The PCIe-based readout system for the LHCb experiment

K.Arnaud, J.P. Cachemiche, P.-Y. Duval,
F. Hachon, M. Jevaud, R. Le Gac, **F. Rethore**
Centre de Physique des Particules de Marseille

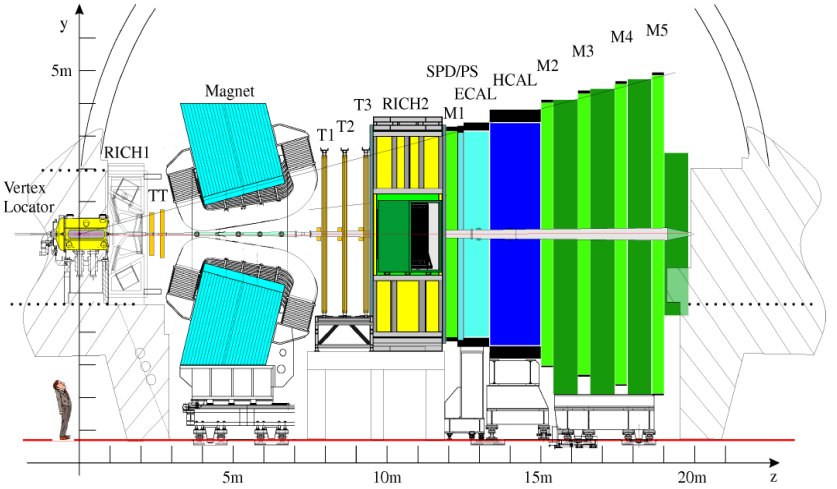
Outline

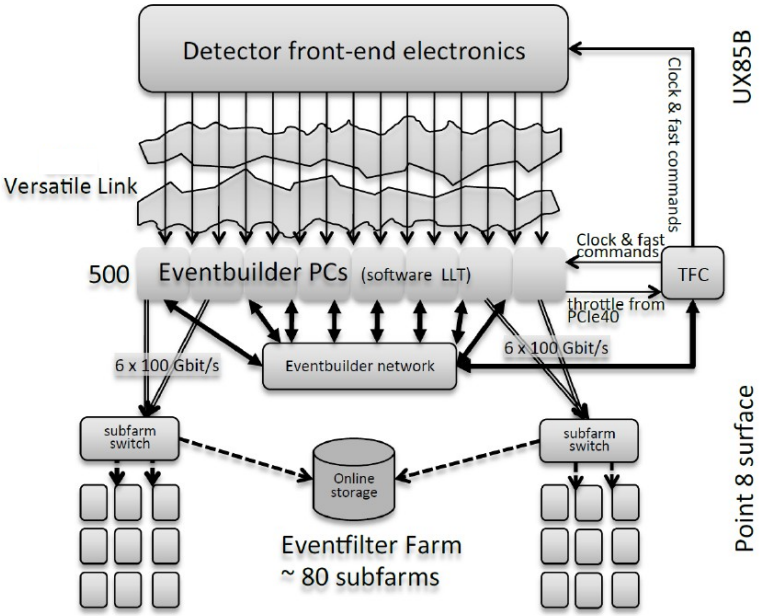
- Triggerless readout
- New readout architecture
- Current prototype PCIe40
- Overall architecture

General architecture

5 main blocks :

- Front End electronics
- Eventbuilder PCs
- TFC system
- Online network
- High Level Trigger PCs Farm

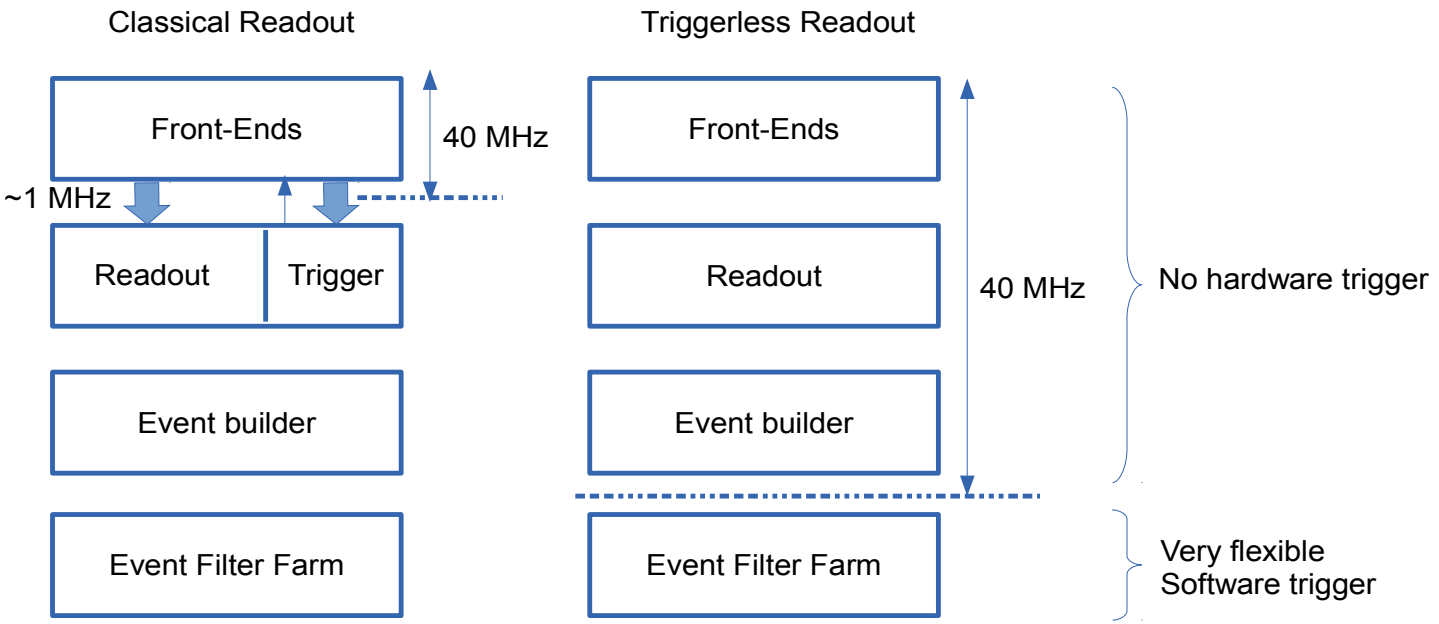




LHCb Readout system

Triggerless readout

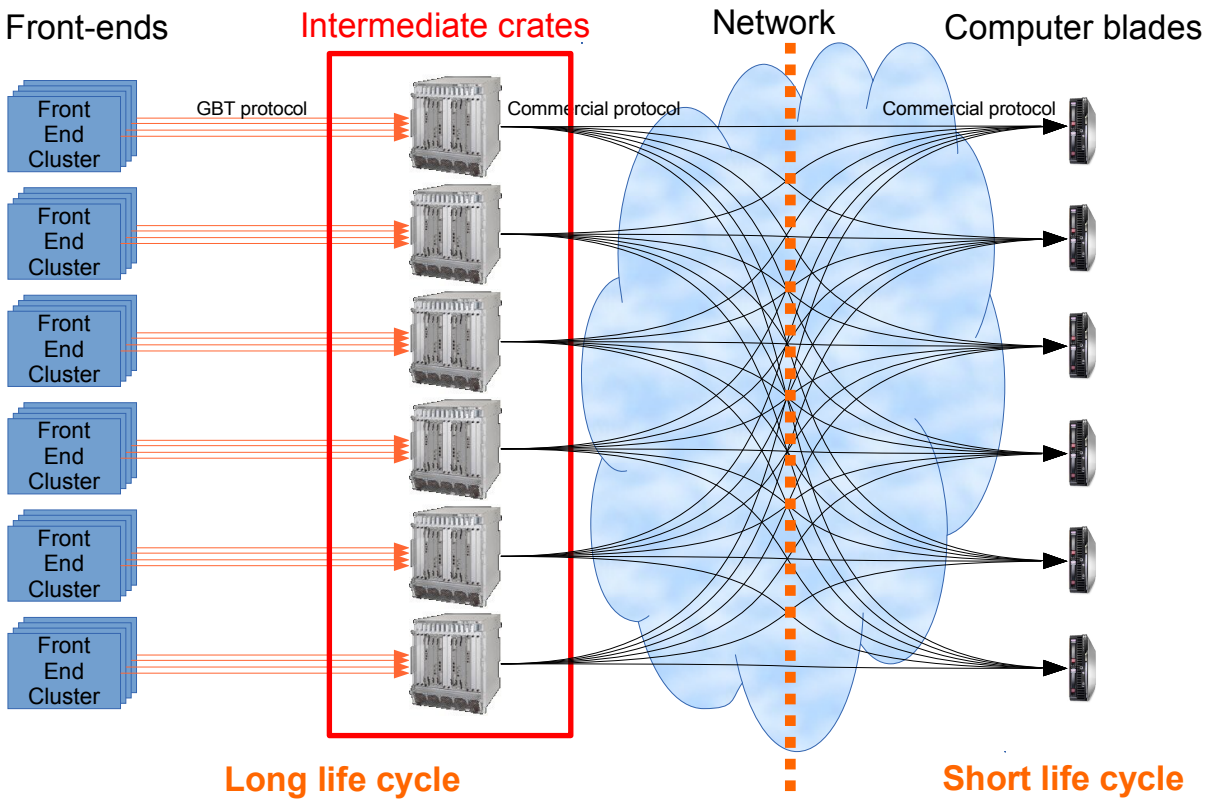
Does not mean there is no trigger



- CPU processing power makes feasible computation of trigger by software :
  - Previous systems were able to process events at 1 Mhz
  - Software trigger implies to acquire data at 40 MHz and therefore « only » 40 times more processing power
    - Much more flexible

