Charge readout front-end electronics, DAQ and online storage/computing facility

Dario Autiero (IPNL)

24/4/2017

WA105 Accessible cold front-end electronics and uTCA DAQ system 7680 ch

Full accessibility provided by the double-phase charge readout at the top of the detector

➢ **Digital electronics at warm on the tank deck:**

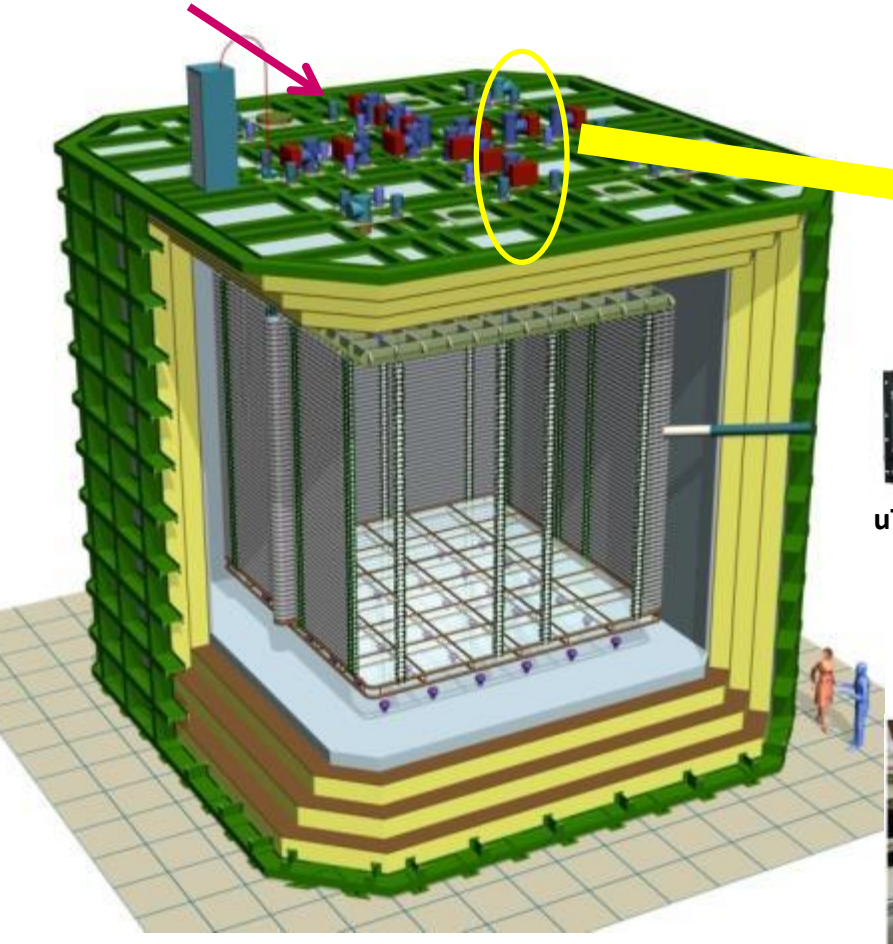
- Architecture based on uTCA standard
- 1 crate/signal chimney, 640 channels/crate

→ 12 uTCA crates, 10 AMC cards/crate, 64 ch/card

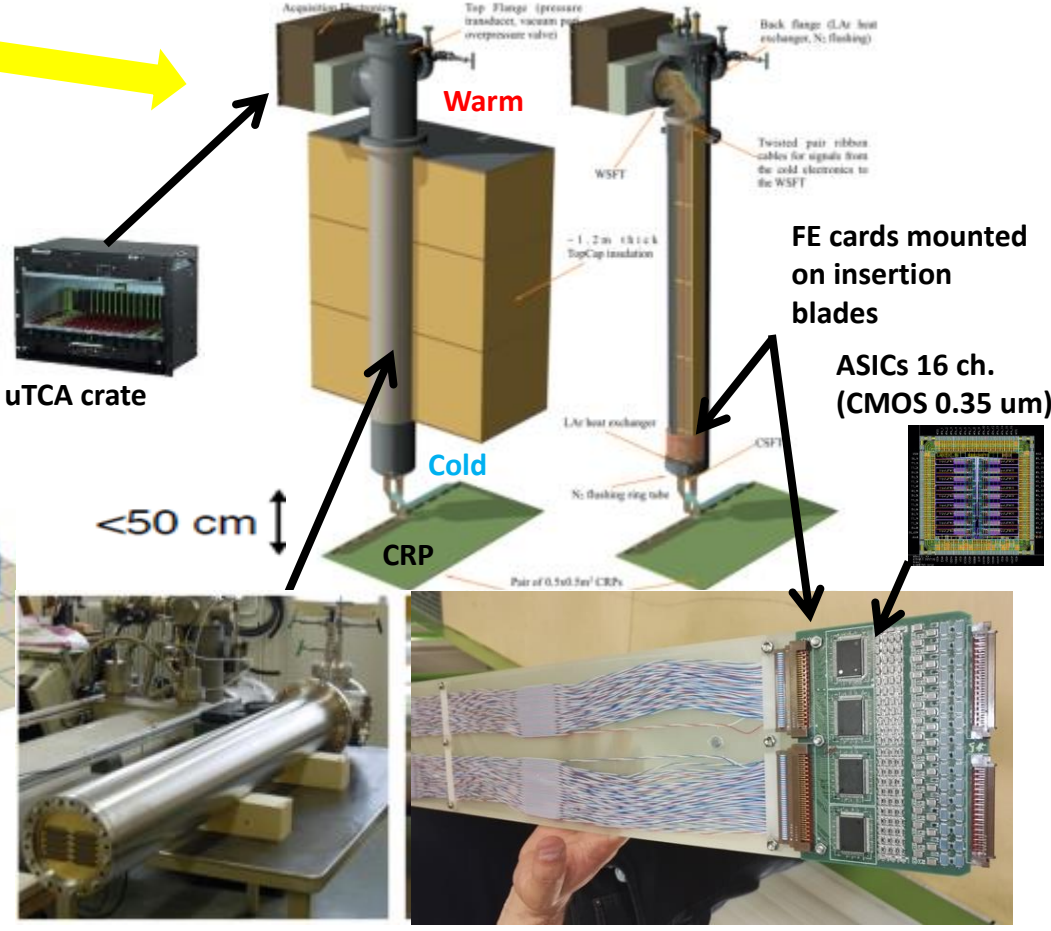
➢ **Cryogenic ASIC amplifiers (CMOS 0.35um) 16ch externally accessible:**

- Working at 110K at the bottom of the signal chimneys
- Cards fixed to a plug accessible from outside

→Short cables capacitance, low noise at low T



Signal chimney



Cost effective and fully accessible cold front-end electronics and DAQ

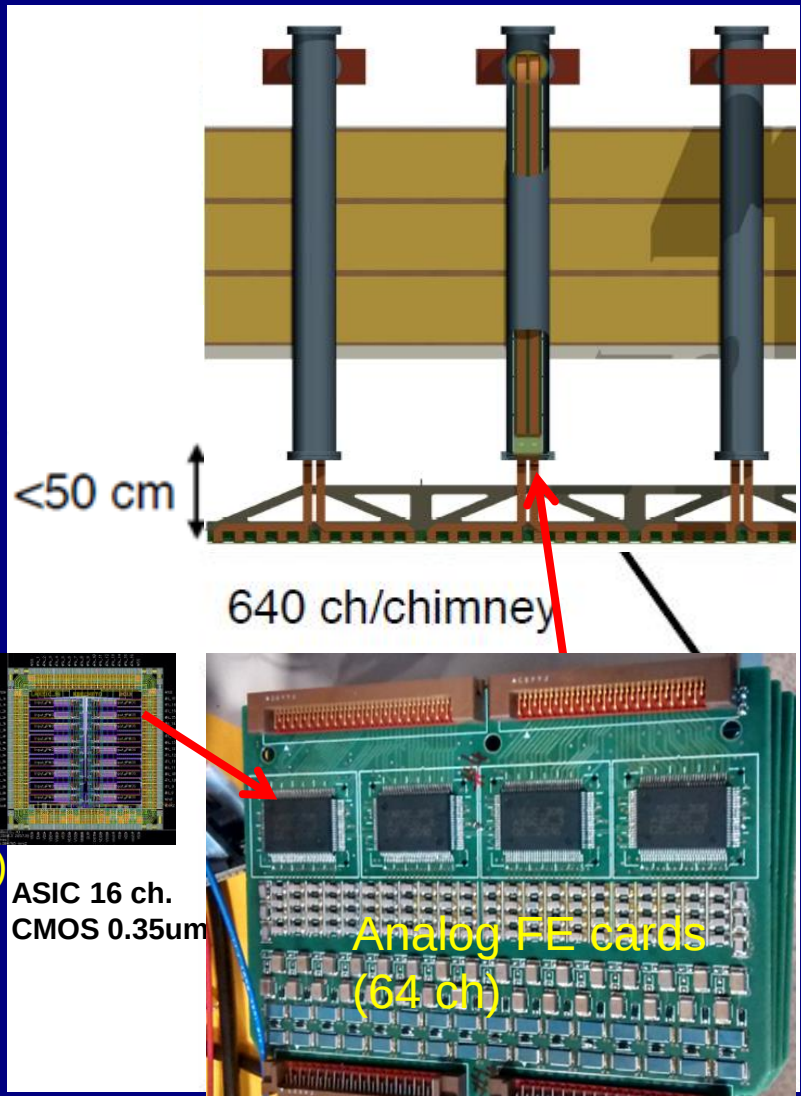
Ongoing R&D since 2006→ in production for 6x6x6 (7680 readout channels)

