24th IEEE REAL TIME CONFERENCE
Quy Nhon, Vietnam

Optimization of the Upgraded Timing Distribution System of the LHCb experiment at CERN

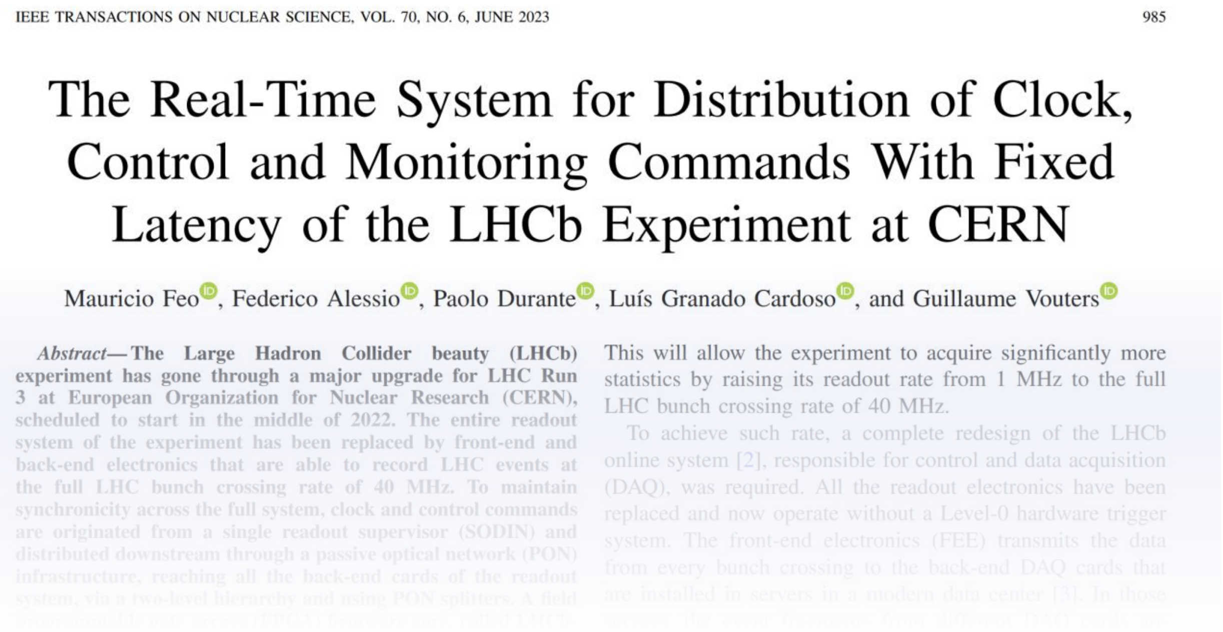
Mauricio Féo – TU Dortmund / CERN

Previously on Real Time Conference… (2022)

Real-Time System for Distribution of Clock and Control Commands with Fixed Latency

Maurício Féo – CERN
on behalf of the LHCb Online team

23rd Real Time Conference



IEEE TRANSACTIONS ON NUCLEAR SCIENCE, VOL. 70, NO. 6, JUNE 2023 985

The Real-Time System for Distribution of Clock, Control and Monitoring Commands With Fixed Latency of the LHCb Experiment at CERN

Mauricio Feo, Federico Alessio, Paolo Durante, Luís Granado Cardoso, and Guillaume Vouters

Abstract—**The Large Hadron Collider beauty (LHCb) experiment has gone through a major upgrade for LHC Run 3 at European Organization for Nuclear Research (CERN), scheduled to start in the middle of 2022. The entire readout system of the experiment has been replaced by front-end and back-end electronics that are able to record LHC events at the full LHC bunch crossing rate of 40 MHz. To maintain synchronicity across the full system, clock and control commands are originated from a single readout supervisor (SODIN) and distributed downstream through a passive optical network (PON) infrastructure, reaching all the back-end cards of the readout system, via a two-level hierarchy and using PON splitters. A field programmable gate array (FPGA) firmware core, called LHCb-**

This will allow the experiment to acquire significantly more statistics by raising its readout rate from 1 MHz to the full LHC bunch crossing rate of 40 MHz.

To achieve such rate, a complete redesign of the LHCb online system [2], responsible for control and data acquisition (DAQ), was required. All the readout electronics have been replaced and now operate without a Level-0 hardware trigger system. The front-end electronics (FEE) transmits the data from every bunch crossing to the back-end DAQ cards that are installed in servers in a modern data center [3]. In those servers, the event fragments from different DAQ cards are

https://indico.cern.ch/event/1109460/contributions/4893298/

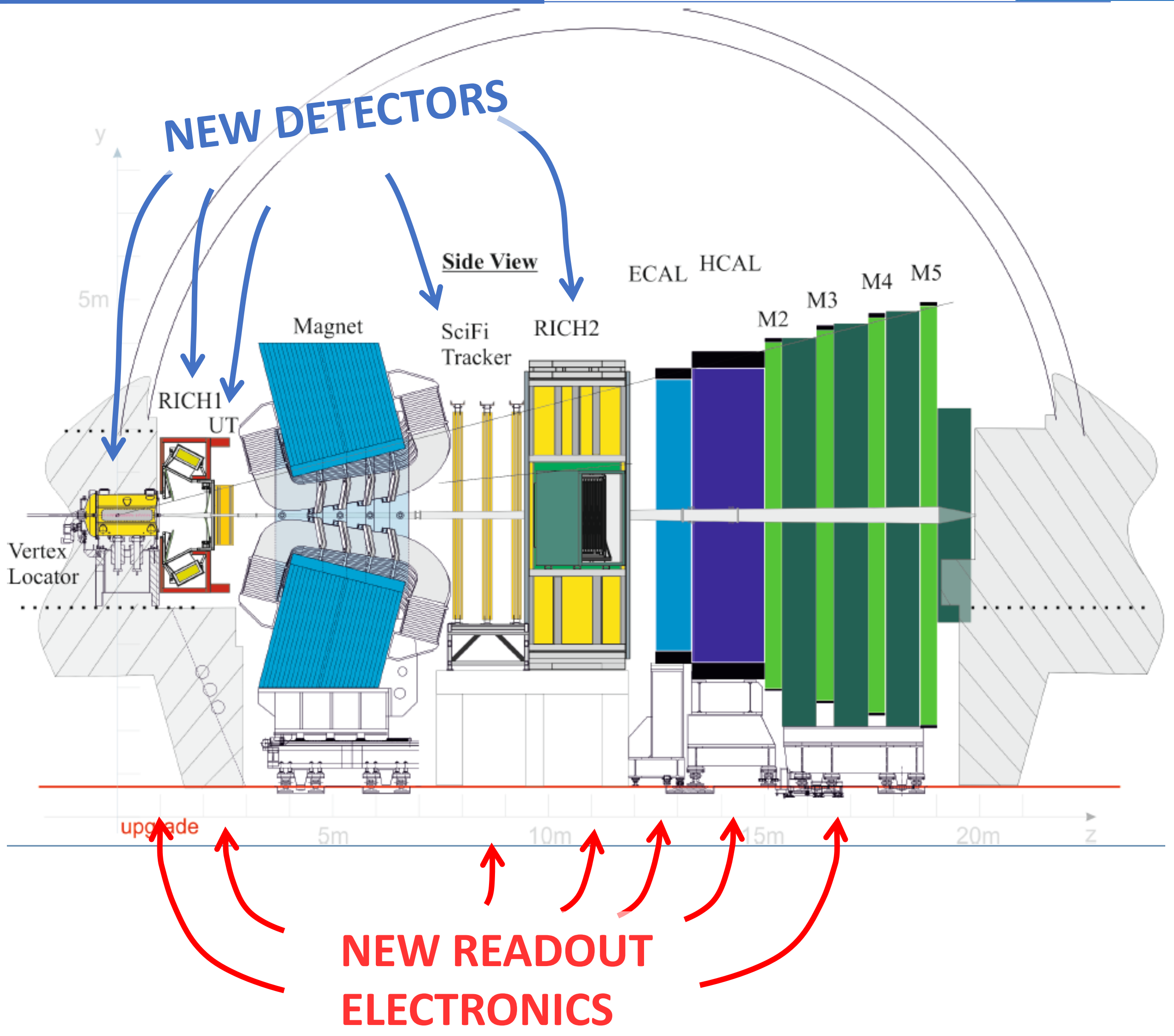
https://ieeexplore.ieee.org/document/10115510

The LHCb Upgrade

When:	LHC Long Shutdown 2 (by 2022) For Runs 3 & 4 (2022 – 2029)
Why:	To increase statistics 9 fb^{-1} (Runs 1-2) → 50 fb^{-1} (Runs 1-4)
How:	Increasing instant. luminosity 5x higher → $L_{inst} = 2x10^{33}$ cm^{-2}s^{-1} Increasing readout rate **1 MHz → 40 MHz**