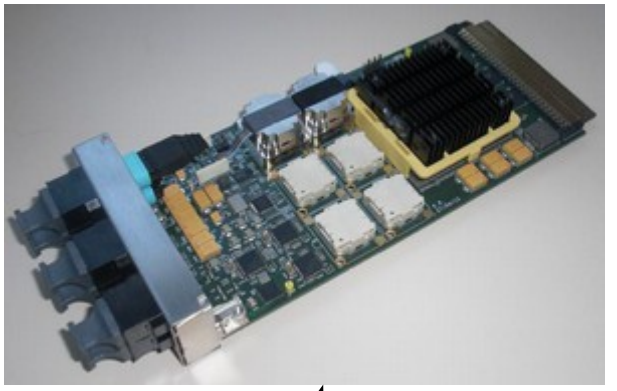
Initial architecture choice for LHCb

Triggerless: All events fragments are routed toward a single CPU blade



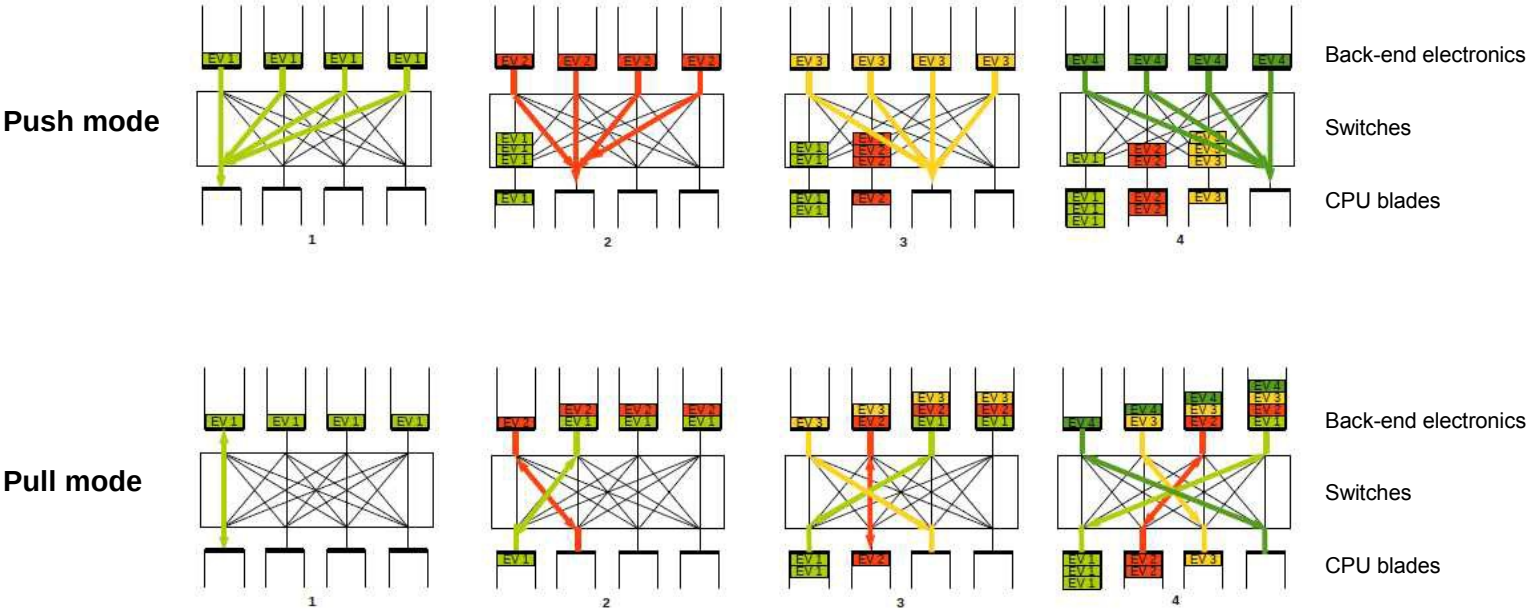
ATCA based

Memory issue



2 methods to make data converge toward a single CPU:
both require memory

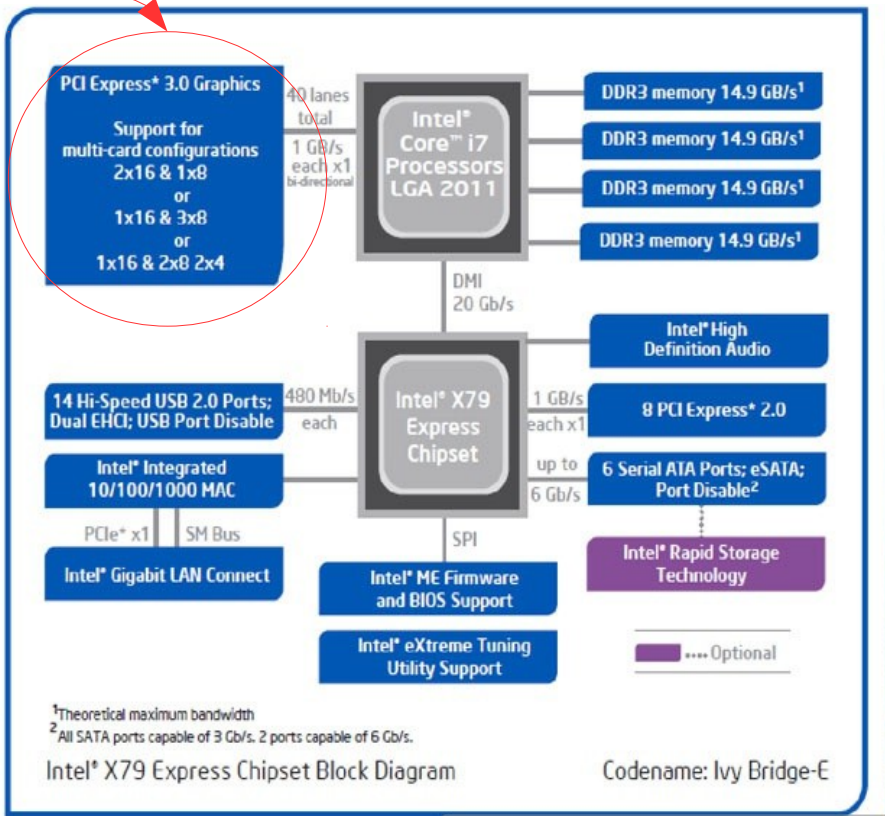
- **Push** : requires expensive switches with **memory** inside
- **Pull** : requires **memory** on already very dense back-end boards