ASIC (CMOS 0.35 um) 16 ch. amplifiers working at ~110 K to profit from minimal noise conditions:

- FE electronics inside chimneys, cards fixed to a plug accessible from outside
- Distance cards-CRP<50 cm
- Dynamic range 40 mips, (1200 fC) (LEM gain =20)
- 1300 e- ENC @250 pF, <100 keV sensitivity
- Single and double-slope versions
- Power consumption <18 mW/ch
- **Produced at the end of 2015 in 700 units (entire 6x6x6)**
- **1280 channels installed on 3x1x1**

DAQ in warm zone on the tank deck:

- Architecture based on uTCA standard
- Local processors replaced by virtual processors emulated in low cost FPGAs (NIOS)
- Integration of the time distribution chain (improved PTP)
- Bittware S5-PCIe-HQ 10 Gbe backend with OPENCL and high computing power in FPGAs
- **Production of uTCA cards started at the end of 2015, pre-batch already deployed on 3x1x1**

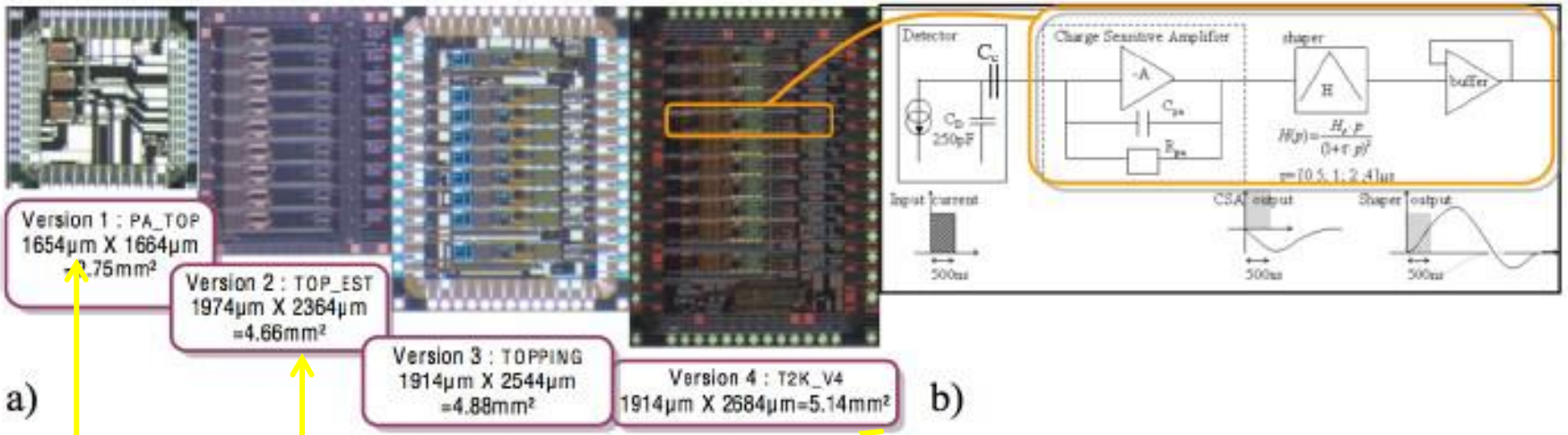


→ Large scalability (150k channels for 10kton) at low costs

Analog FE electronics (Design)

Since 2006 6 versions of the analog ASIC (CMOS 0.35 um) at cold were developed for single-phase + 3 for dual-phase dynamics

2007 2008 2009 2010



Amplifier
+ test components

8 channels shaper + buffer, xtalk optimization

8 channels shaper + buffer, selectable configuration

8 channels shaper + buffer
Phase margin optimization

First dual-phase version DP-1

Extended dynamics for LEM (produced at the end of 2013 and tested in 2014)

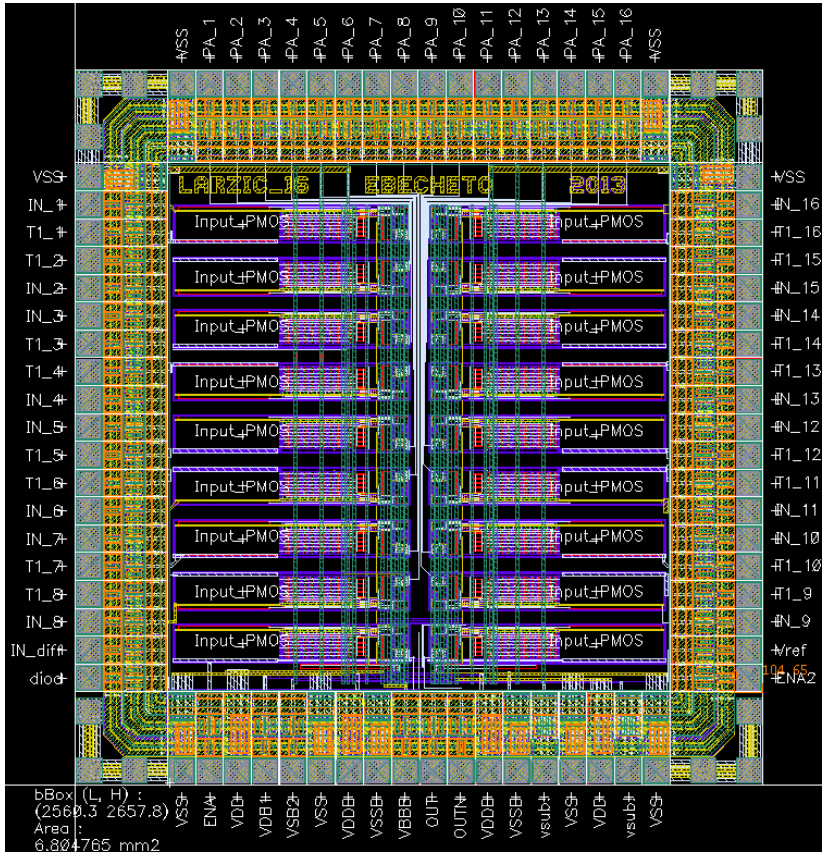
- 16 channels
- Evolution from single-phase versions: end of dynamic range 150 fC → 1200 fC (to cope with gain in double-phase charge readout**)** 40 mip
- Single slope gain

Other characteristics:

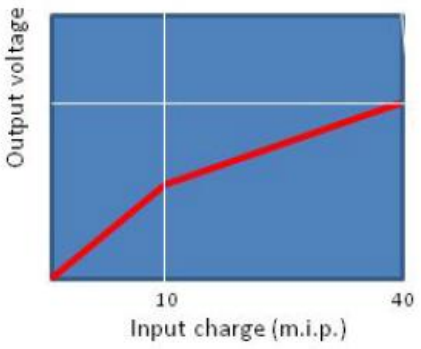
- Power consumption 18 mW/ch
- ENC 1300 e- @ Cdet=250 pF

Basic assumptions:

- LEM gain ~20, split in two collection views
- Sensitivity ~100 keV
- Dynamics 40 mips



Dual-phase double-slope gain DP-2

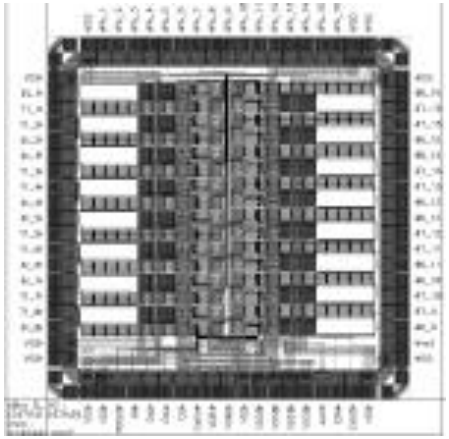


Adapted to LEM dynamics like previous version 1200 fC single slope

Larger gain up in the few mip region, kink point point at 10 mip and reduced gain by a factor 3 up to 40 mips max dynamic range

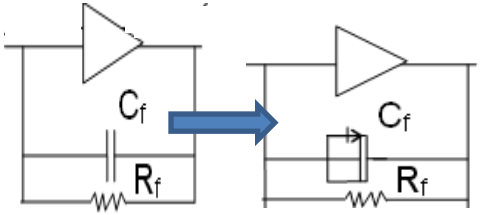
→ **Produced in fall 2014, received at beginning of January 2015**

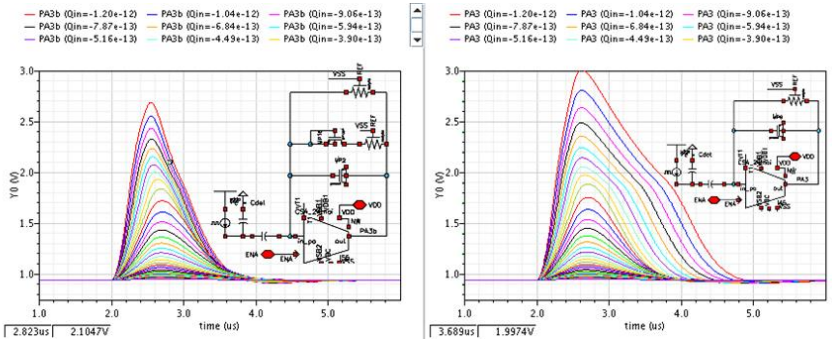
Double slope implementation (starting from previous ASIC version DP-1):

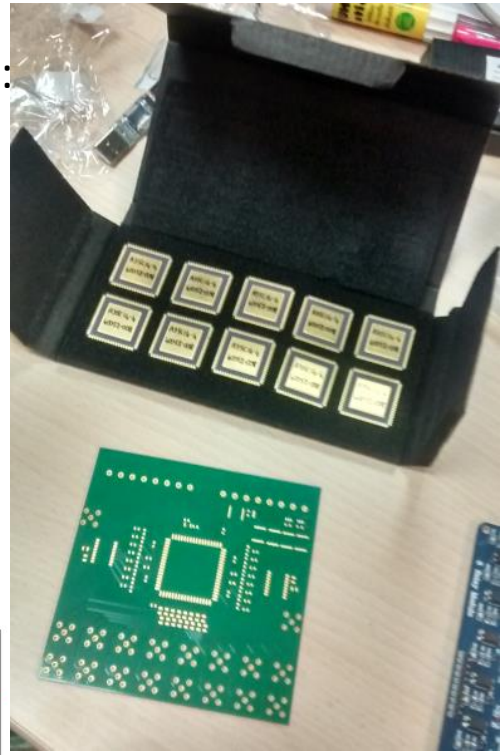


→ Replace feedback capacitor of the preamplifier with a MOS capacitance which changes the C value above a certain threshold voltage (gain ~ 1/C).