This requires some **main changes:**

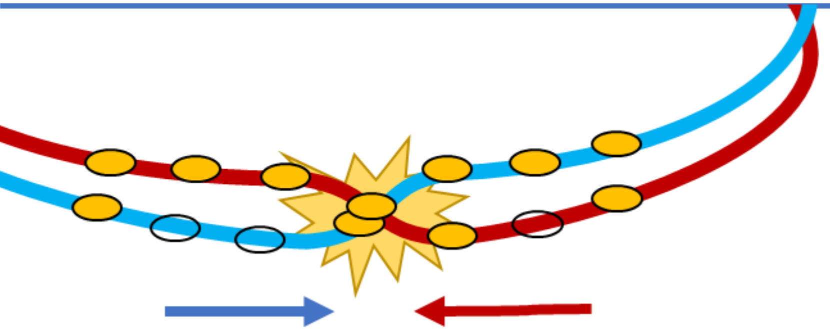
- Replace many sub-detectors:
  - New Tracking System
    (VeLo, UT, SciFi)
  - Partially new Particle ID System
    (RICH1 + RICH2)
- Replace of ALL the electronics:
  - No more hardware trigger
  - Event selection in software
  - **Completely new DAQ system**

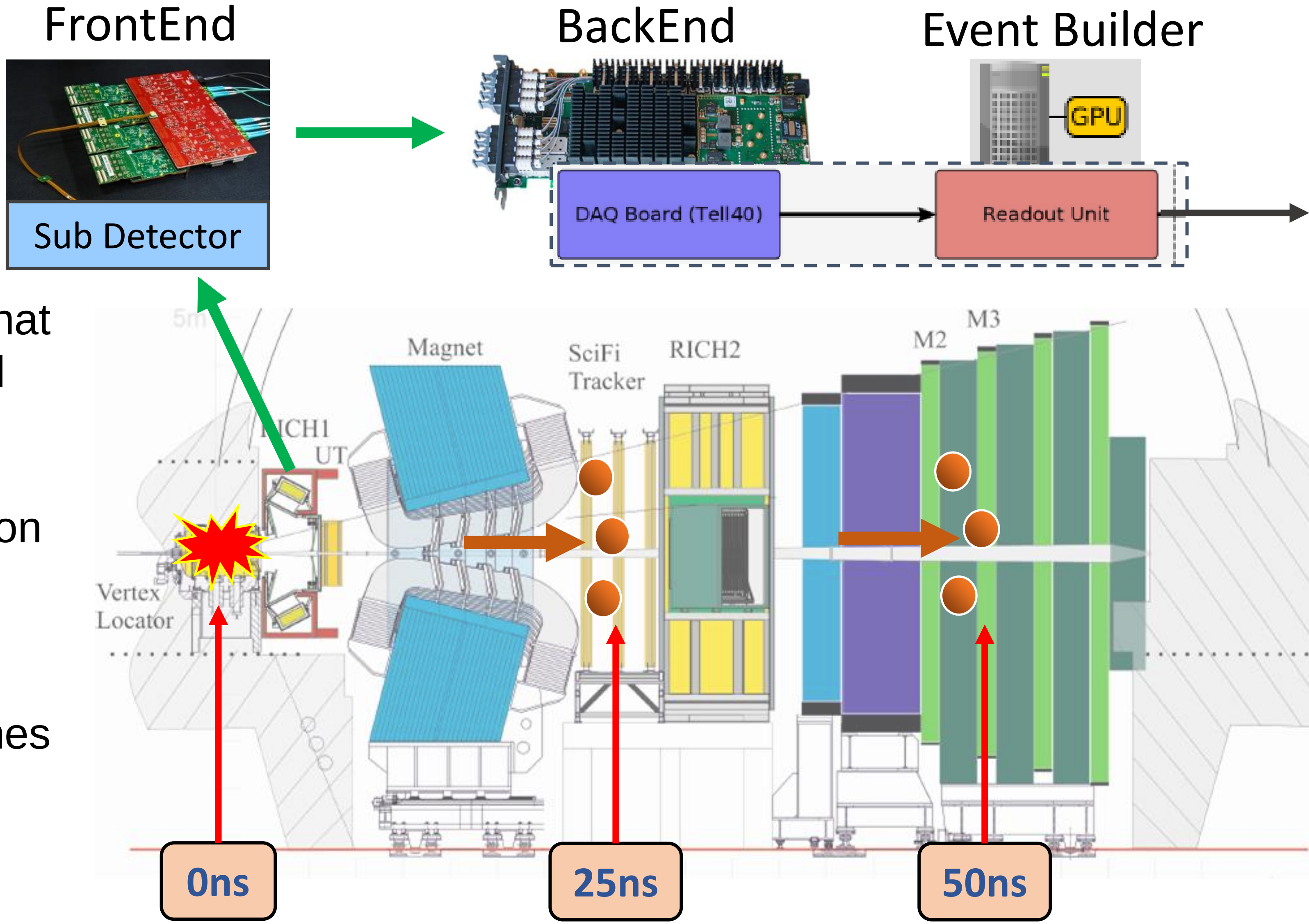

NEW DETECTORS

NEW READOUT ELECTRONICS

From Collision to Memory

The data path
before the software





- LHC beams are divided in bunches that cross in synchrony with a clock signal

  ~40MHz = 25ns / cross

- Not all bunches are filled → no collision

  We need to select which bunch crossings to save

- Particles arrive at different times depending on the subdetector

  We need to phase-align the clocks per subdetector and have fixed latency

The TFC System: A Real-Time Architecture

- Generates Timing and Fast Commands
- Provides clock with fixed and deterministic latency
- Distributes them to all the detector electronics
  - BackEnd cards via TTC-PON (Project link)
  - FrontEnd modules via GBT (Project link)

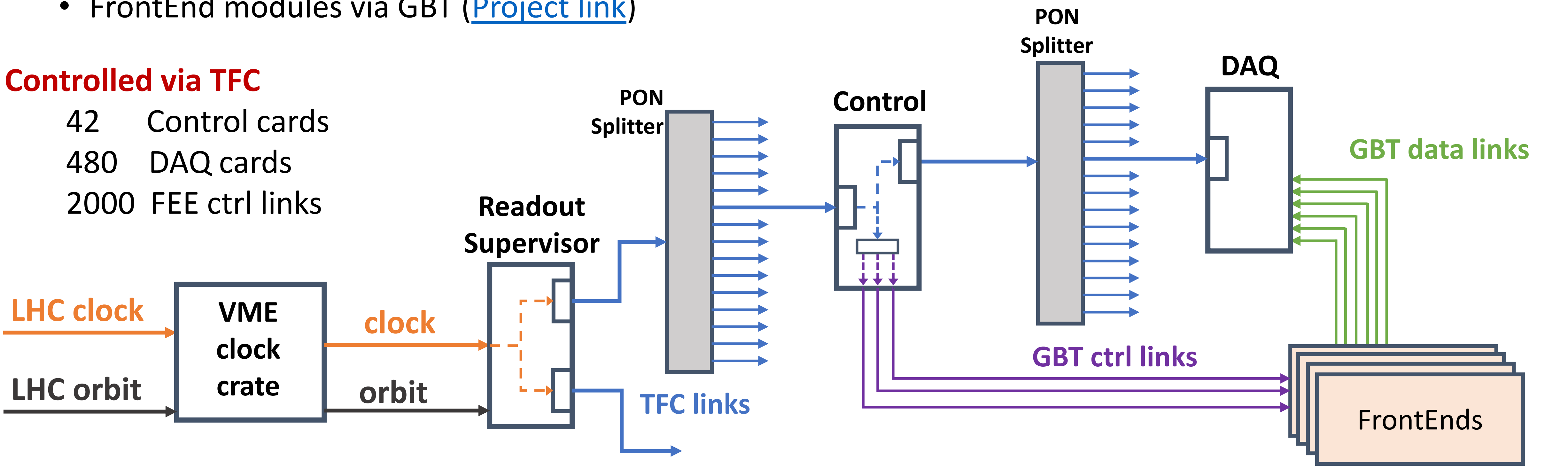
All cards are PCIe40s with different firmwares