Major technological breaktrough

New architecture of Intel chips includes very large bandwidth data paths

- 40 lanes PCIe GEN3 at 8 Gbits/s
- Independent paths to memory

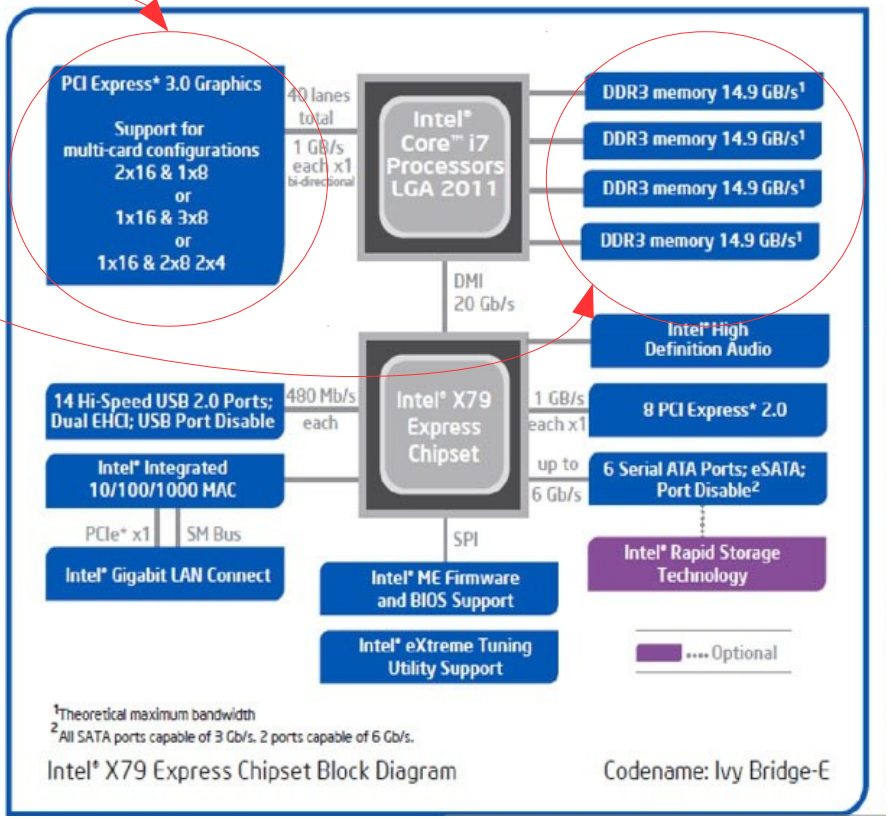


Ivy Bridge architecture

Major technological breaktrough

New architecture of Intel chips includes very large bandwidth data paths

- 40 lanes PCIe GEN3 at 8 Gbits/s
- Independent paths to memory
- Large memory available

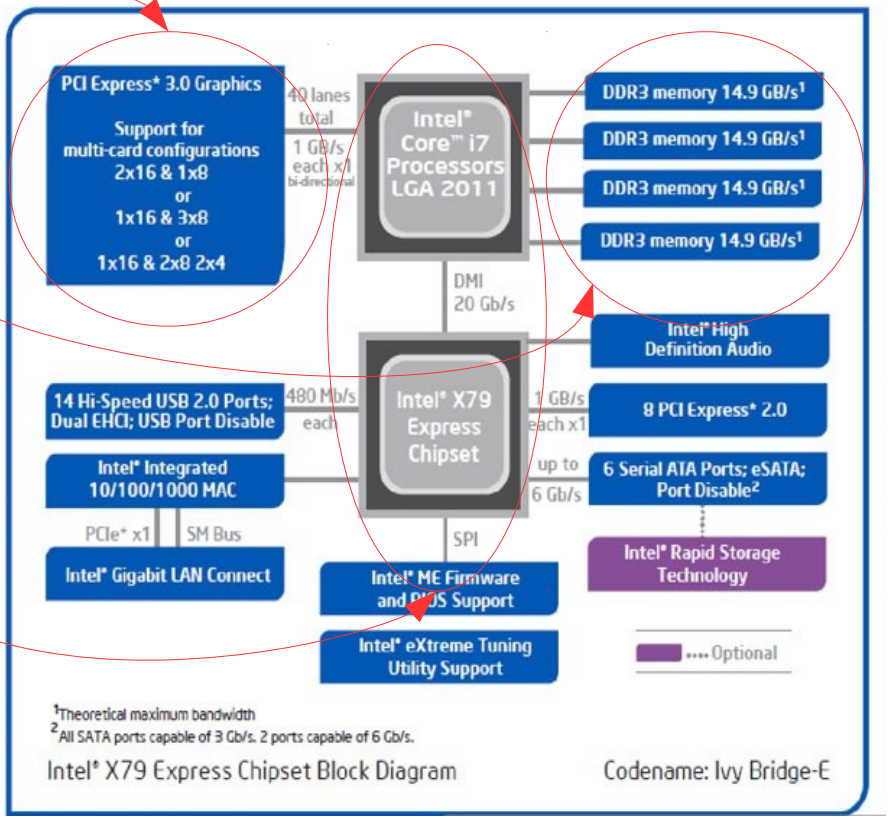


Ivy Bridge architecture

Major technological breaktrough

New architecture of Intel chips includes very large bandwidth data paths

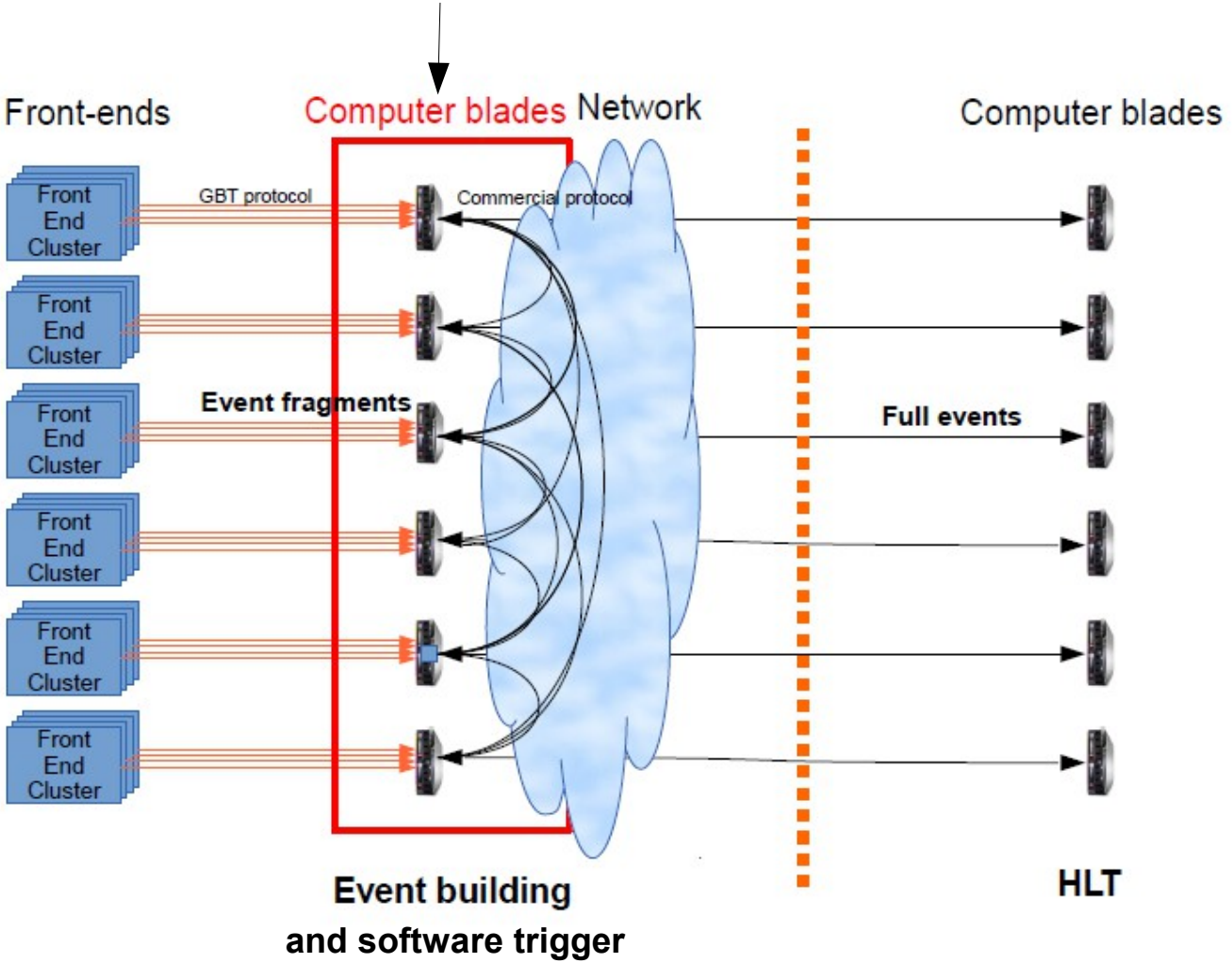
- 40 lanes PCIe GEN3 at 8 Gbits/s
- Independent paths to memory
- Large memory available
- Multicore CPUs
  - Powerful enough to handle both Event building and Software Trigger



