and Upgrade II.
- The PCIe400 will feature a larger FPGA and a higher-bandwidth PCI Express interface.
- The number of spare PCIe40 boards for Belle II DAQ is limited, and no additional production of PCIe40 is foreseen.
- The PCIe400 could serve as a spare for the PCIe40 and may also be applicable to other high-speed readout systems.

backup

Evolution of LHCb

Pre-Upgrade I (Run 1 and 2):

- $L_{\text{peak}} = 4 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$
- $L_{\text{int}} = 9 \text{ fb}^{-1}$

After Upgrade I (in LS2):

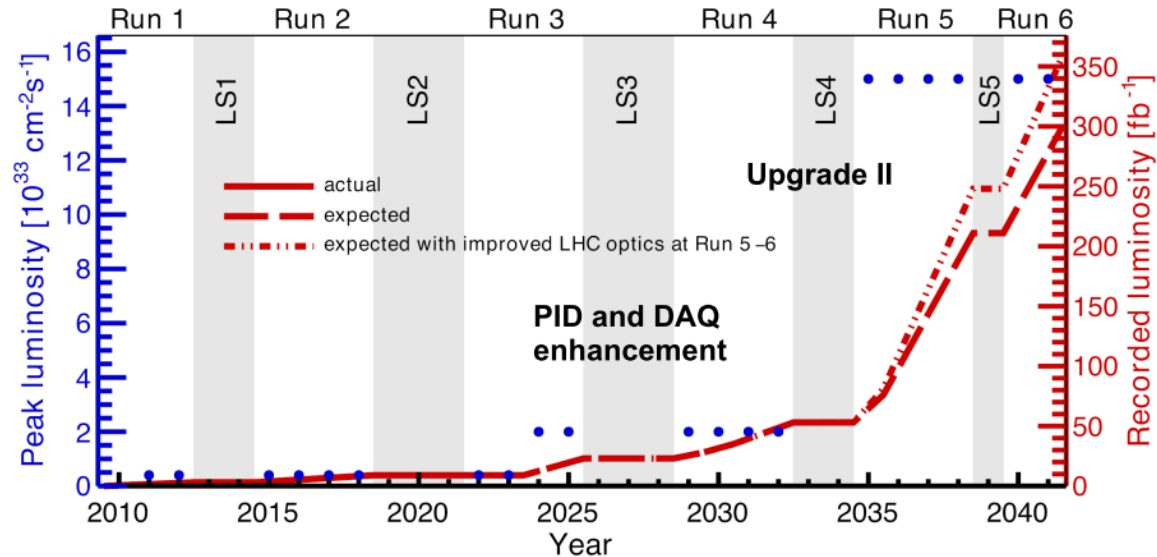
- $L_{\text{peak}} = 2 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$
- $L_{\text{int}} = 50 \text{ fb}^{-1}$ (Run 3 & 4)

PID and DAQ Enhancement (in LS3):

- Improved ECAL granularity
- Improved RICH timing
- Improved DAQ and trigger

Upgrade II (in LS 4):

- L_{peak} up to $1.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$
- $L_{\text{int}} = 300 \text{ fb}^{-1}$ (Run 5 & 6)