Supervisor = SODIN
Control = SOL40
DAQ = TELL40

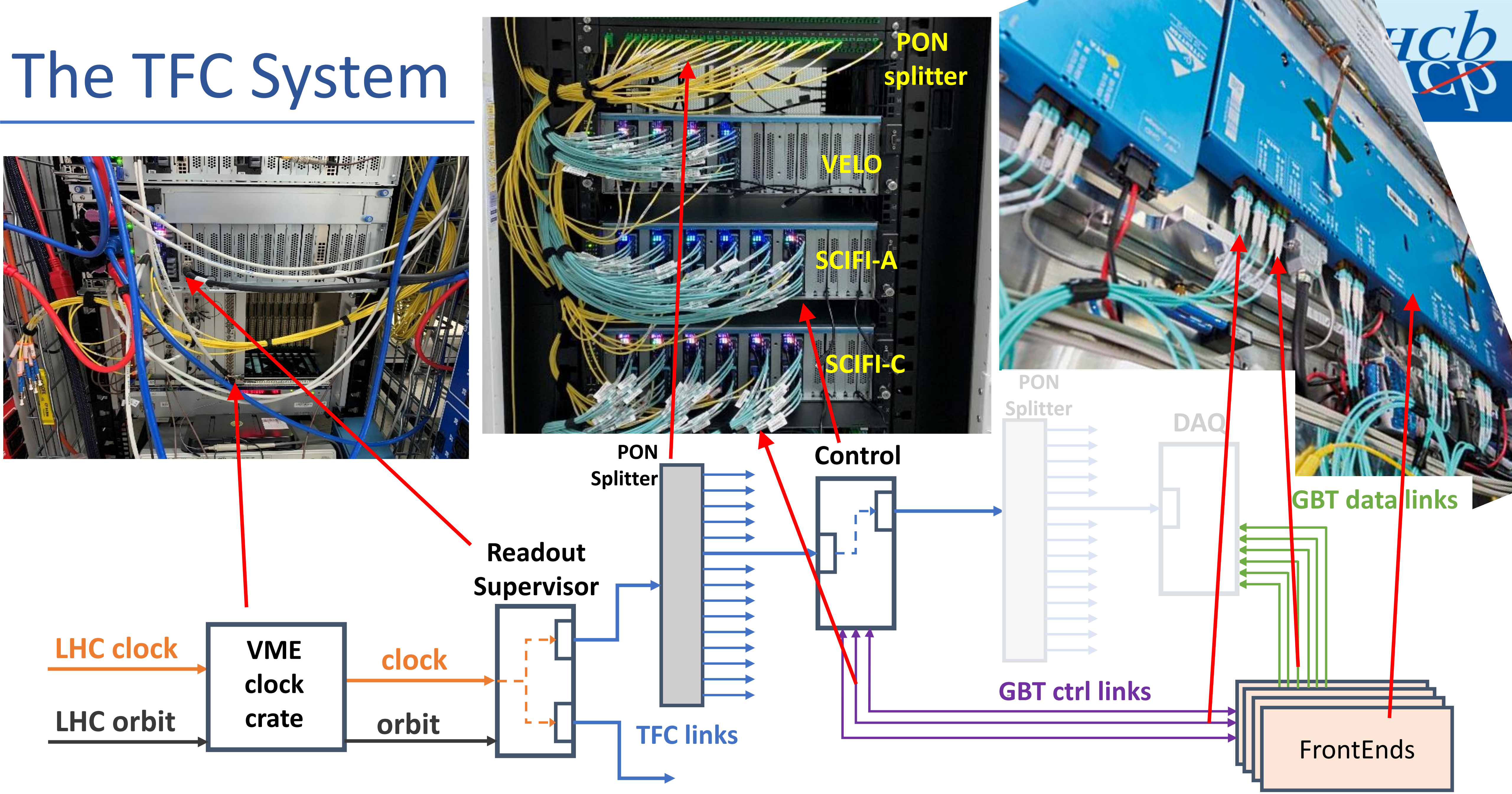
LHC clock
TFC links
GBT ctrl links
GBT data links

Controlled via TFC

42	Control cards
480	DAQ cards
2000	FEE ctrl links



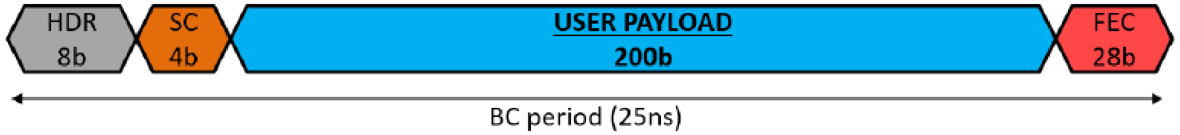
The TFC System



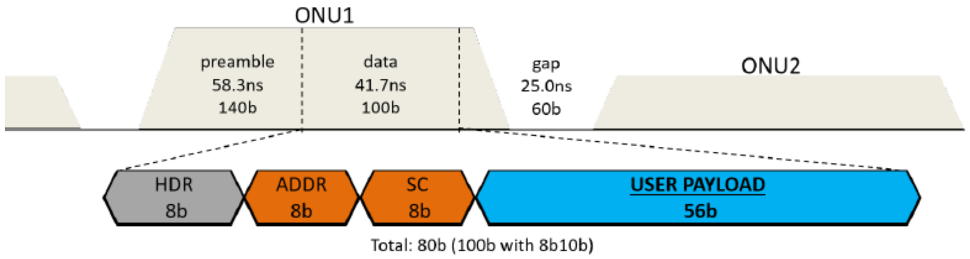
The TTC-PON Project

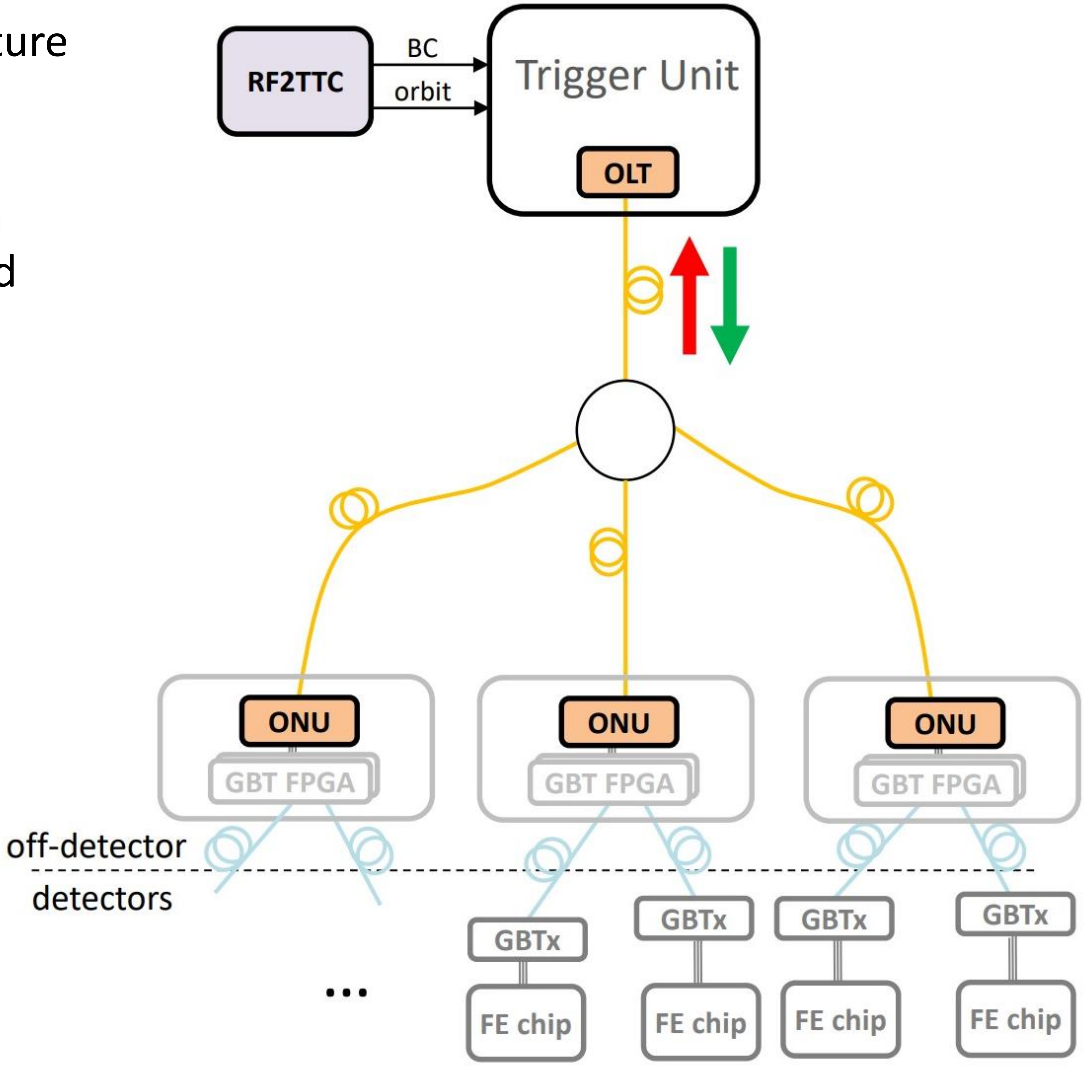
- Timing, Trigger and Control for Passive Optical Networks
  (Link to project)
- FPGA-Based System for TFC distribution with fixed latency
- 9.6 Gbps downstream with FEC
  (8.0 for the user)

- Master/Follower architecture
  - OLT = Master
  - ONU = Follower
- Slow Control Implemented
  - Everything controllable from the master