Ivy Bridge architecture

New readout/trigger scheme

Move back-end card into already existing CPU blades



Advantages & Issue

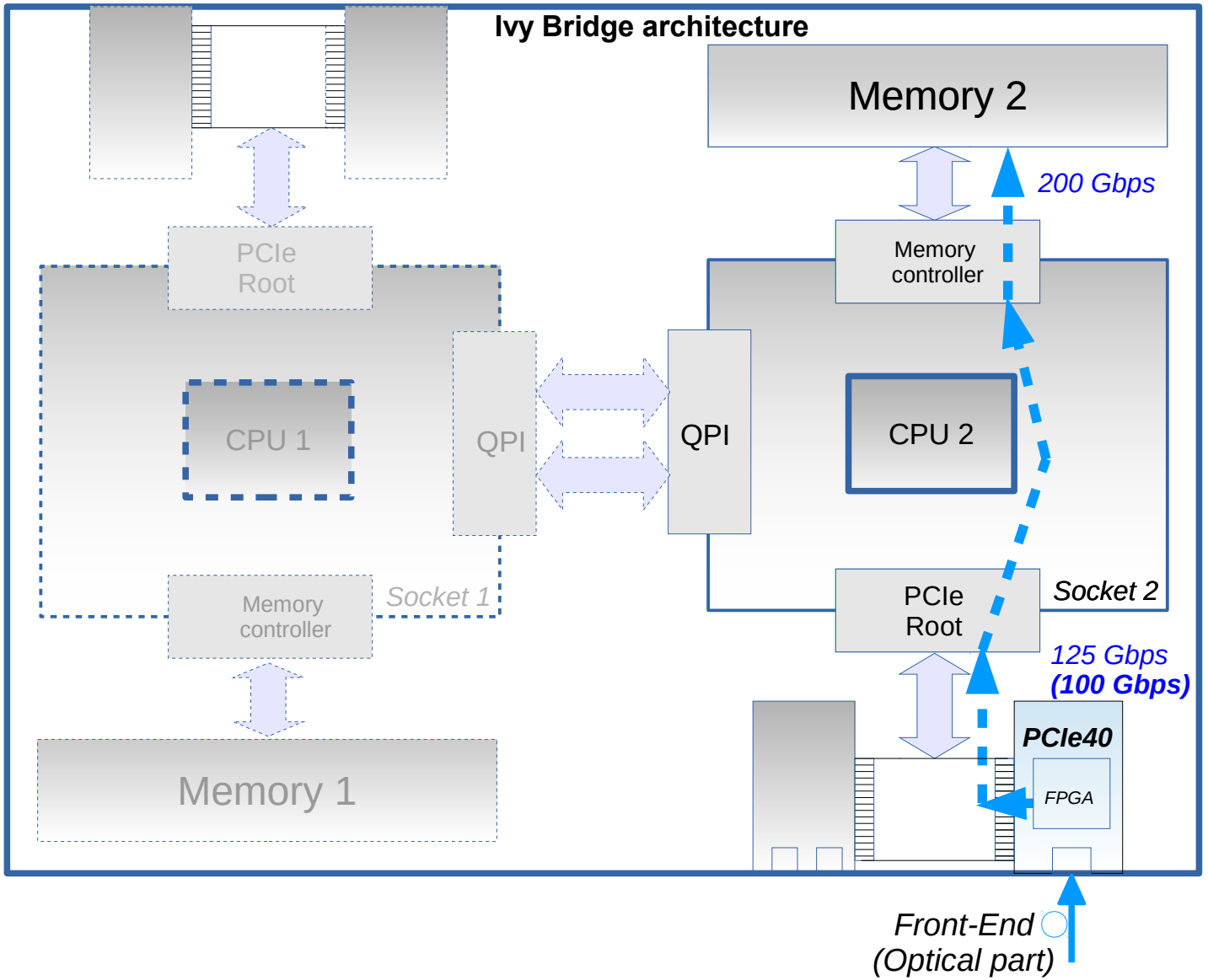
☺ Advantages

- Large memory in the CPU → simpler acquisition board, cheaper switches
- Possibility to run LLT in the Event Building CPU blades
- No more intermediate crates
- Less optical links

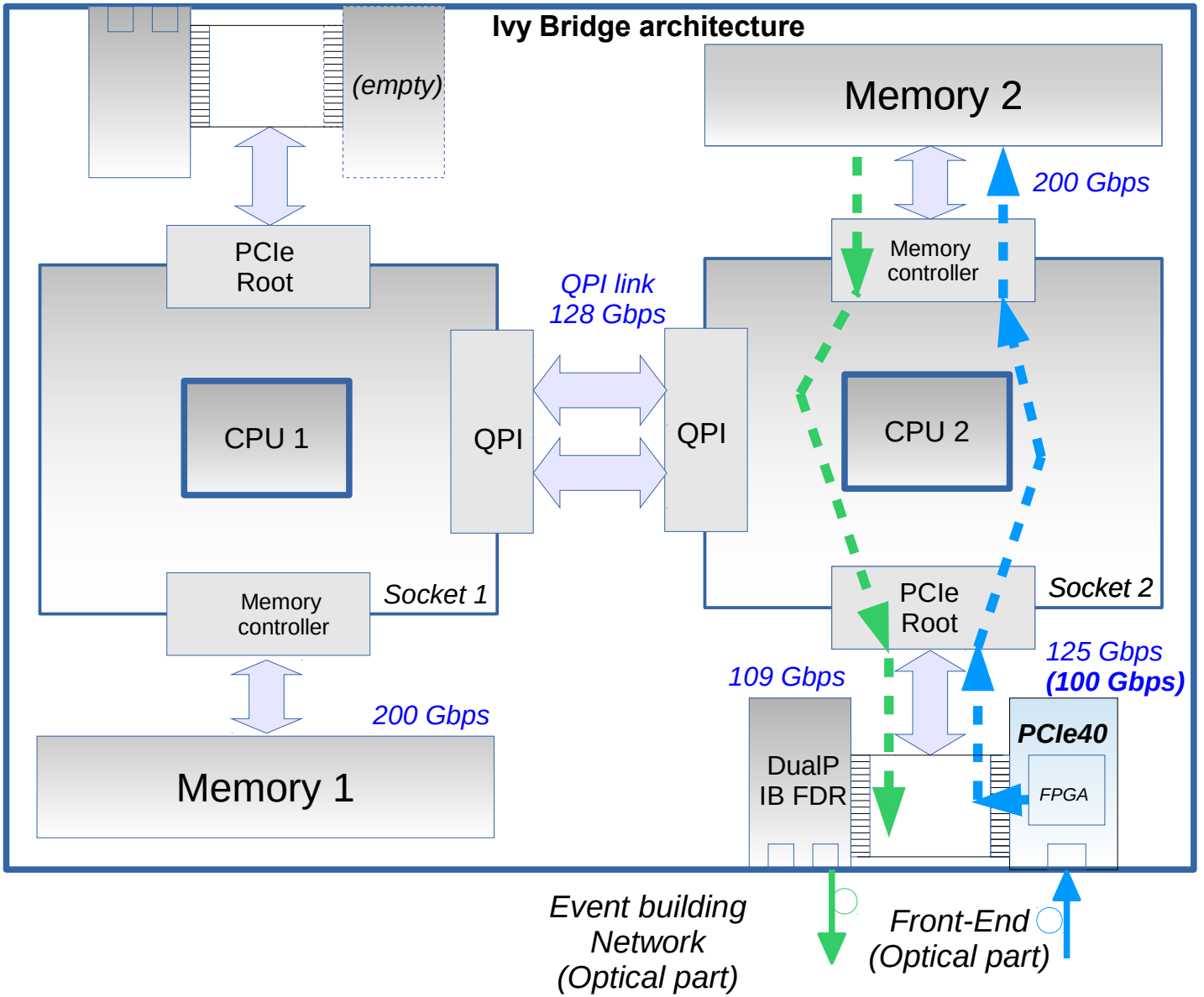
☹ System life duration

- Average life of a PC = ~4y (up to 8y according to CERN statistics)
  - What if PCIe slots not anymore supported by future GENx on PCs?
    Risk to redesign and remanufacture an expensive PCIe40 board.
  - BUT :
    - GEN4 mechanicaly compatible with GEN3
    - GEN5 ? Probably the futur CPU mother board will be equipped with slots GEN3 or 4