Selectable double time constant in discharge or single one with diode +resistor to keep constant RC



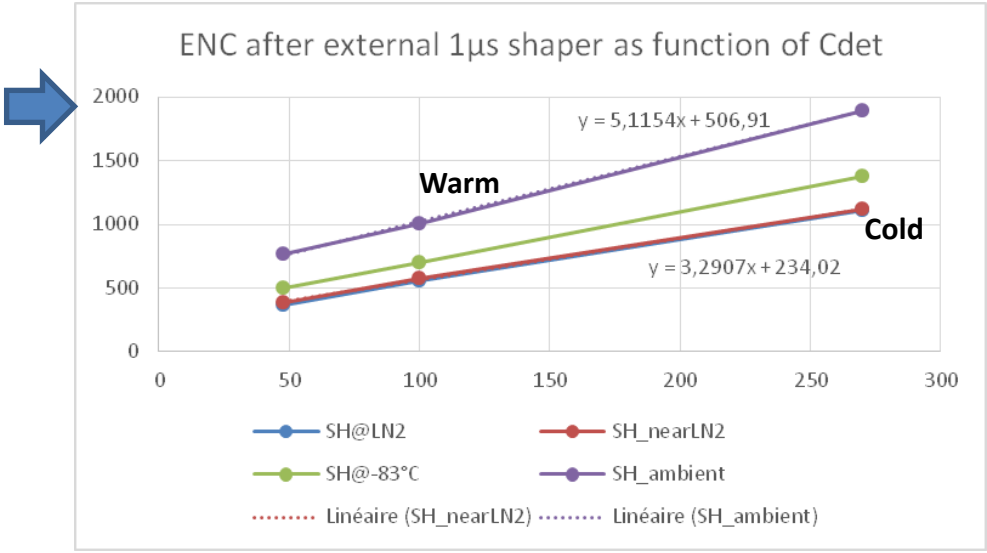


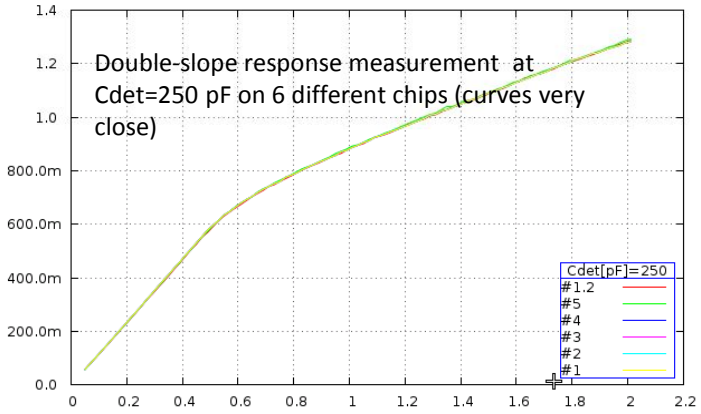


DP-2

Dual-phase single-slope version produced in in 2013 (up to 1200 fC dynamic range) and a double-slope version at the end of 2014 → Both versions have noise within specifications and working correctly at cold:

Noise measurements as a function of Cdet and various temperatures. At cold around 100k: ENC=3.3*Cdet+234 e- → 1200 e- at 300 pF





Some improvements for DP-3:

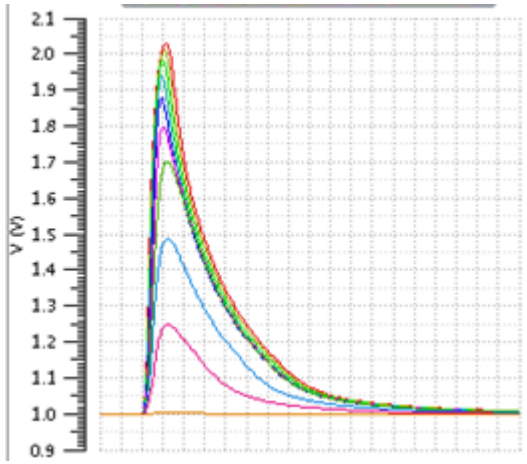
- The value of the MOSCAP capacitance physically resulting in the circuit implementation smaller than the one in the submitted design (150 pF vs 250 pf) due to process dependence

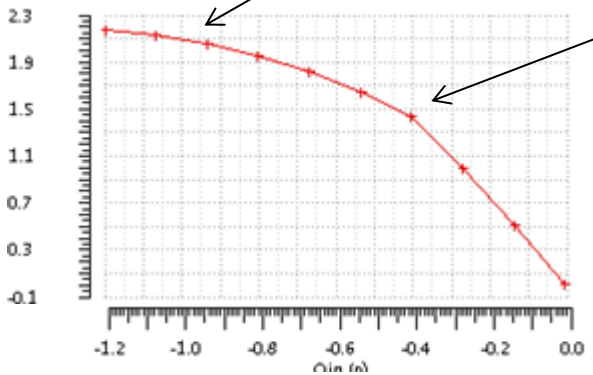
- The smaller value of Cf plus a parasitic capacitance effect on the feedback resistor branch introducing a dependence of the response on Cdet (lower signal by increasing Cdet and longer peaking time) → fix parasitic capacitance effects

DP-3 submitted in fall 2015

- Design of MOSCAP less process dependent
- Removal of parasitic capacitance on feedback resistor branch
- Better differential driver integrated from another IN2P3 development

Simulation at cold of DP-3 :



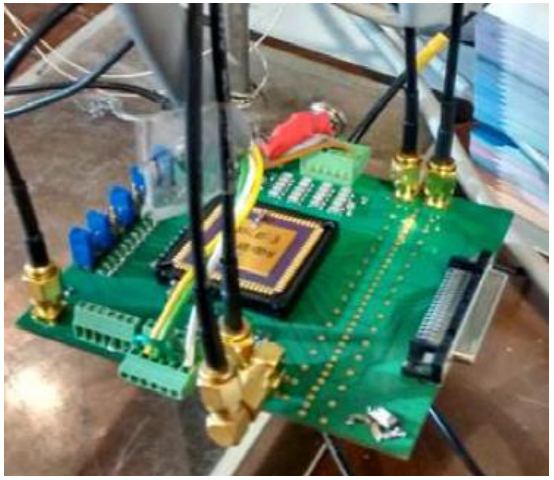


~log regime up to 1200 fC

Kink at 400 fC
~13 mip with
LEM gain = 20

- Circuit produced as a test batch (25 units) with purchase option for already produced 600 units (entire WA105 production) if the tests on the 25 ordered units confirmed expectations

Cryogenic FE electronics :