- 2.4 Gbps upstream 8b10b (Time Division Multiplexing)





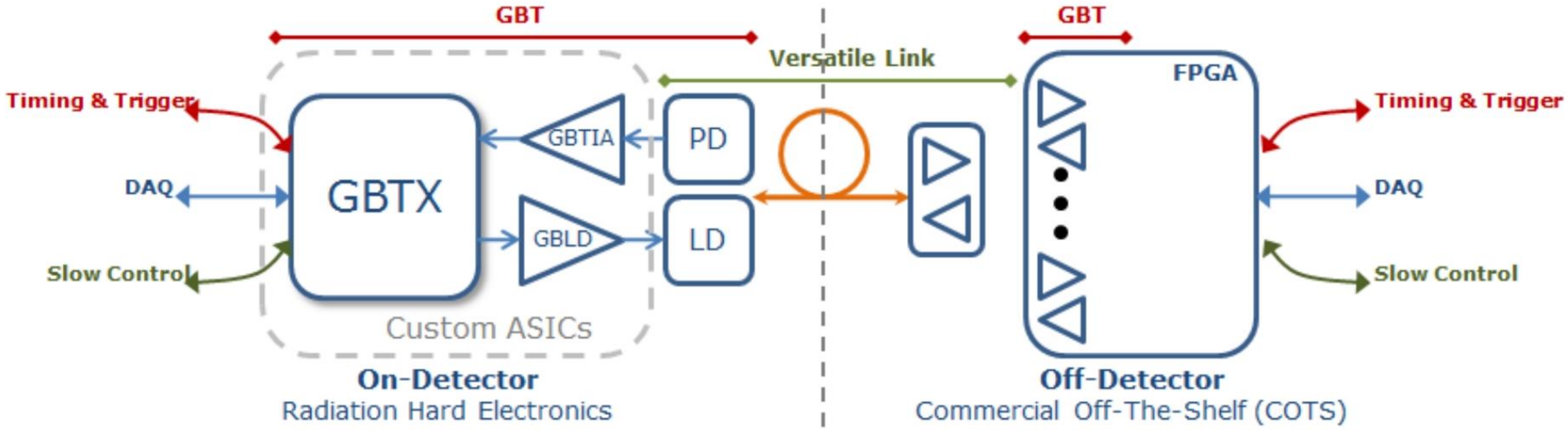
OLT: Optical Line Terminal / ONU: Optical Network Unit

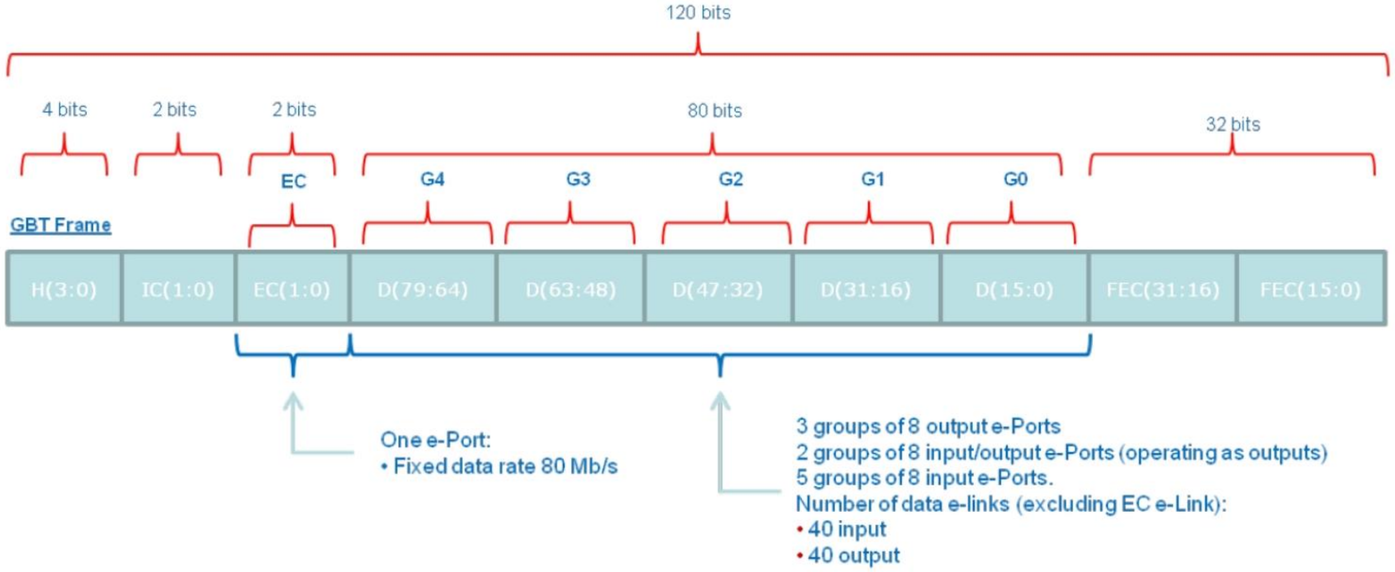
GBT Project: The GigaBit Transceiver

- The GBT Project provides a radiation hard chipset for handling control and data acquisition on frontend boards

- Designed at CERN, it is widely used on the upgrade of its experiments

- It also provides a firmware component (GBT-FPGA) that allows FPGAs to interface directly with the GBTx chip.





One e-Port:
- Fixed data rate 80 Mb/s

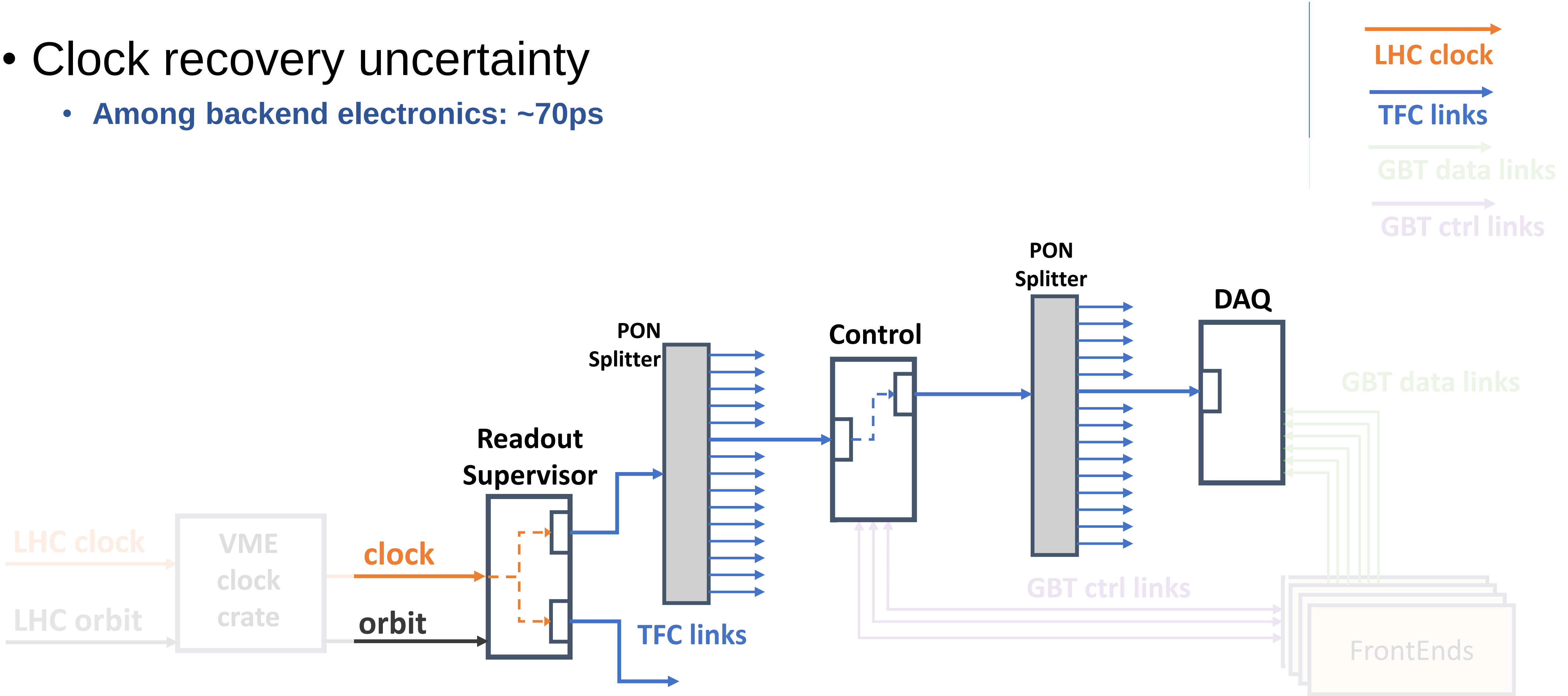
3 groups of 8 output e-Ports
2 groups of 8 input/output e-Ports (operating as outputs)
5 groups of 8 input e-Ports.
Number of data e-links (excluding EC e-Link):
- 40 input
- 40 output