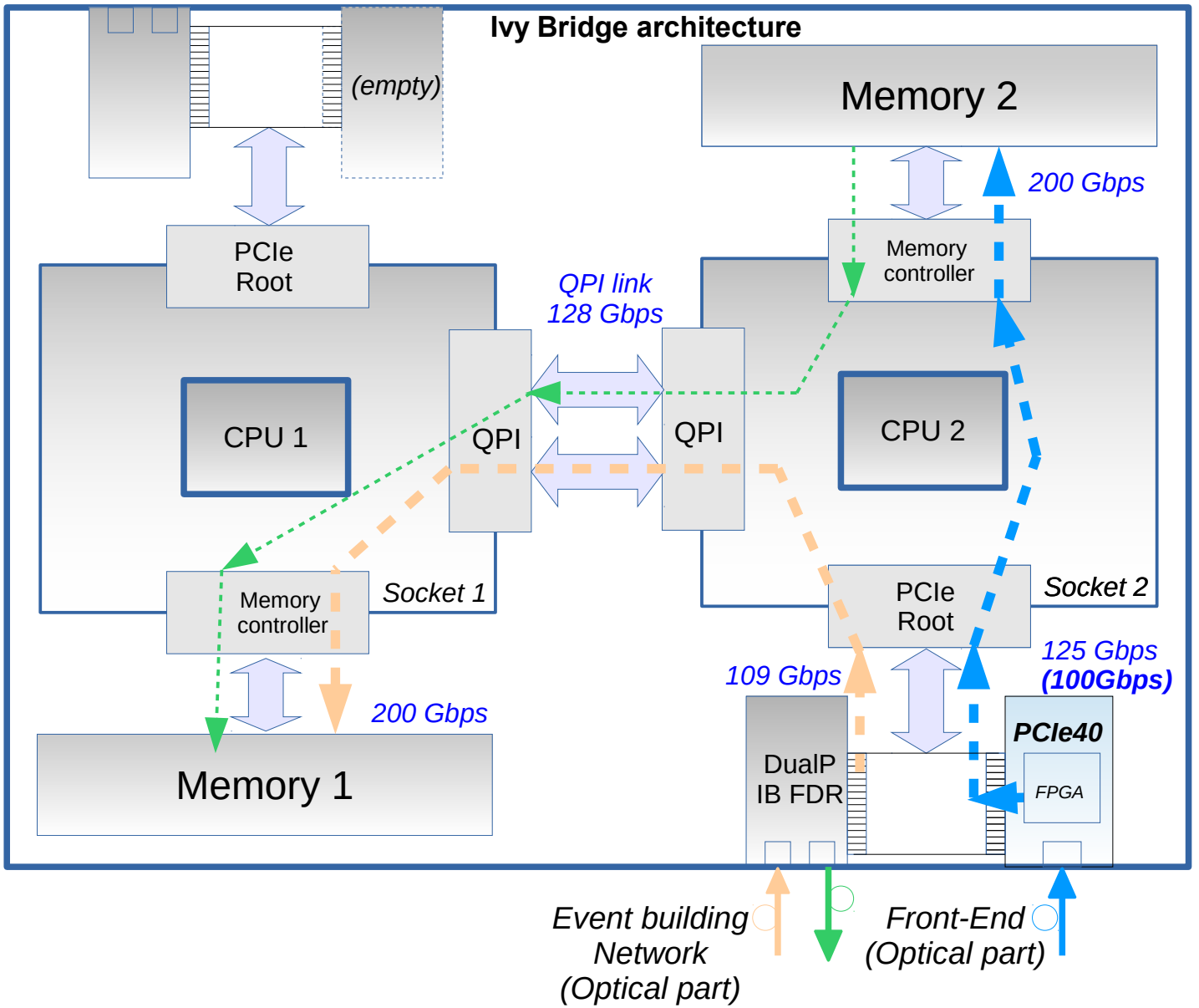
Data path into the computer



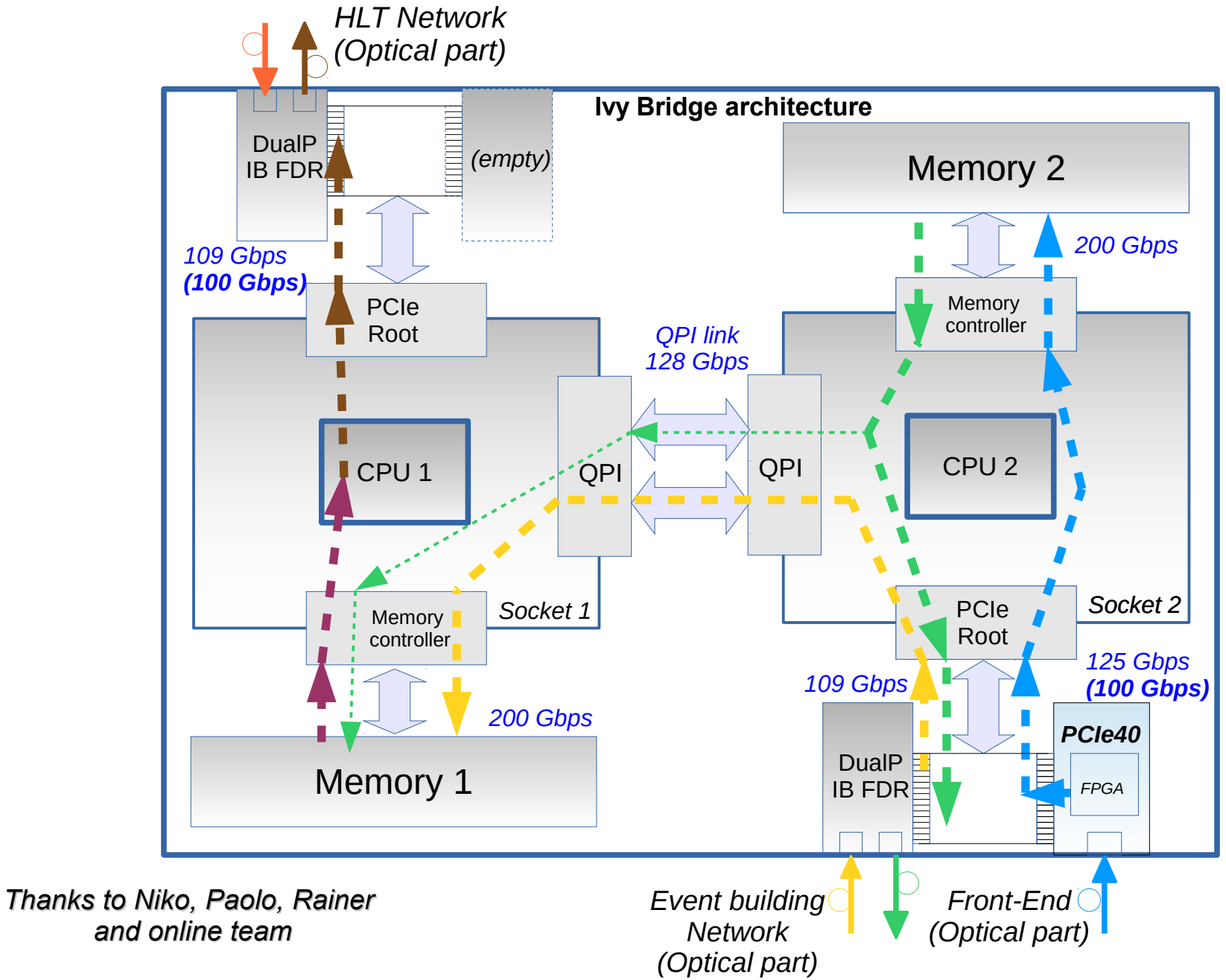
Data path into the computer



Data path into the computer



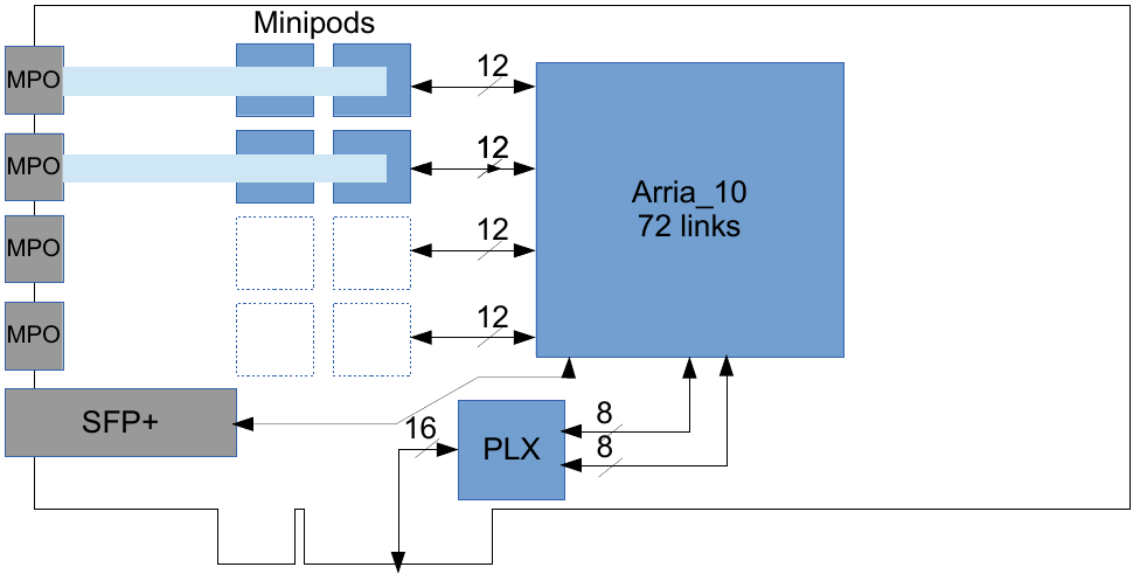
Data path into the computer



PCIE40 Synoptic view

Up to 48 optical inputs/outputs
Serial signal up to 10 Gbps

Bidirectional serial links
for the TFC



output bandwidth ~100 Gbps though PCIe Gen3 x16

- ***Generic board:***
  - **Data Acquisition from the FE boards (@100Gbps)**
  - **TFC supervisor**
  - **TFC & ECS distribution to/from the FE boards**

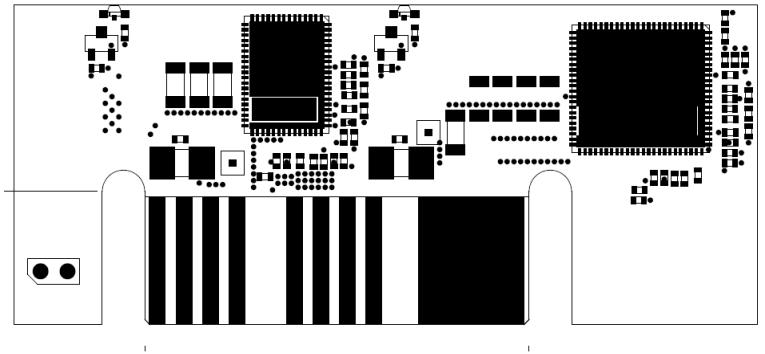
PCIE40 Challenging board

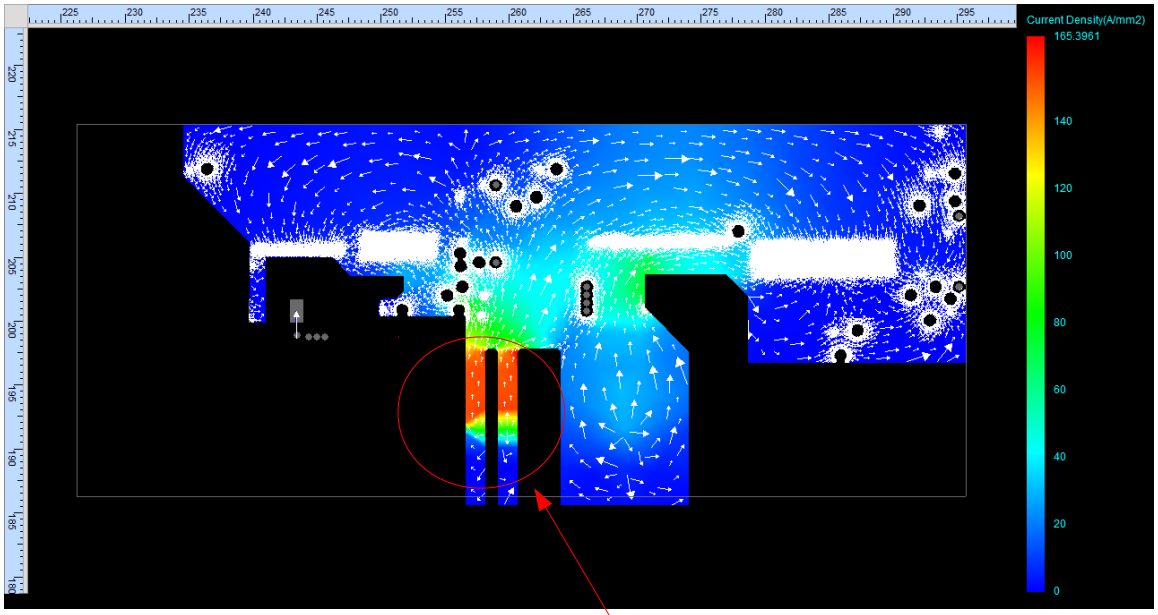