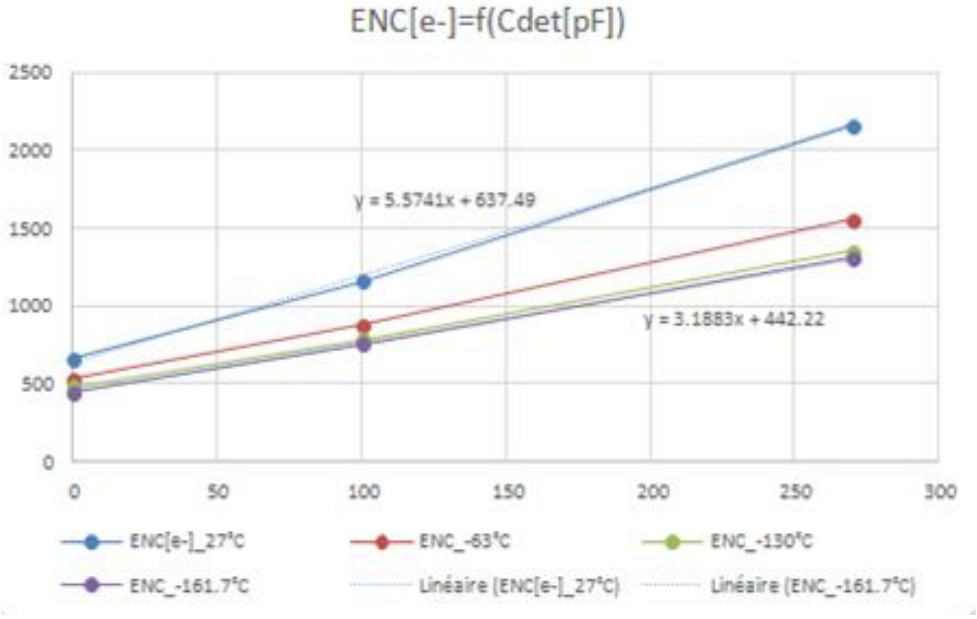


Dual-slope ASICs final version DP-3

- 16 channels
- Double slope gain with “kink” at 400 fC
- 1200 fC dynamic range

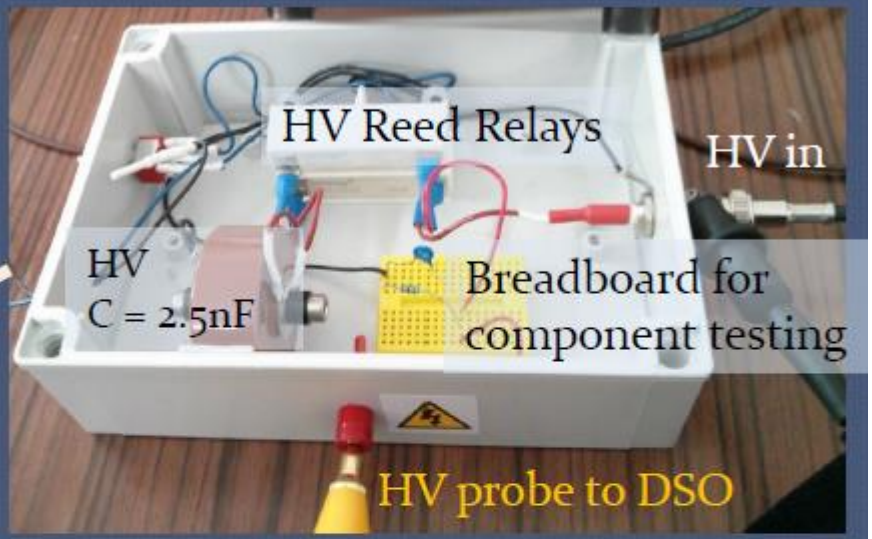
(batch of 25 circuits) tested in January 2016
→ fully satisfactory.

Full production for 6x6x6 produced and purchased (700 chips) in March 2016



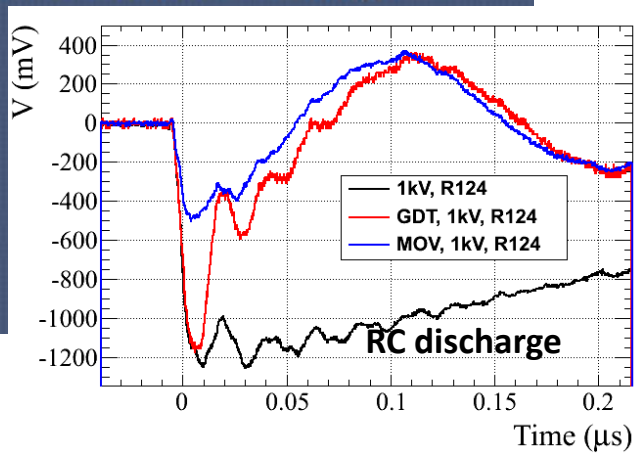
Test-bench for ESD protection studies

Each discharge is recorded with high bandwidth DSO



Test-bench to evaluate & stress-test different ESD protection technologies

Raspberry Pi gives programmable control of relay switching



Double opposite diodes components preferred for capacitance and performance

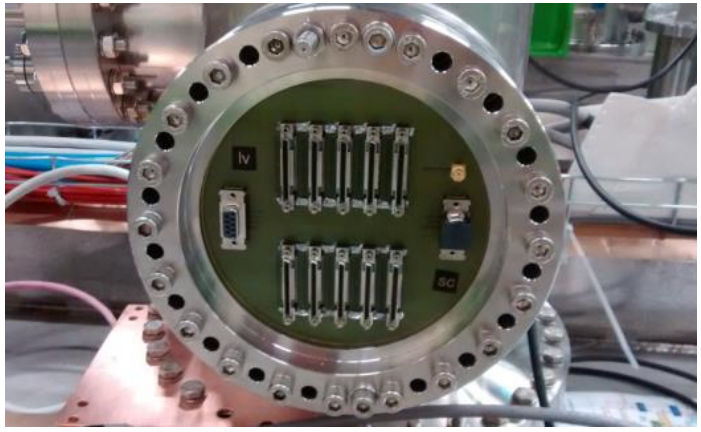
FE-cards designed in 2016 together with chimneys warm flanges PCBs

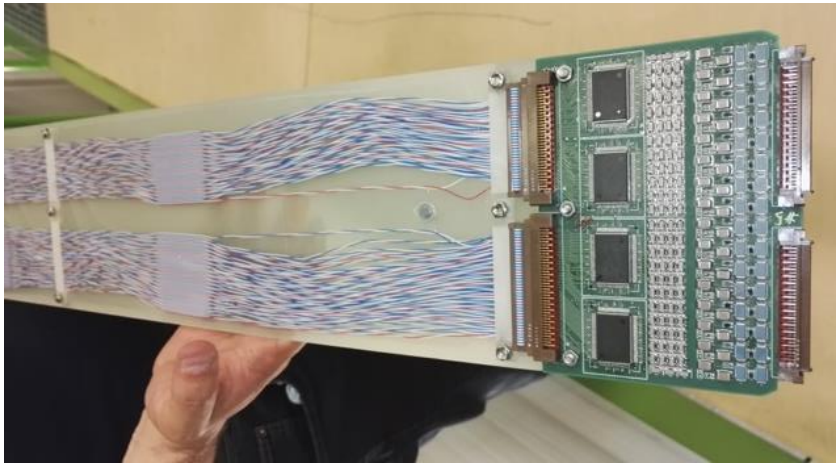
4 double-slope ASICS DP-3 → 64 channels + protection components

20 FE cards (1280 channels) produced and installed on 3x1x1 pilot detector at CERN since September 2016

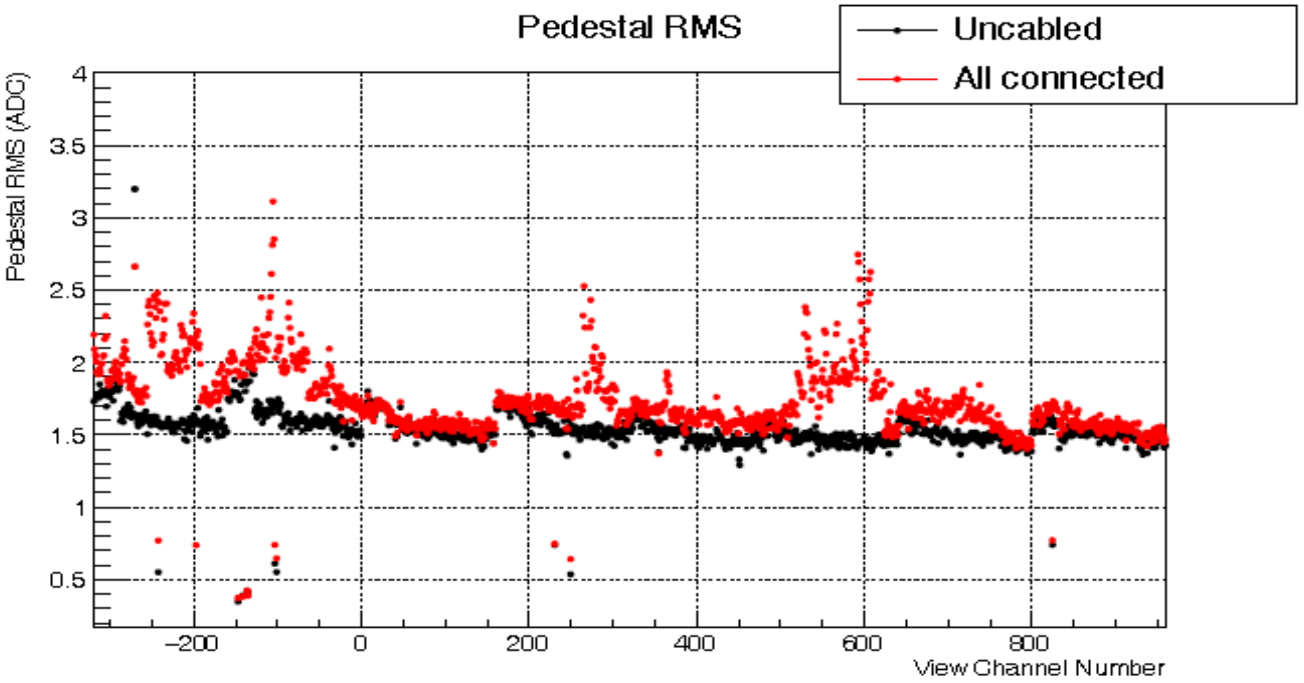
Insertion and extraction in the chimneys with blades tested over many months

Warm flange PCB





Several campaigns of checking of the grounding conditions/noise measurements since June 2016.
Good noise conditions with some residual small issues related to slow-control/HV grounding and cabling
→ Average RMS noise 1.7 ADC counts (0.82 mV) at warm with all systems active and cabled 1.5 ADC counts with slow control/HV cables disconnected from flanges



The grounding scheme for the 6x6x6 is more sophisticated with the cryostat, FE electronics and slow control completely insulated from external environment and only referred to cryostat ground (same as for single-phase).