For more info:

GBT Project link (requires CERN login)
The GBT-SerDes ASIC prototype (public paper)

Optimization of the TFC System

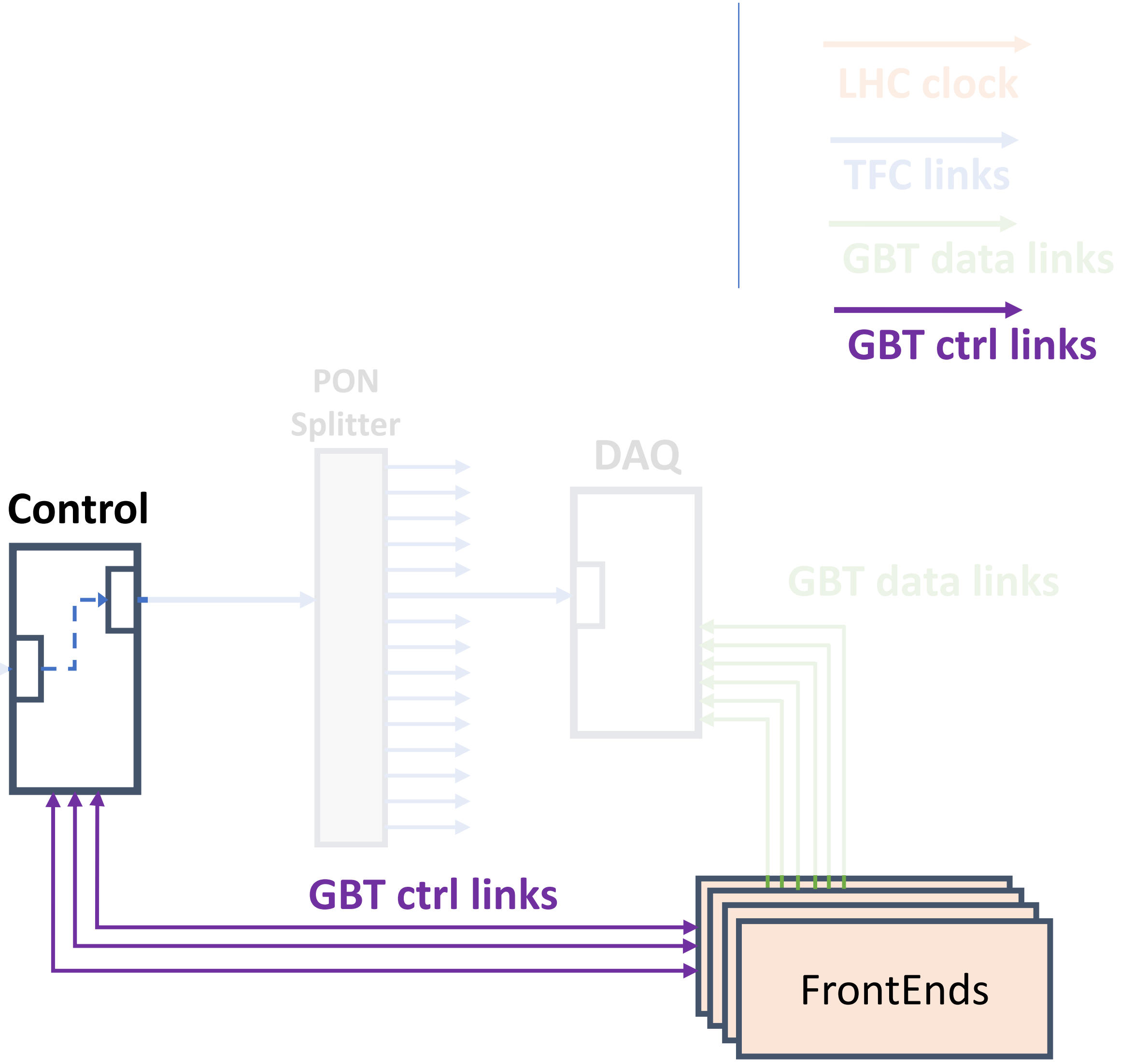
- Clock recovery uncertainty
  - **Among backend electronics: ~70ps**



Optimization of the TFC System

- Clock recovery uncertainty
  - **Among backend electronics: ~70ps**
  - **Between backend and frontend: ~4ns**

- The optimization aims at **reducing the timing uncertainty** between the Control cards and the Frontend electronics by an order of magnitude
  - For all ~2000 FEE GBT ctrl links
  - From **4ns** to less than **~500ps**

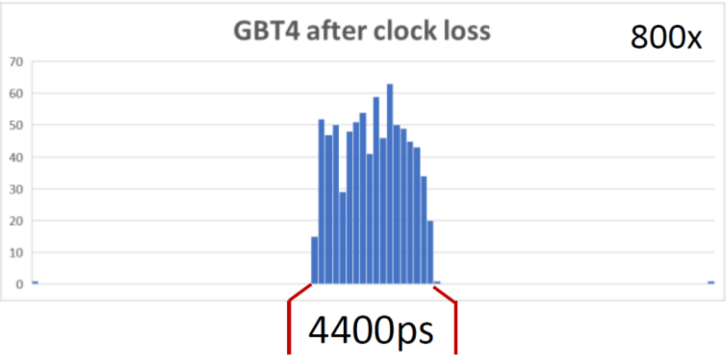


Control Card: Previous Clock Architecture

- But why 4ns?

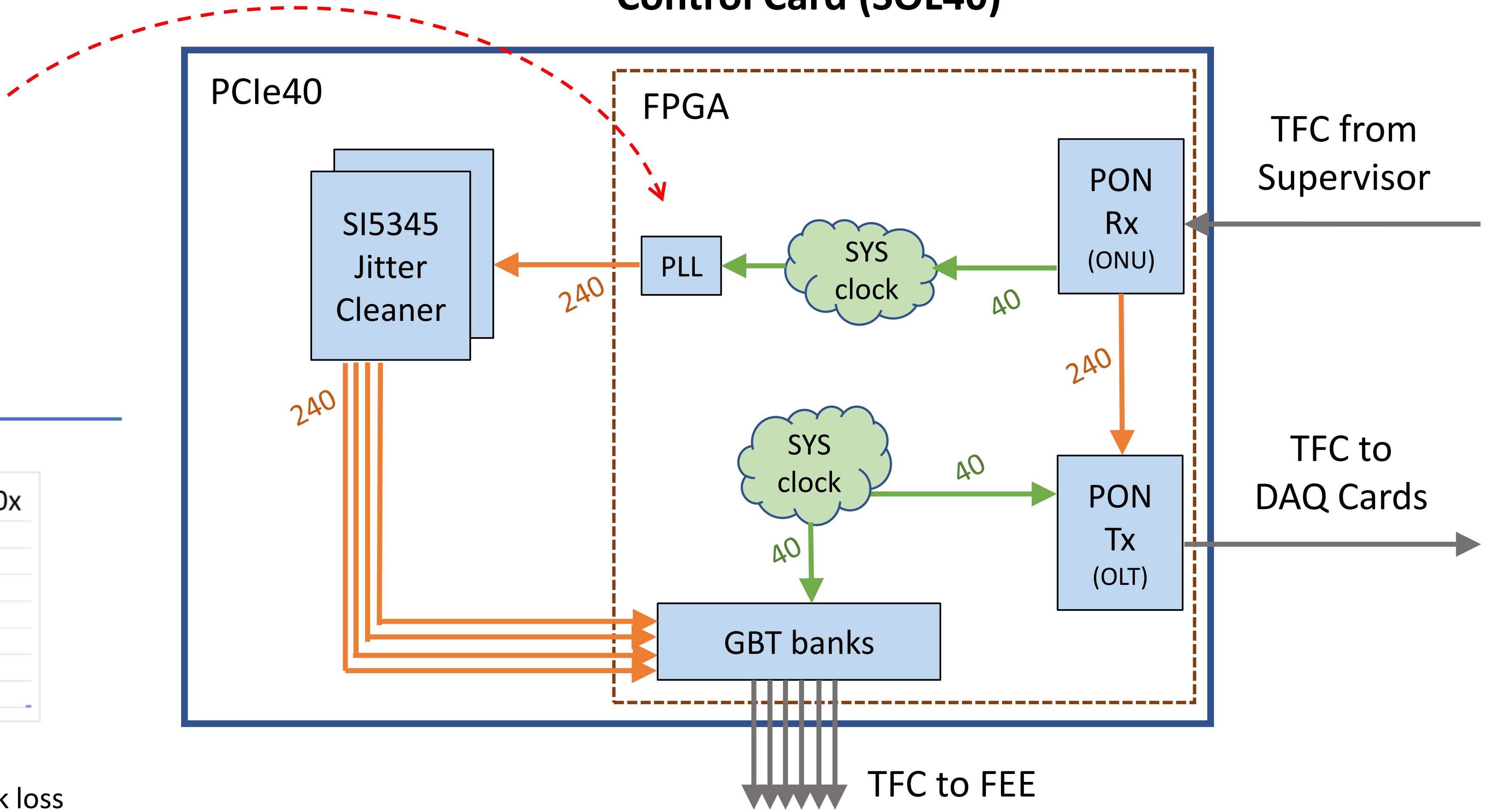
The FPGA internal PLL locks at a random phase after a clock loss