Design of:

- 114 high speed serial signals up to 10 Gbps including PCI Express GEN3x16.
- Reduce form factor to be compatible with most CPU blades
- Developed with the largest ALTERA FPGA currently available (Arria 10)
  - Use of engineering sample (ES1 & ES2)
  - Production chip still not released
- Power dissipation up to 160W
  - 10 different voltages.
  - Precise power sequencing required.
  - Transport high current (50A/0.9V) to the FPGA core.
    - Risk of PCB delamination by thermal effect

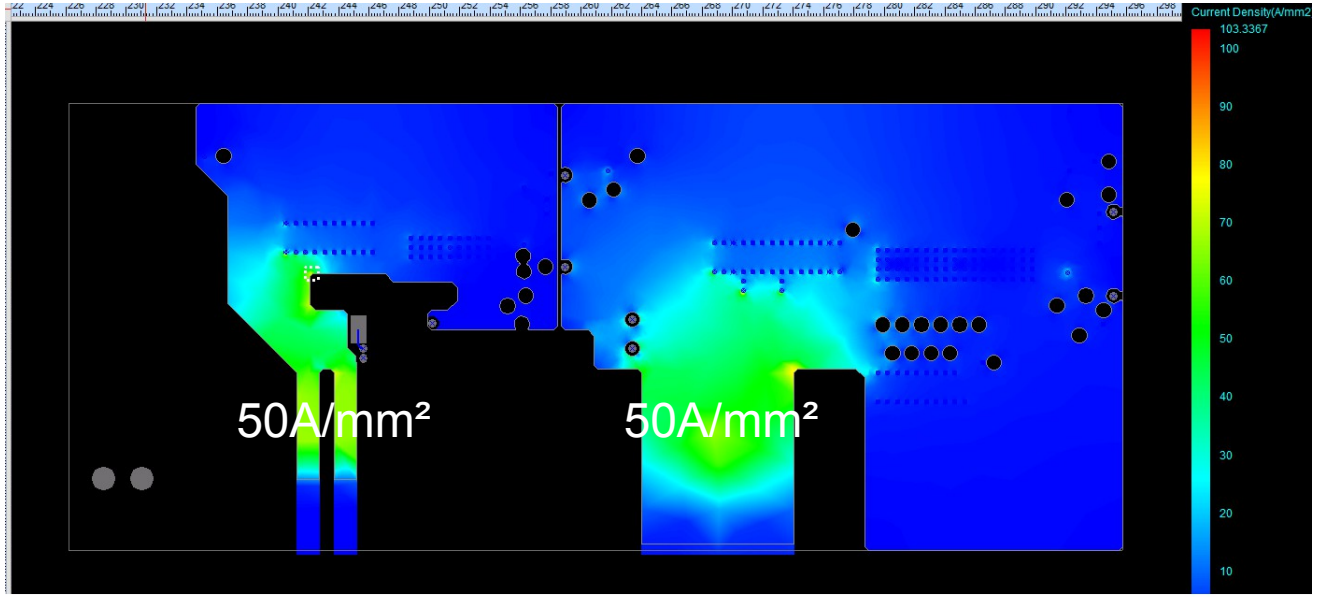
Simulation of current propagation on the ground shape of the power supply Mezzanine board – ***CADENCE – SIGRITY suite***