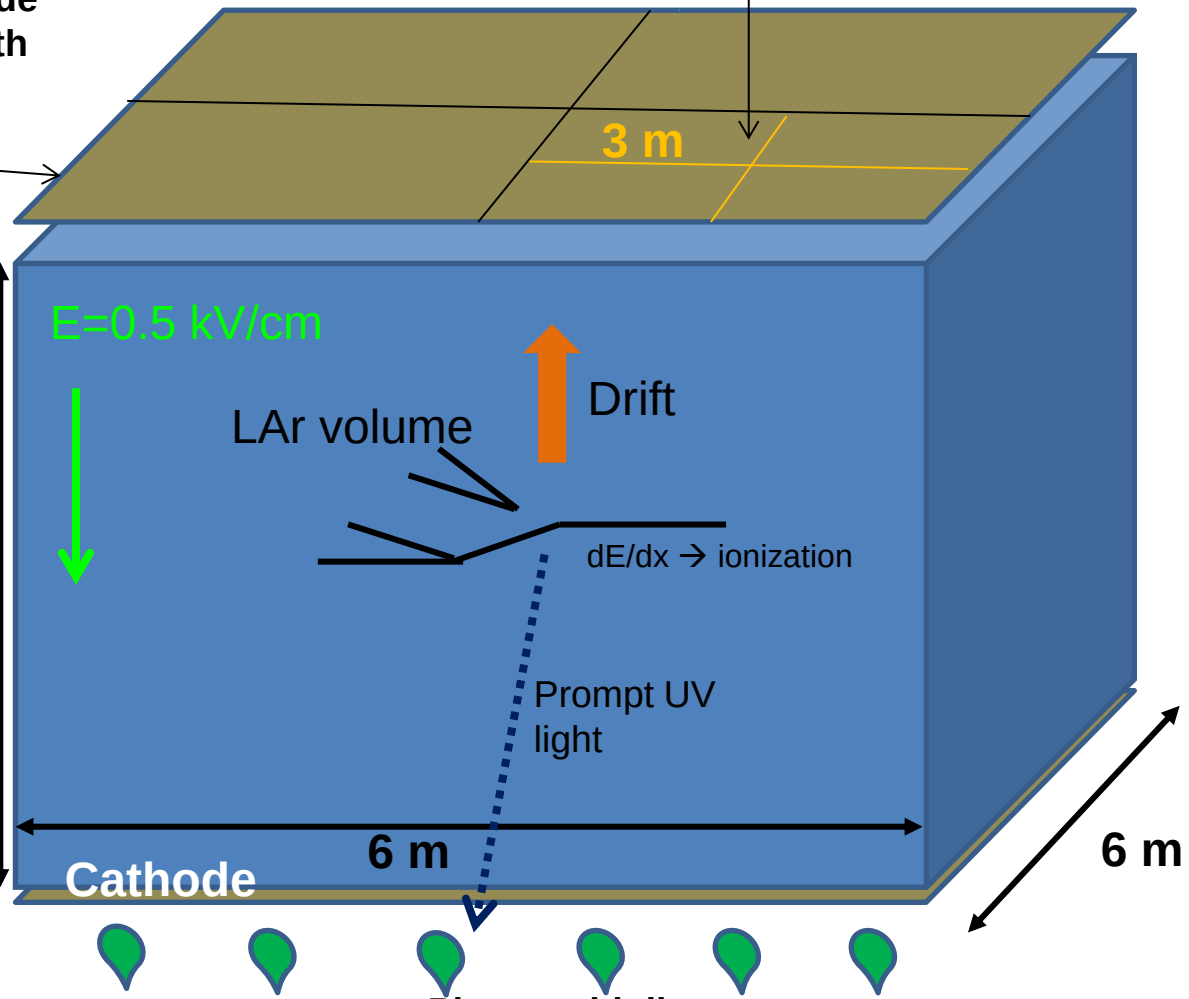
Digital FE electronics and DAQ (Design)

Double phase liquid argon TPC
6x6x6 m3 active volume

→ **Event size: drift window of**
7680 channels x 10000 samples = 146.8 MB

X and Y charge collection strips
3.125 mm pitch, 3 m long
→ 7680 readout channels

Segmented anode
In gas phase with
double-phase
amplification



Drift coordinate
6 m = 4 ms
sampling 2.5 MHz
(400 ns), 12 bits

→ 10000 samples
per drift window

- **Dual phase ProtoDUNE detector characteristics:**

- Two views with 3.125 mm pitch → 7680 channels
- Long drift 4 ms → 10000 samples at 2.5 MHz
- High S/N~100

- **All electronics at warm, accessible**

→ **Costs minimization**, massive use of **commercial** large bandwidth standards in telecommunication industry, **uTCA**, Ethernet networks, massive computing

→ Easy to follow technological evolution, benefit of costs reduction and increase of performance in the long term perspective

- **Non-zero suppressed data flow** handled up to computing farm back-end which is taking care of final part of event building, data filtering, online processing for data quality, purity, gain analysis, local buffer of data and data formatting for transfer to EOS storage in files of a few GBs

- Signal lossless compression benefits by high S/N ratio, developed an optimized version of Huffman code reducing data volume by at least a factor 10

- Timing and trigger distribution scheme based on White Rabbit (became commercial hardware too); thought since the beginning for **a beam application** (handles beam window signals, beam trigger counters, external trigger counters for cosmics) → Components of the timing chain purchased and uTCA slave card for signal distributions on crates backplane developed

- FE based on microTCA standard and 10 Gbit/s ethernet

→ 120 uTCA digitization boards went under production for the 6x6X6 since 2015, 20 card already installed on 3x1x1

- Light readout fully integrated in charge readout uTCA scheme and with different operation modes in-spill out-spill

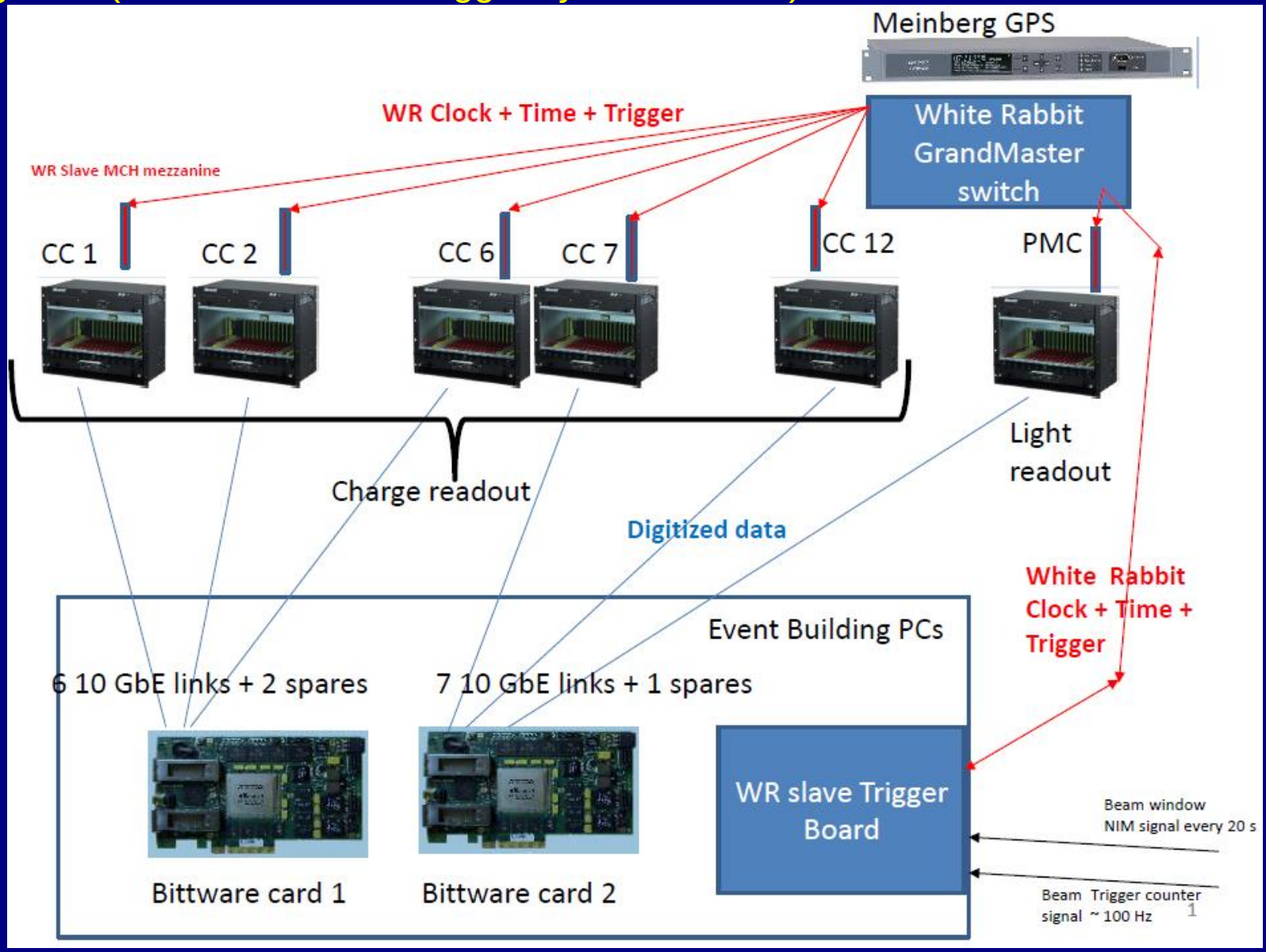
- Back-end actually based on two commercial Bittware cards with x8 10 Gbit links with high computing power and event building capabilities. Each card performs event building for ½ of the detector charge readout, one card deals with light signals too