Clock period = **4.16ns**



Recovered clock phase at the FEE after a clock loss

Control Card (SOL40)

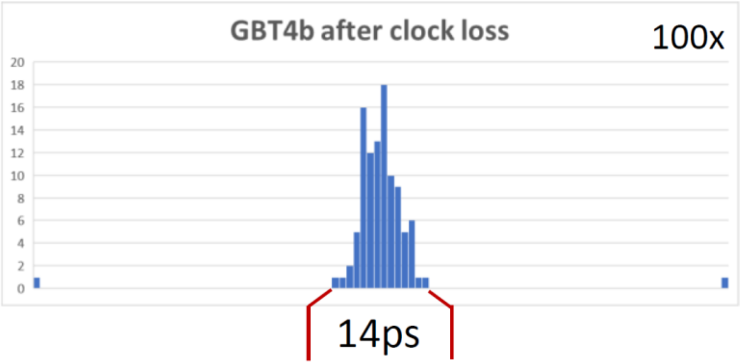


Alternative #1: keeping all clocks inside

AGAINST RECOMMENDATIONS!!

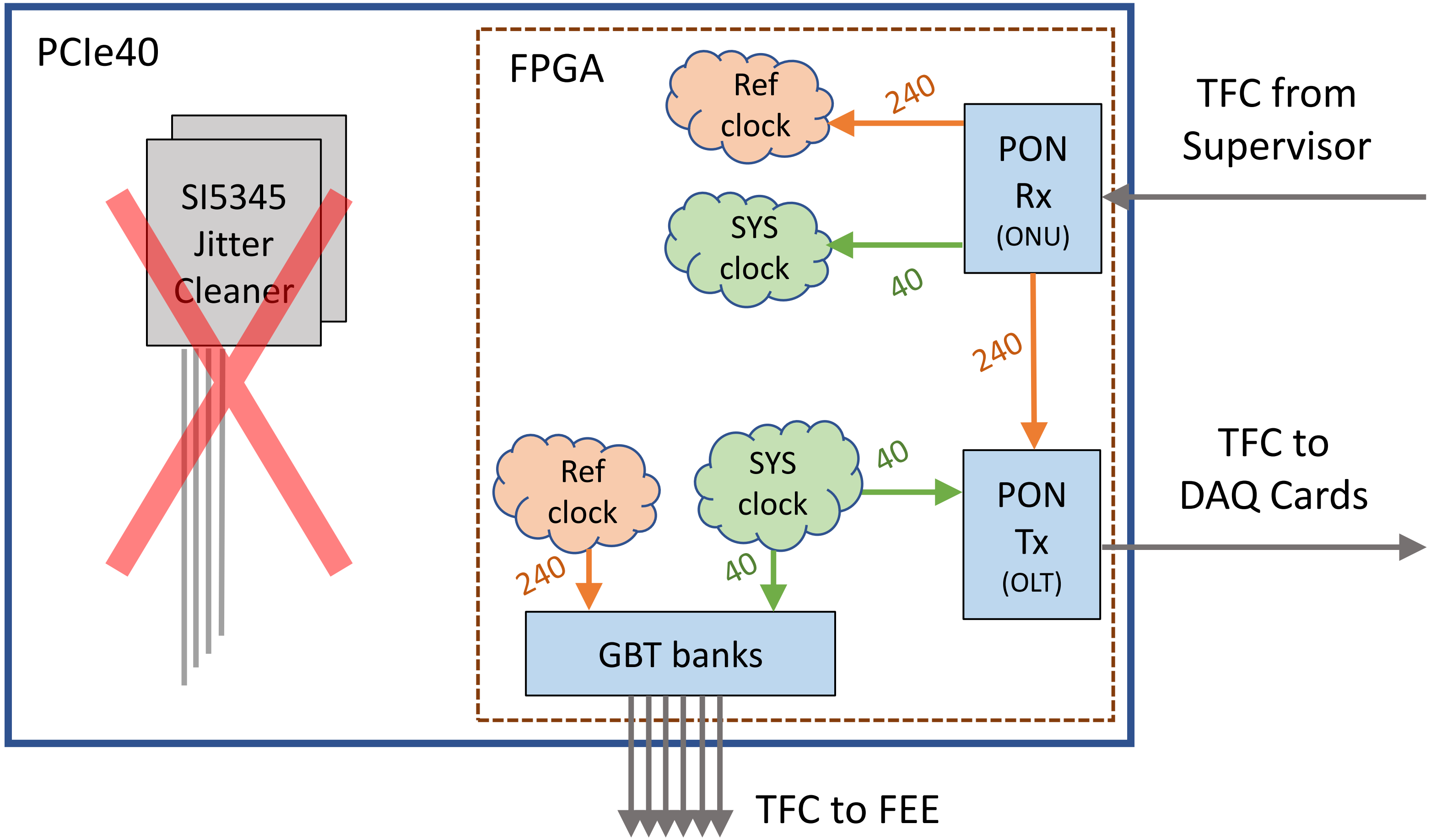
Inject the 240MHz CDR clock into the global clock network and use it for all GBT XCVR banks

- Phase uncertainty falls to **~15ps** !!
- Poor clock quality causes some FEE links to lose lock eventually
- Metastability on the CDC 40→240 at the GBT banks



Recovered clock phase at the FEE after a clock loss

Control Card (SOL40)

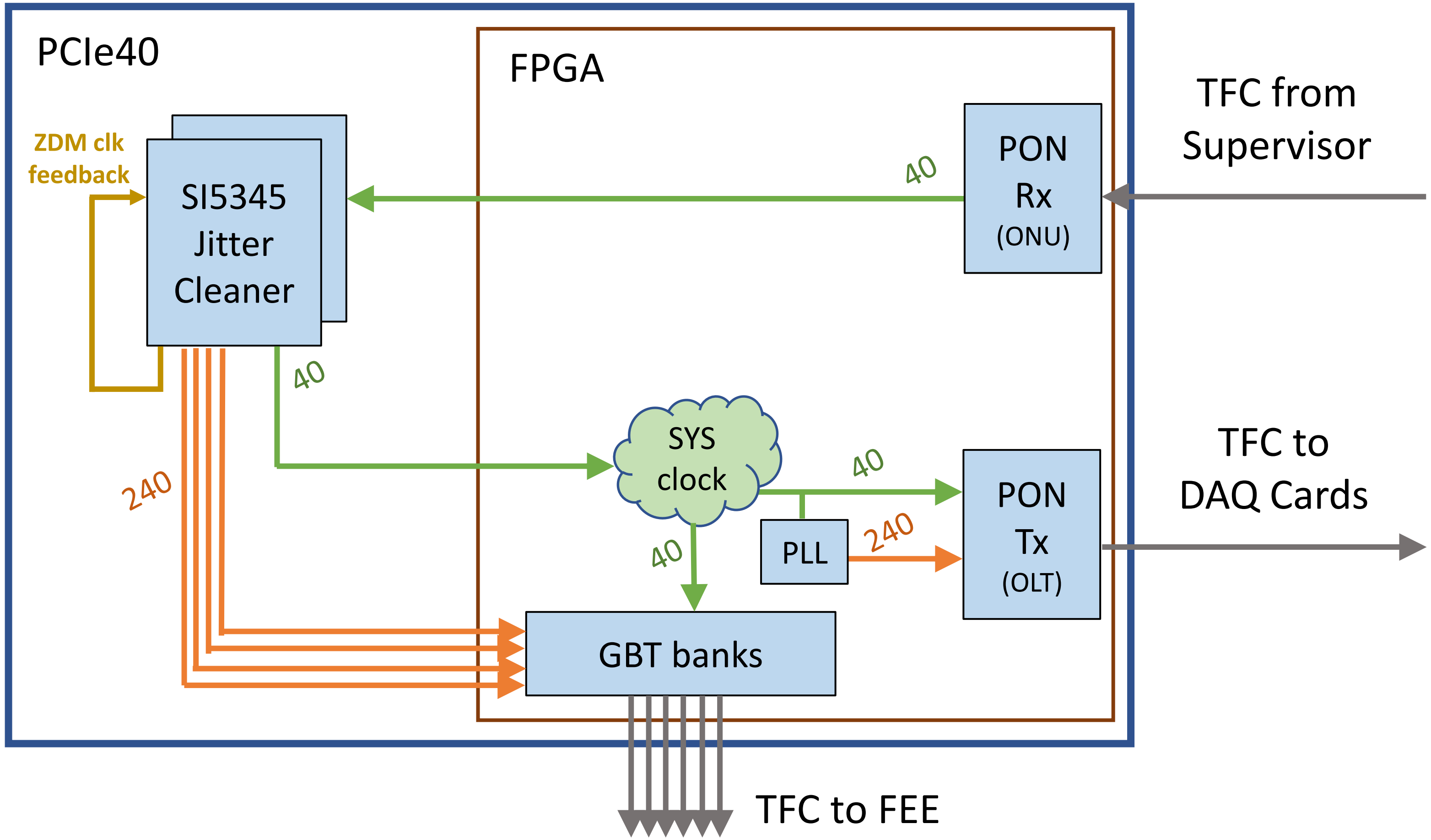


Alternative #2: PLLs in Zero-Delay Mode

- Configure the SI5345s PLLs to receive 40MHz in ZDM mode.
- Borrow an unused clean 40MHz output to be used as SYS clock.
- Carefully apply timing constraints