Previous routing





Final routing

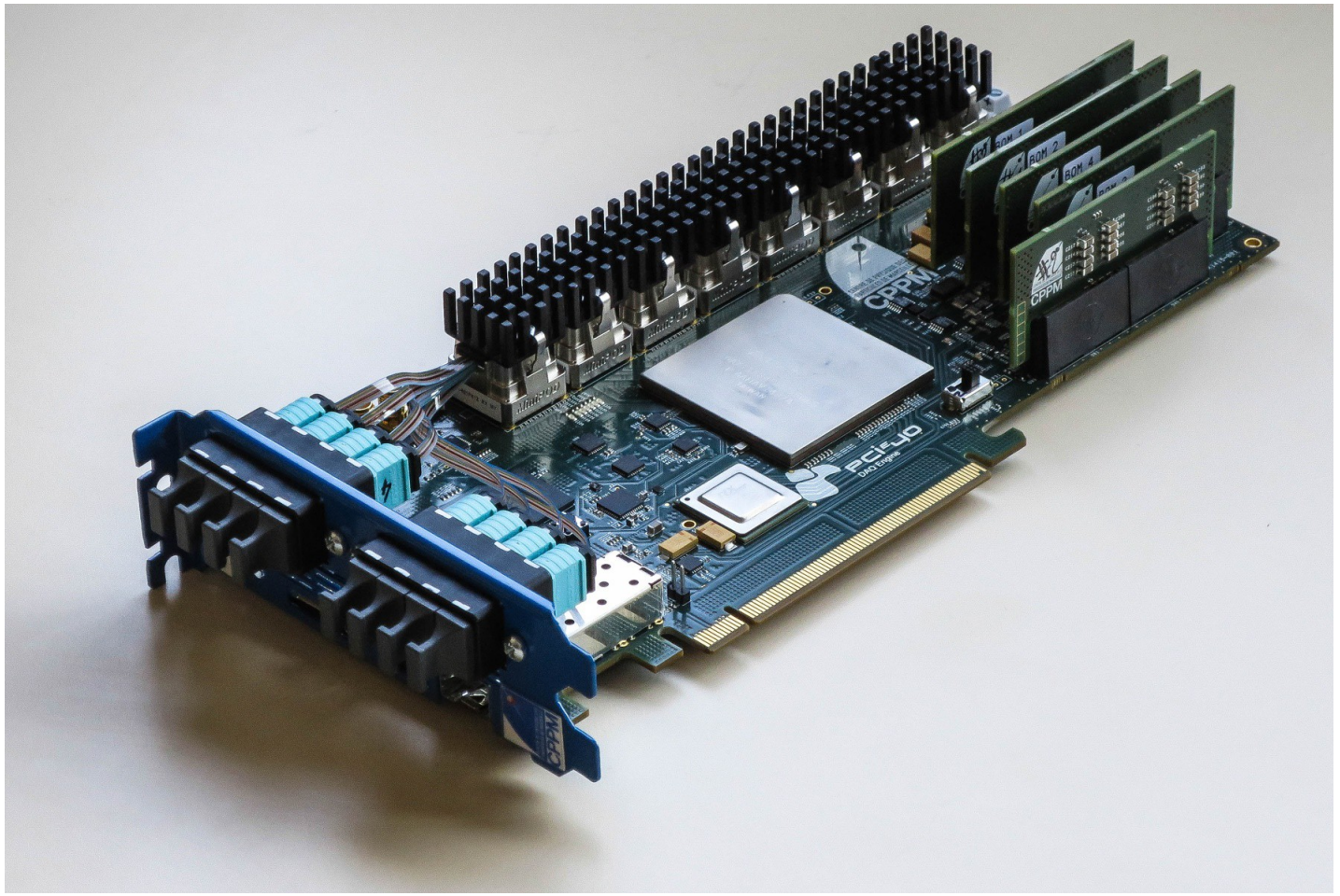


Very high density
Of current (150A/mm²)