- Online storage and computing facility is an important part of the system, a possible implementation has been designed and costed with DELL, it has been implemented on a smaller scale for the 3x1x1 (September 2016)

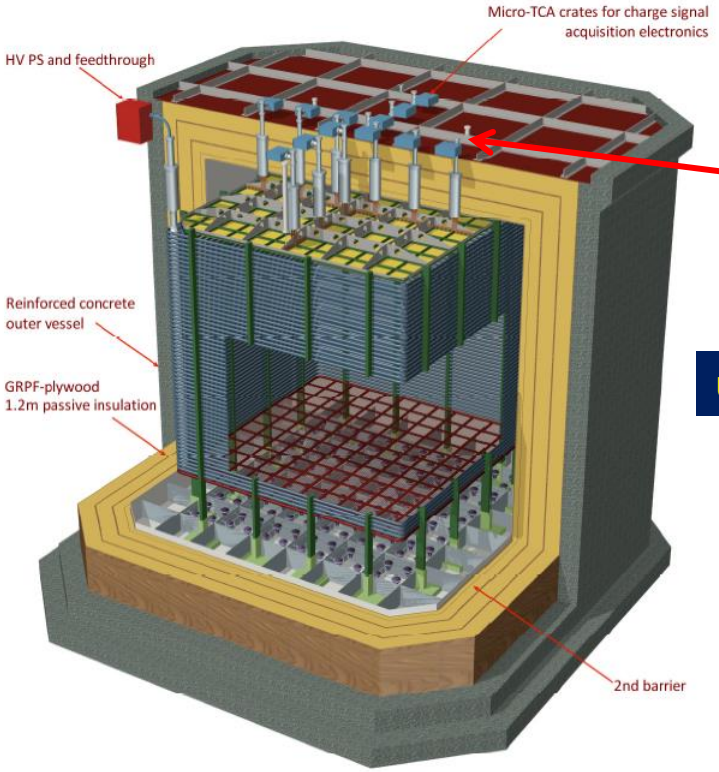
- 20 Gbit/s data link foreseen for data transfer to computing division (final implementation 2x20 Gbit/s links)

Global uTCA DAQ architecture

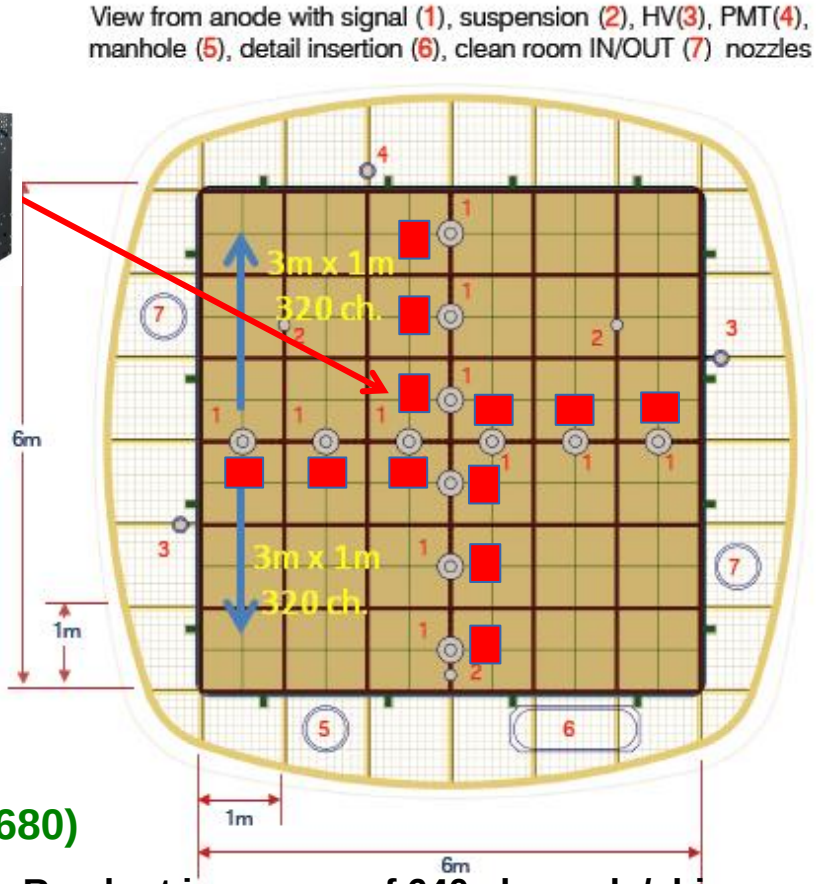
integrated with « White Rabbit » (WR) Time and Trigger distribution network
+ White Rabbit slaves nodes in uTCA crates +
WR system (time source, GM, trigger system, slaves)



uTCA charge readout architecture



uTCA crates



Readout in groups of 640 channels/chimney

- 12 chimneys, 640 channels/chimney (640*12=7680)
- 1 uTCA crate/chimney, 10 Gb/s link
- 10 AMC digitization boards per uTCA crate, 64 readout channels per AMC board

→ 12 uTCA crates for charge readout + 1 uTCA crate for light readout

19

- White Rabbit developed for the synchronization of CERN accelerators chain offers sub-ns synchronization over ~10 km distance, based on PTP + synchronous Ethernet scheme previously developed in 2008 (http://arxiv.org/abs/0906.2325)

- White Rabbit chains can be now set up with commercial components:
  - Network based on Grand Master switch
  - Time tagging cards for external triggers
  - Slave nodes in piggy back configuration to interface to uTCA

- Transmission of synchronization and trigger data over the WR network + clock

- Slave uTCA nodes propagate clock + sync + trigger signals on the uTCA backplane, so that the FE digitization cards are aligned in their sampling, can know the absolute time and can compare it with the one of transmitted triggers

- FE knows spill time and off spill time and can set up different operation modes

- Trigger timestamps may be created by beam counters, cosmic counters, light readout system in uTCA

- The White Rabbit network provides the distribution of a common time base to align all the elements of the DAQ system: the 400 ns sampling on the uTCA digitization boards of the charge readout and the light readout digitization boards

- The White Rabbit network is also used to distribute triggers (timestamps of trigger signals in this common time base) to the elements of the DAQ chain. The uTCA digitization boards have a large circular memory buffer (larger than the drift time) and associate to the drift window the samples starting from the time stamp of the triggers

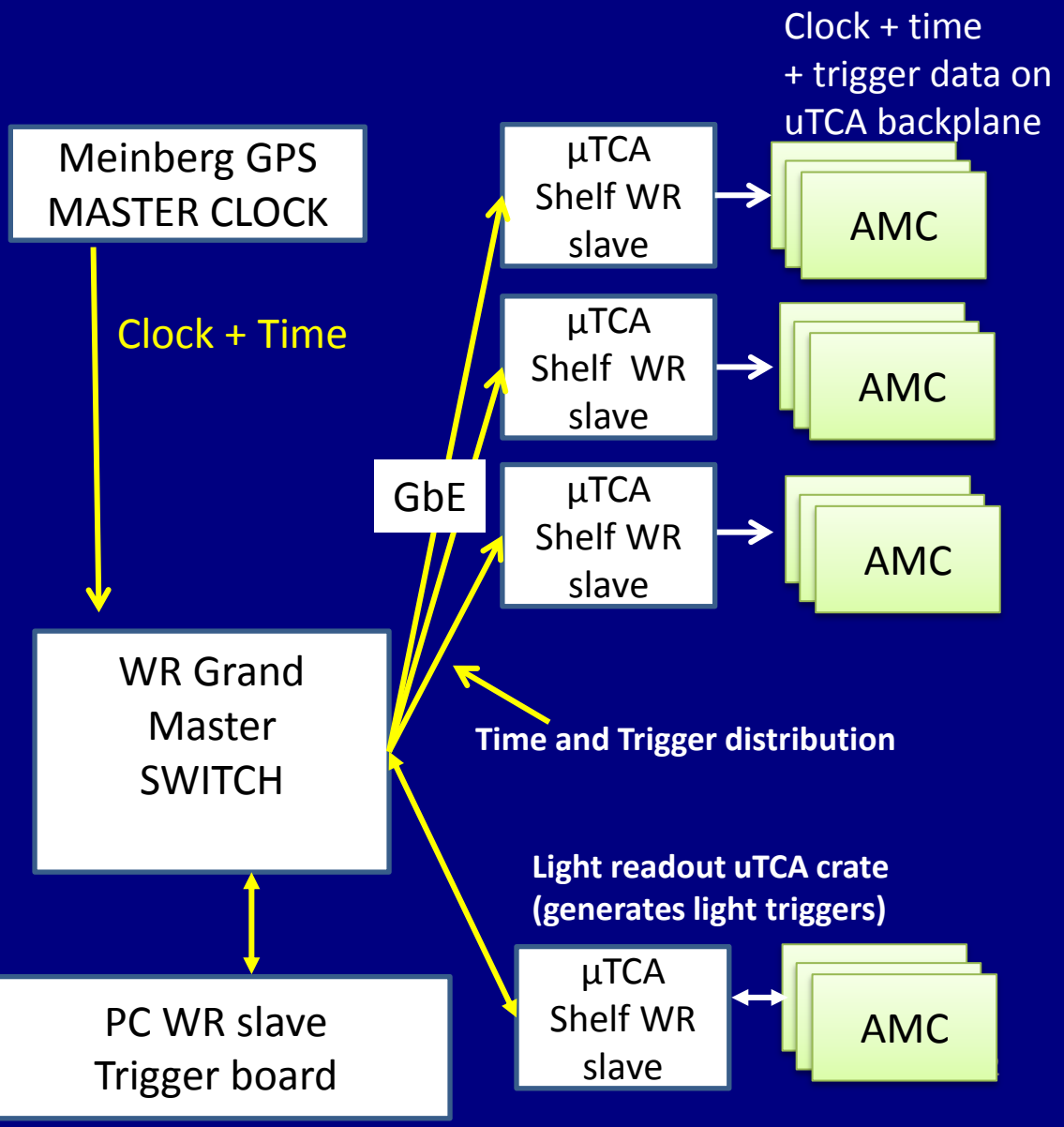
- Triggers are created and injected in the network by a WR timestamping board in the trigger PC. This looks at 3 input logic signals (connections via LEMO cables):

a) The start of spill signal (the FE needs to know if it is taking data during or out of the spill in order to deal with different triggers and set different sampling modes of the LRO)
b) Beam trigger (from the scintillators along the beam line)
c) Cosmic ray taggers