- Excellent phase uncertainty:
  - **~15ps** after clock loss
  - **~30ps** after FPGA reprogramming
- TIE jitter significantly reduced across the whole system
- Worked perfectly!
  - until it did not… ☹

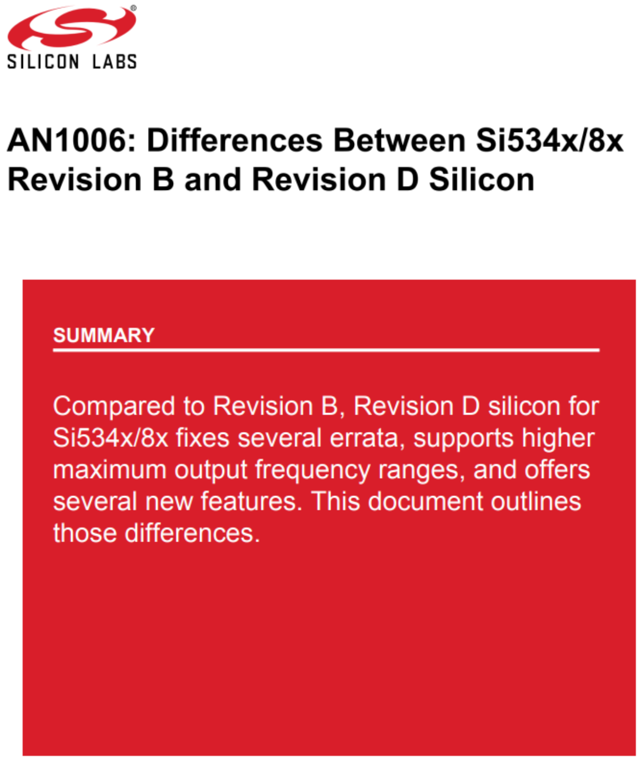
Control Card (SOL40)



Alternative #2: The Problem

Si5345 Rev B presenting weird symptoms…

- The ZDM option worked well across LHCb until a major test where the FE control links of a specific subdetector would not work anymore.
- FE links losing lock and other symptoms of too many bit errors
- It happens after configuring the Si5345 in ZDM mode.
- Only in the PCIe40s of a specific server. (identical to working ones)
- The PLL remains responsive. You can reset, reconfigure, etc. It appears normal but it will only work again after a full power cycle.
- Due to lack of time and difficulty in debugging, the idea was put on hold.



SILICON LABS

AN1006: Differences Between Si534x/8x Revision B and Revision D Silicon

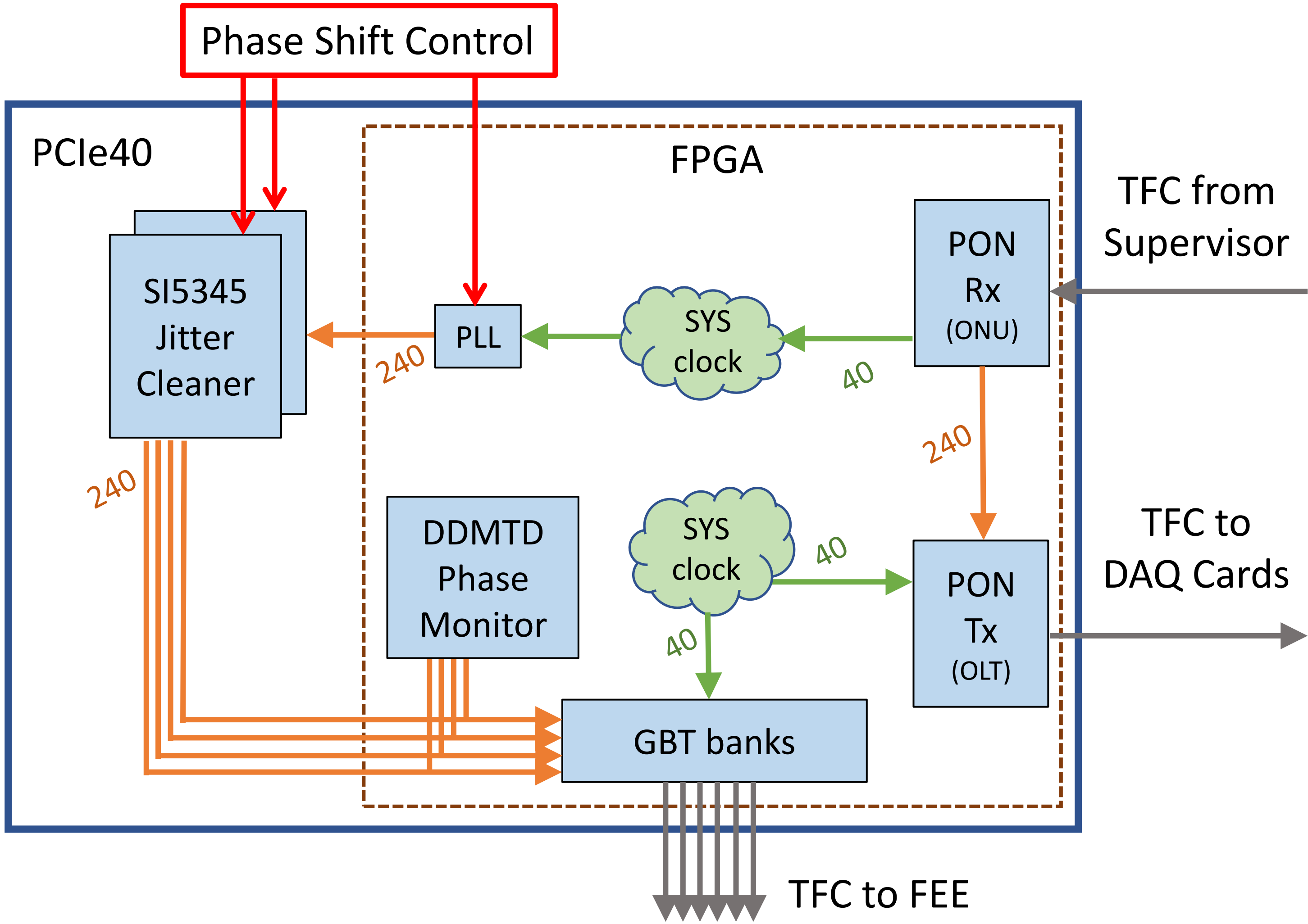
SUMMARY

Compared to Revision B, Revision D silicon for Si534x/8x fixes several errata, supports higher maximum output frequency ranges, and offers several new features. This document outlines those differences.

Alternative #3: Measure and Shift the Clocks

- We measure the clock phases using a DDMTD* implemented in firmware and shift the PLL clocks to a specific setpoint.
- The setpoint is the middle of the stability windows found when scanning the phase space.
- The phase measurement needs to compensate for internal delays in the FPGA.
  - Timing reports need to be generated for every compilation.
- Complex operation when compared to previous alternatives.