Solution: two separated ground shape

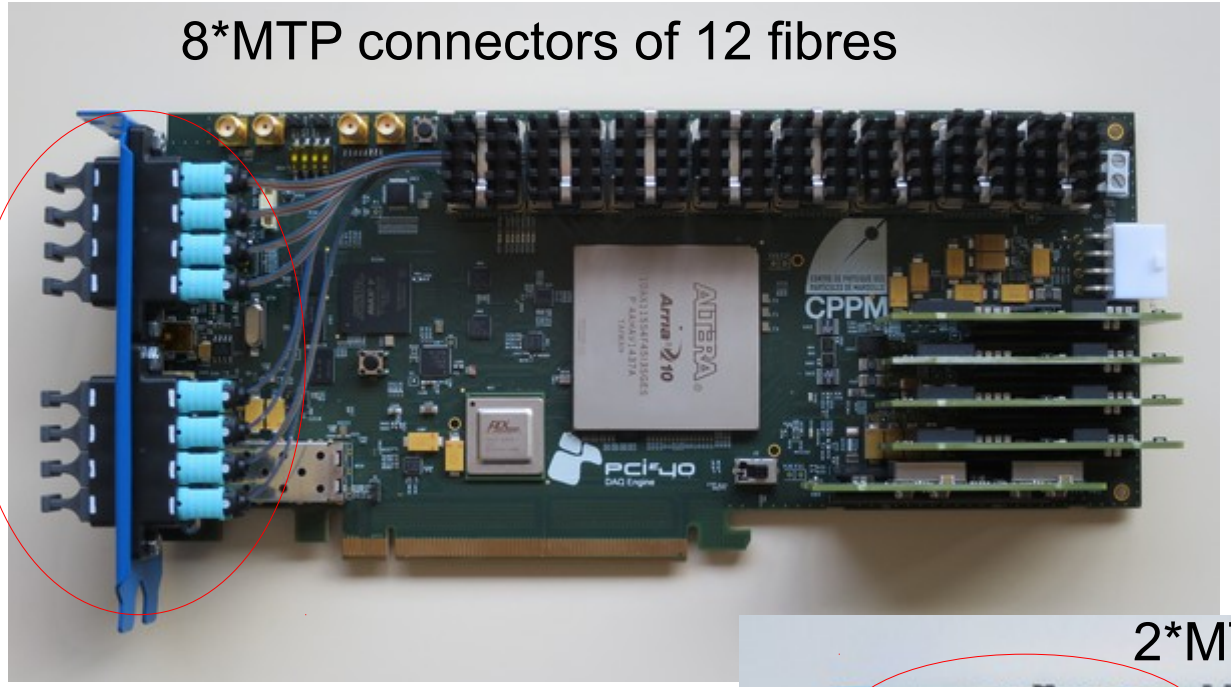
PCIE40 Prototypes

2 prototypes: equipped with Arria10 ES1 & ES2

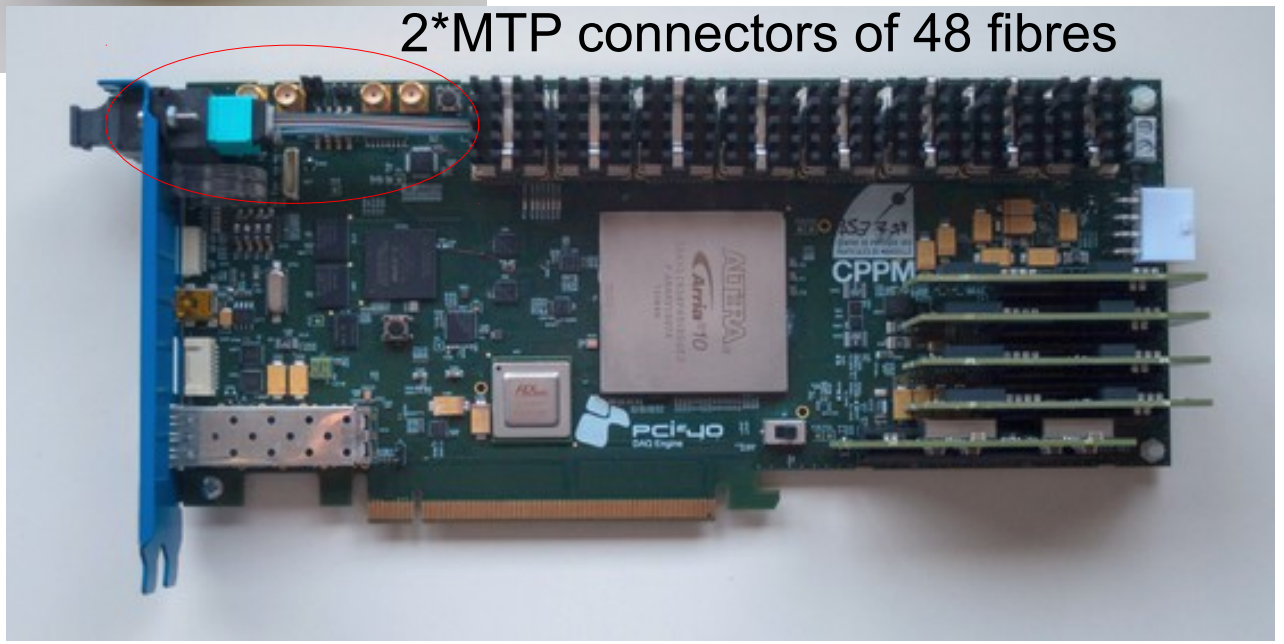


PCIE40 Flexible front panel optical configuration

8*MTP connectors of 12 fibres



2*MTP connectors of 48 fibres

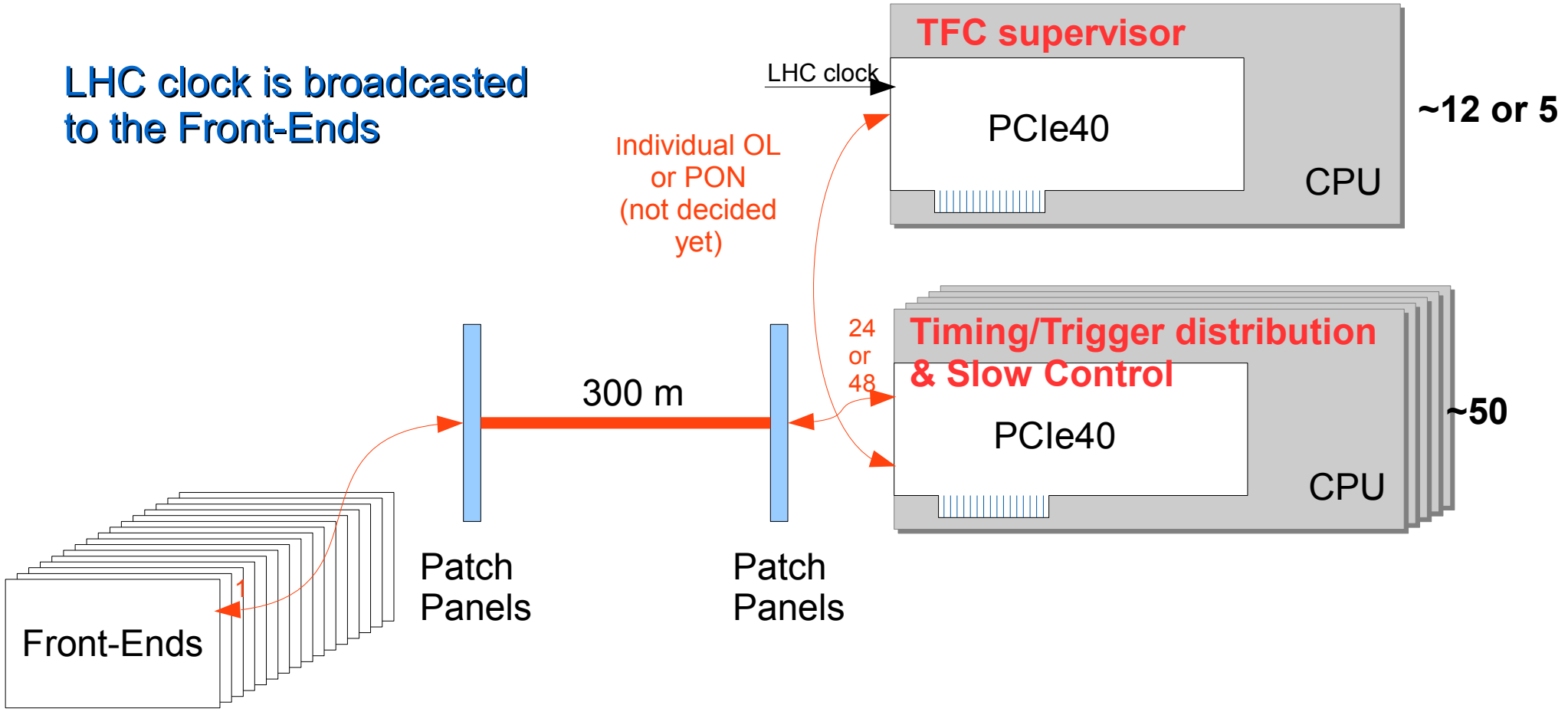


Server integration

Board is operational in the server

Overall architecture

LHC clock is broadcasted to the Front-Ends

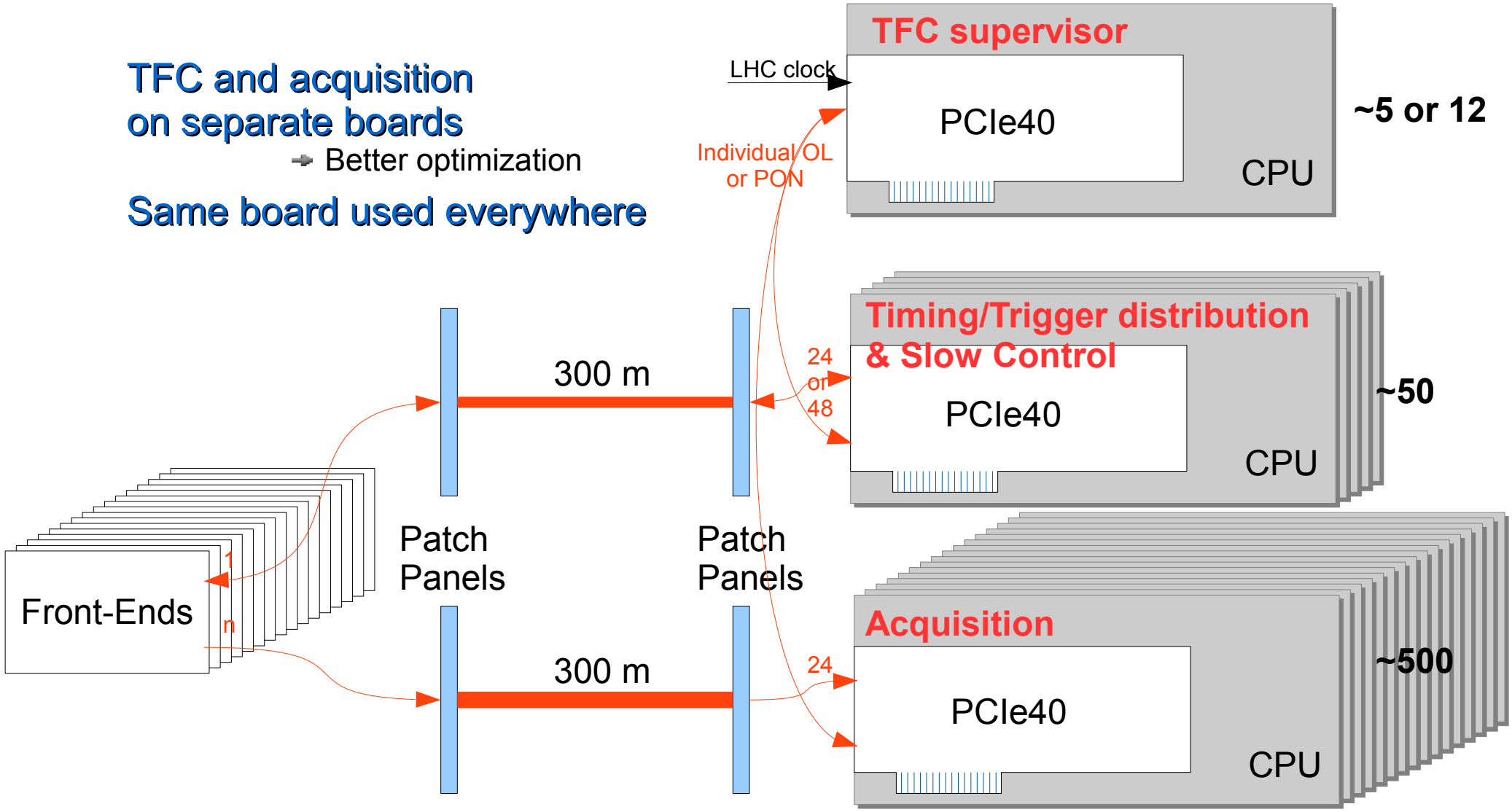


Overall architecture

TFC and acquisition on separate boards

- Better optimization

Same board used everywhere



Conclusion

Feasibility of a triggerless readout system with event building in the farm has been demonstrated

Now the base line solution for LHCb experiment.

Two prototypes of PCIe40 boards are fully operational

- *Robustness tests (Temperature, cooling, BER, …) to be done in the next months*

20 Boards will be duplicated early 2016 to provide « Mini-daq » setups for the collaboration

Full production foreseen 2016/2017