The LRO triggers should be directly injected in the WR network

White Rabbit based Time/trigger distribution scheme

- **No need to develop analog clock distribution system and microTCA receiver cards**
- **Beam counters/large area cosmic counters trigger board also in WR standard → generates trigger timestamps transmitted on WR network**
- **Light readout DAQ uTCA → generates trigger timestamps transmitted on WR network**
- **Development of the WR slave as MCH mezzanine from a commercial WR node**
- **Sub-ns sync accuracy**



Commercial components for the WR network:

- **WR Grand Master Switch WRS-3/18**

18 SFP fiber cages, 1GbE
Supports connections up to 10 km distance
Connected to Meimberg GPS as time source

- **PC time-tagging trigger board**

SPEC carrier board+ Fine Delay FMC

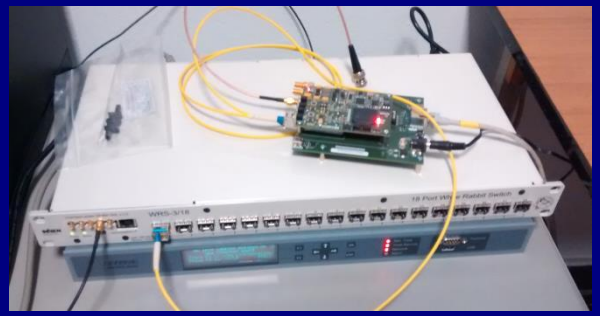
Can time-stamp with 1ns accuracy beam trigger signals or external large area scintillator counters → Trigger time-stamps data are transmitted over the WR network to the microTCA MCH

FMC Fine Delay 1 ns 4 channels

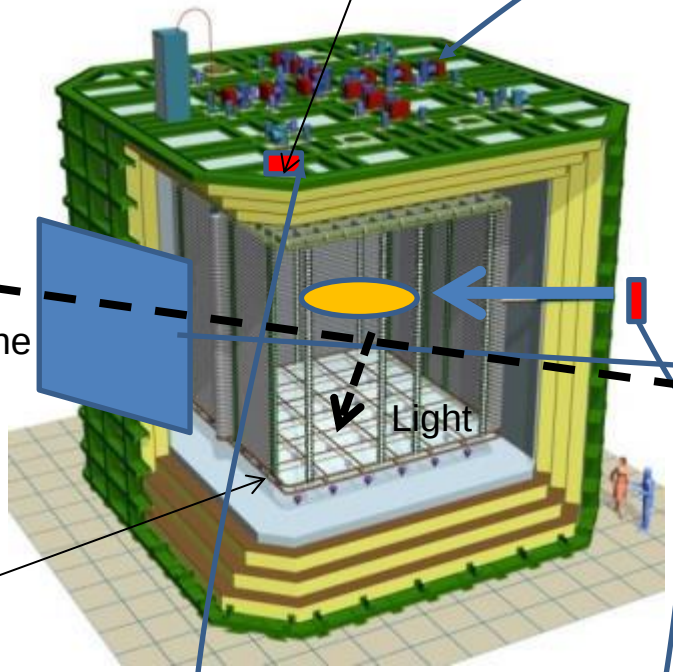
SPEC FMC PCIe carrier V4

- **WR-LEN (White rabbit Light Embedded Node)** →
  WR slave node cards basis of WA105 development for a MCH WR uTCA mezzanine

Test-bench with Meimberg + switch + slave node



DAQ timing-trigger integration



uTCA crate of LRO

x12 Charge readout WR links

Trigger counter

Cosmic muon

Start of spill signal (cable)

External trigger plane for cosmics

Light

Crate with NIM logic

PMTs for light readout

Light readout WR link

PC with WR time stamping card

Time stamps:

- **Beam trigger**
- **Start of spill**
- **External cosmic trigger**
- **Calibration pulses**

WR switch (optical links 1 Gbps)

Beam triggers from SPS
Tertiary beam line ~100Hz

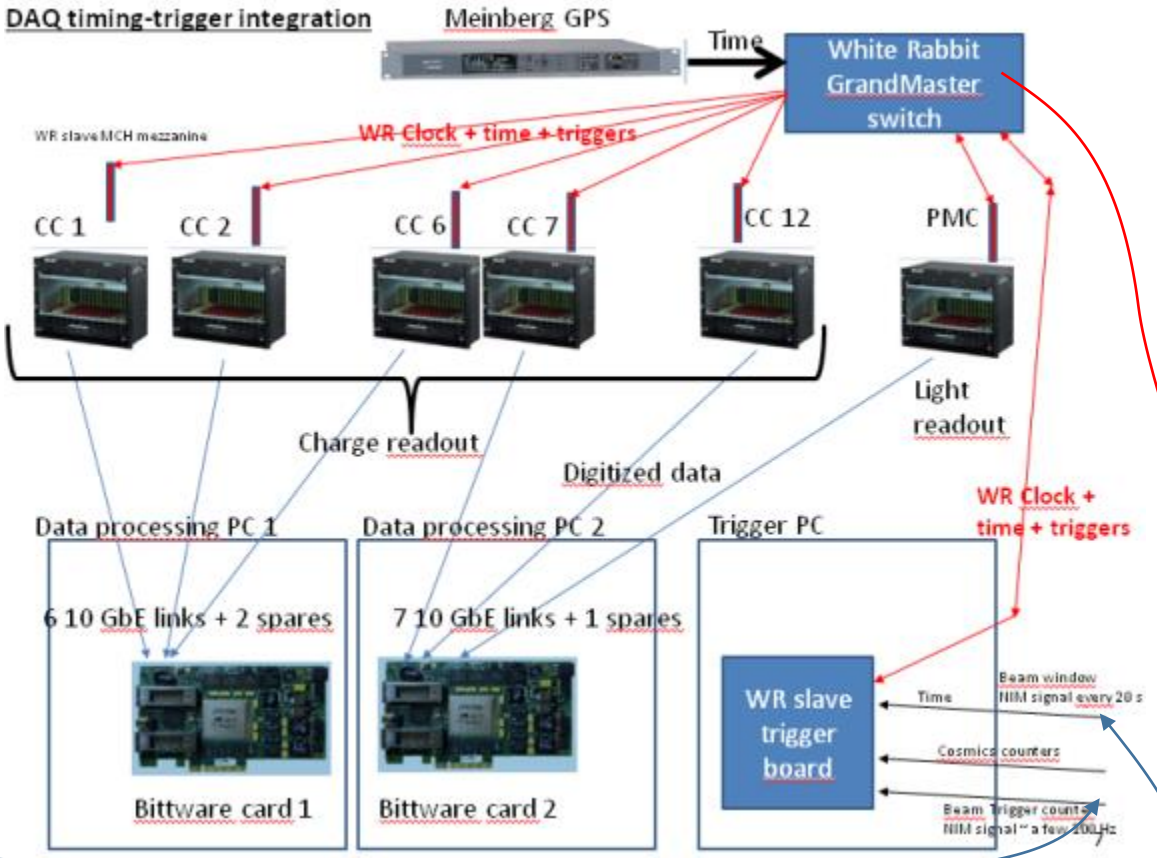
The DAQ system of the beam instrumentation will also be triggered by the NIM signal from trigger scintillators and will take data for the different ADC/TDC/IO registers (in VME) used to read the beam instrumentation. The NIM crate which defines the coincidence for the NIM trigger will have a fan-out in order to distribute the trigger signal to both the local beam DAQ PC, the beam DAQ crate and to the ProtoDUNE-DP trigger PC

The DAQ VME system will have also a WR time tagging card installed which looks at the beam trigger. This WR card is connected via an optical fiber to the WR grand master of WA105 so that it is aligned on the common time base and it is read out with the beam DAQ. Once the beam DAQ system reads the event from the beam instrumentation it will also read the timestamp of the beam trigger and associate to the event structure

The beam DAQ will write the beam trigger data from the local beam instrumentation DAQ which includes the time tagging WR card on the beam instrumentation database

The online computing farm (or any offline process) can access the beam instrumentation database in order to fish the beam instrumentation data related to a given timestamp for a ProtoDUNE-DP trigger