Phase shift resolution	
Internal PLL	104ps
Si5345	72ps



*DDMTD = Digital Dual Mixer Time Difference

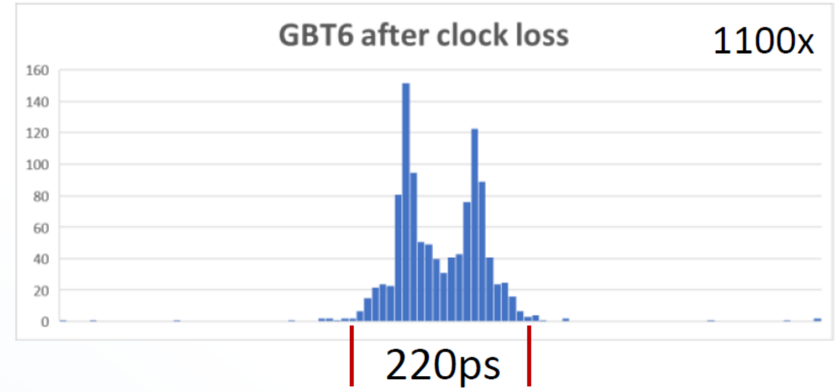
Alternative #3: Measure and Shift the Clocks

- Phase uncertainty within the experiment requirements

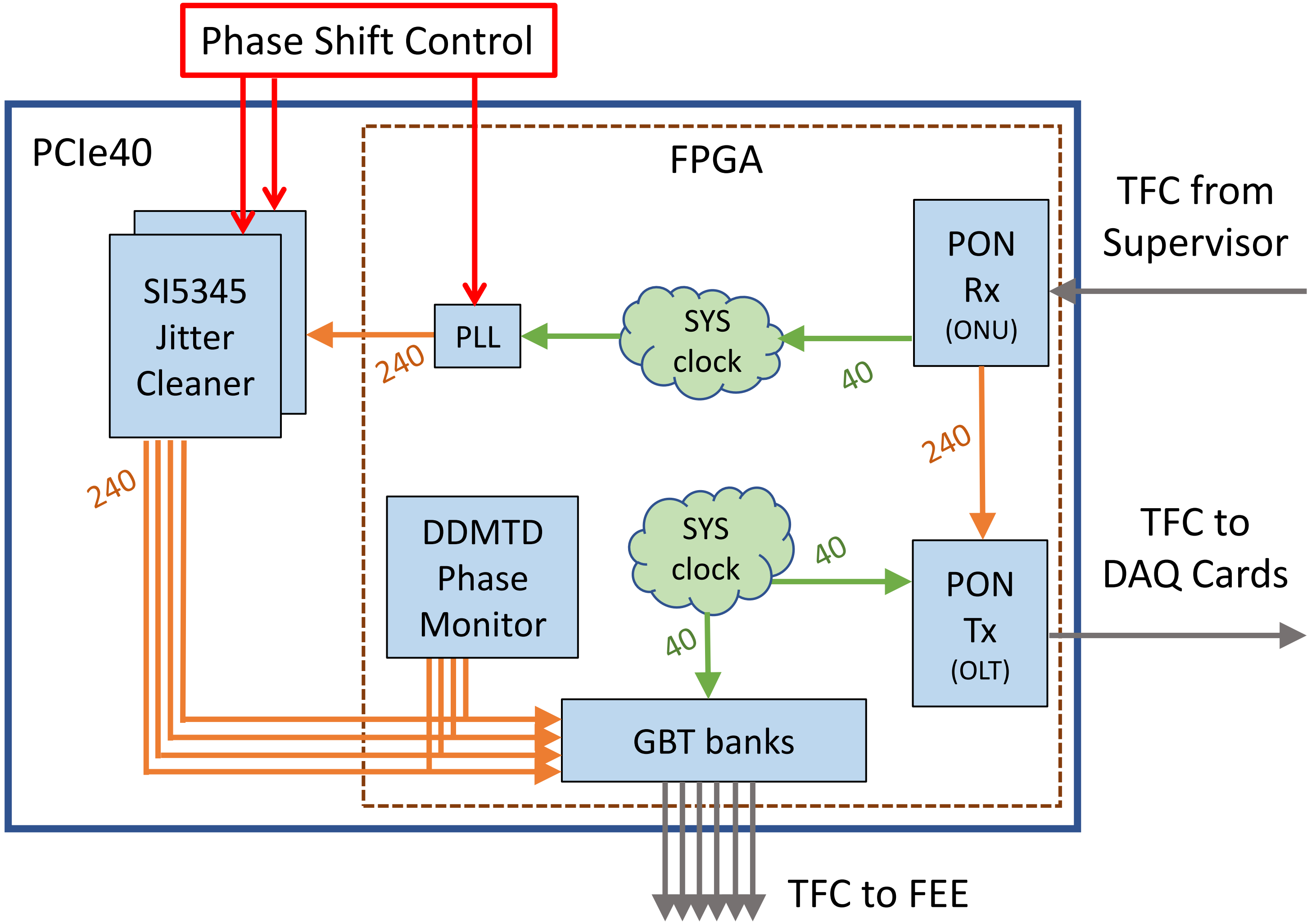
- FE control links stable and reliable

- A lot of effort to automatize the phase shifting mechanism but it works

- **Currently in use at LHCb**



Recovered clock phase at the FEE after a clock loss



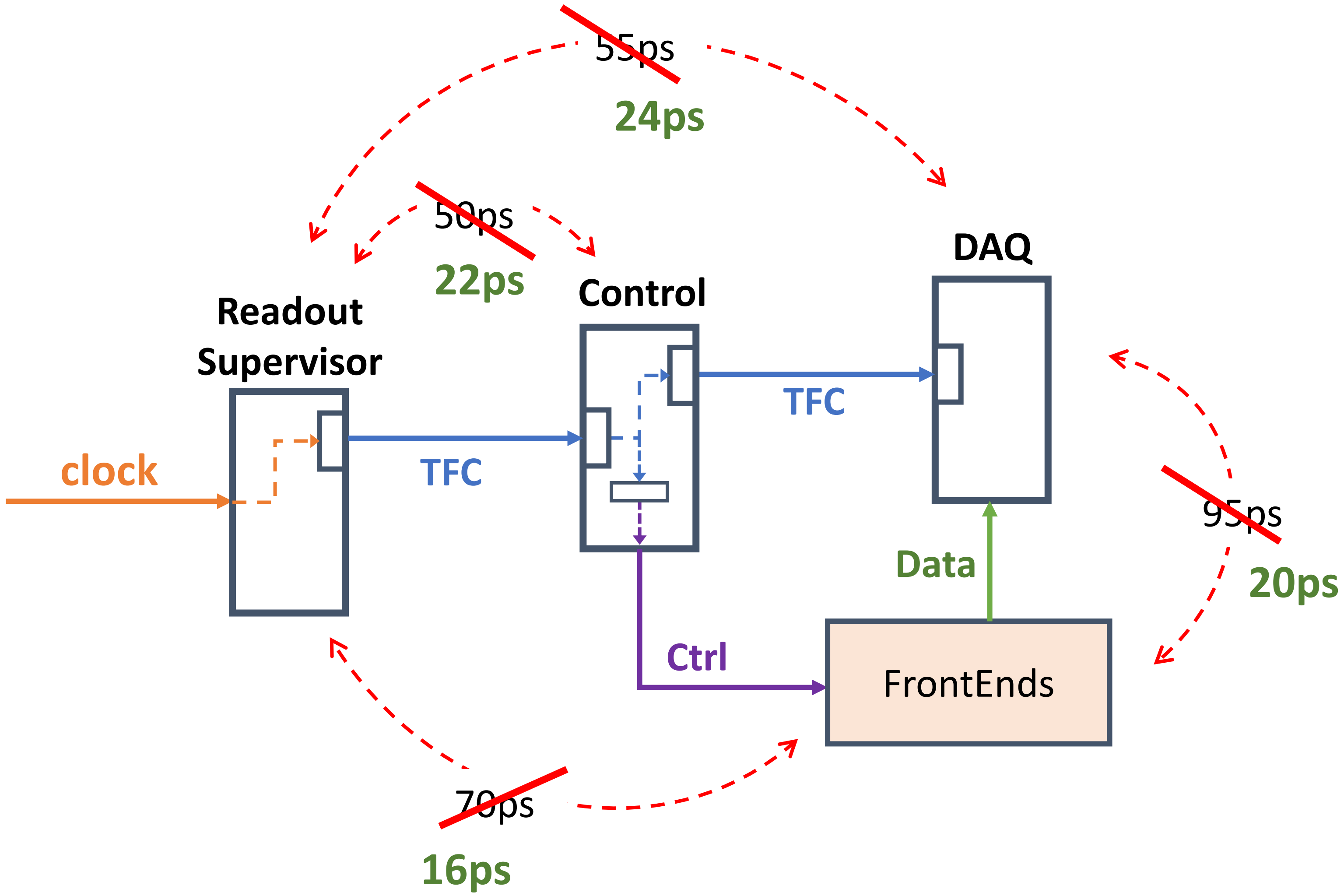
Conclusion

- Initially, any clock loss would cause the LHCb subdetectors to lose time alignment due to a phase change in the clock propagated to the FEE.
- Different implementations were studied to provide fixed latency and deterministic phase after the event of a clock loss.
- Alternative #2 (PLLs in ZDM mode) showed the best results, however, a mysterious problem when configuring the PLLs does not allowed us to use it with all subdetectors.
- Alternative #3 (clock phase shifting) proved to work reliably and has been chosen as the solution.
- Since then, the LHCb subdetectors are able to keep proper time alignment even after the eventual loss of clock or similar problem.

Obrigado!

Questions are welcome :)

Alternative #2: TIE Jitter


~~55ps~~
24ps
~~50ps~~
22ps
DAQ
Control
Readout
Supervisor
TFC
TFC
clock
~~95ps~~
20ps
Data
Ctrl
FrontEnds
~~70ps~~
16ps