Product Line		AGM 032	AGM 039
Resources	Specification Options ¹	A, B, E, F, G, H	A, B, E, F, G, H
	Logic elements (LEs)	3,245,000	3,851,520
	Adaptive logic modules (ALMs)	1,100,000	1,305,600
	ALM registers	4,400,000	5,222,400
	M20K memory blocks	15,932	18,960
	M20K memory size (Mb)	311	370
	MLAB memory count	55,000	65,280
	MLAB memory size (Mb)	33	40
	HBM2E High Bandwidth DRAM Memory (Gbytes) Size	16 / 32	16 / 32
	Fabric PLL	8	8
	I/O PLL	16	16
	Variable-precision digital signal processing (DSP) blocks	9,375	12,300
	18 x 19 multipliers	18,750	24,600
	Single-precision or half-precision tera floating point operations per second (TFLOPS)	14 / 28	18.4 / 37
	Maximum Available Device Resources	Maximum EMIF x72	4
Memory devices supported		DDR5, LPDDR5, and DDR4	
Maximum AIB Interfaces		4	
Secure device manager (SDM)		AES-256/SHA-256 bitstream encryption/authentication, physically uncloneable function (PUF), ECDSA 256/384 boot code authentication, side channel attack protection	
Hard processor system (HPS)		Quad-core 64 bit Arm Cortex®-A53 up to 1.41GHz with 32KB I/D cache, Neon® coprocessor, 1 MB L2 Cache, direct memory access (DMA), system memory management unit, cache coherency unit, hard memory controllers, USB 2.0 x2, 1G EMAC x3, UART x2, SPI x4, I2C x5, general purpose timers x7, watchdog timer x4	

Arria® 10 FPGA and SoC Product Table

Version 2025.09.23

Product Line	GX 160	GX 220	GX 270	GX 320	GX 480	GX 570	GX 660	GX 900	GX 1150	GT 900	GT 1150
Part number reference	10AX016	10AX022	10AX027	10AX032	10AX048	10AX057	10AX066	10AX090	10AX115	10AT090	10AT115
LEs (K)	160	220	270	320	480	570	660	900	1,150	900	1,150
System logic elements (K)	210	288	354	419	629	747	865	1,180	1,506	1,180	1,506
Adaptive logic modules (ALMs)	61,510	83,730	101,620	118,730	181,790	217,080	250,540	339,620	427,200	339,620	427,200
Registers	246,040	334,920	406,480	474,920	727,160	868,320	1,002,160	1,358,480	1,708,800	1,358,480	1,708,800
M20K memory blocks	440	588	750	891	1,438	1,800	2,133	2,423	2,713	2,423	2,713
M20K memory (Mb)	9	11	15	17	28	35	42	47	53	47	53
MLAB memory (Mb)	1.0	1.8	2.4	2.8	4.3	5.0	5.7	9.2	12.7	9.2	12.7
Hardened single-precision floating-point multipliers/adders	156/156	192/192	830/830	985/985	1,368/1,368	1,523/1,523	1,687/1,687	1,518/1,518	1,518/1,518	1,518/1,518	1,518/1,518
18 x 19 multipliers	312	384	1,660	1,970	2,736	3,046	3,374	3,036	3,036	3,036	3,036
Peak fixed-point performance (GMACS) ¹	343	420	1,826	2,167	3,010	3,351	3,714	3,340	3,340	3,340	3,340
Peak floating-point performance (GFLOPS)	140	172	747	887	1,231	1,371	1,519	1,366	1,366	1,366	1,366
Global clock networks	32	32	32	32	32	32	32	32	32	32	32
Regional clocks	8	8	8	8	8	8	16	16	16	16	16
I/O voltage levels supported (V)	1.2, 1.25, 1.35, 1.8, 2.5, 3.0										
I/O standards supported	3 V I/O pins only: 3 V LVTTTL, 2.5 V CMOS										
Maximum LVDS channels (1.6 G)	120	120	168	168	222	324	270	384	384	312	312
Maximum user I/O pins	288	288	384	384	492	696	696	768	768	624	624
Transceiver count (17.4 Gbps)	12	12	24	24	36	48	48	96	96	72	72
Transceiver count (25.78 Gbps)	-	-	-	-	-	-	-	-	-	6	6
PCI Express hardened IP blocks (3.0 x8)	1	1	2	2	2	2	2	4	4	4	4
Maximum 3 V I/O pins	48	48	48	48	48	48	48	-	-	-	-
Memory devices supported	DDR4, DDR3, DDR2, QDR IV, QDR II+, QDR II+ Xtreme, LPDDR3, LPDDR2, RLD RAM 3, RLD RAM II, LLD RAM II, HMC										