From the beam line to protoDUNE-DP two cables one for the beam trigger signal and one for the start of spill signal. From WA105 to the Beam DAQ we will deploy an optical fiber with the WR network connection → General scheme in the next page:



Beam instrumentation detectors

electrical signals

NIM crate for trigger logic

Beam trigger logic signals

VME crate for beam DAQ

WR card

SPS start of spill

WR network

Beam data output

Beam instrumentation DB

queries

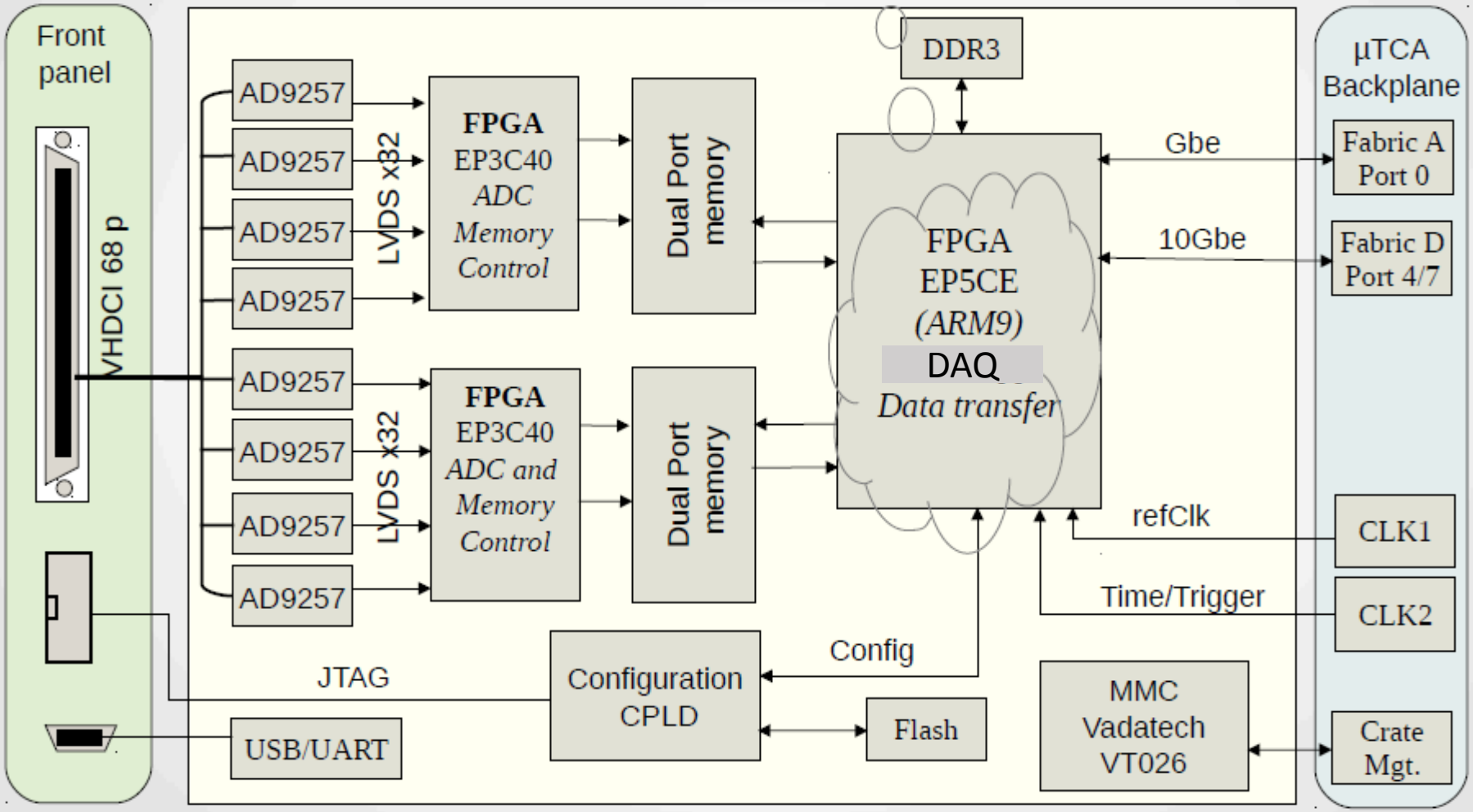
Protune-DP offline or online processing

AMC digitization card (2013)

- Board characteristics
  - µTCA standard (double width , full height)
  - 64 input channels (2V / 14 bits (12 bits effective) / 2.5 Msps to 20 Msps)
  - Control through MCH 1GbE backplane link port 0/1
  - Data transmission through 10GbE backplane and MCH 10GbE SW
  - Local buffer dual port memory

AMC card design : block diagram 2013

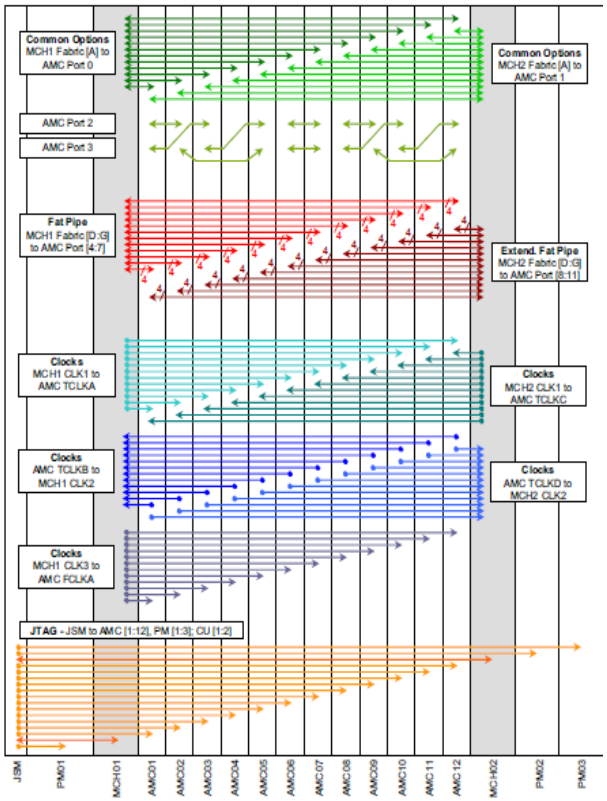
Under design



µTCA crate

- Schroff reference : 11850-015, 8 U µTCA Shelf, 12+2+3+1 slot for AMC double Mid-size modules.
- 19” rack mountable
- 12 AMC Double Mid-size slots
- 2 redundant MicroTCA Carrier Hub (MCH) slots (Double Full-size)
- 2 Power Module (PM) slots (6 HP Double) at the right side
- 1 Power Module (PM) slots (12 (9) HP Double) at the left side
- 1 slot for a JTAG module (Double compact)
- 5 splitting kits to install single module in a double slot

Crate

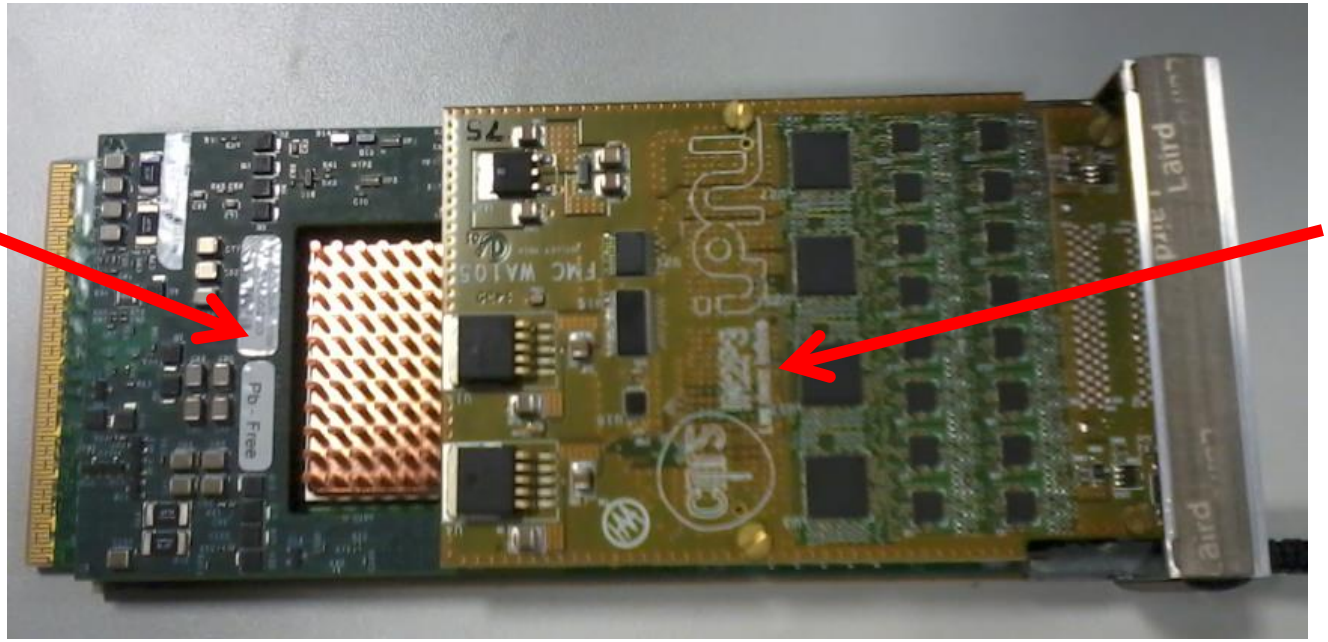


Backplane technology

Data acquisition demonstrator

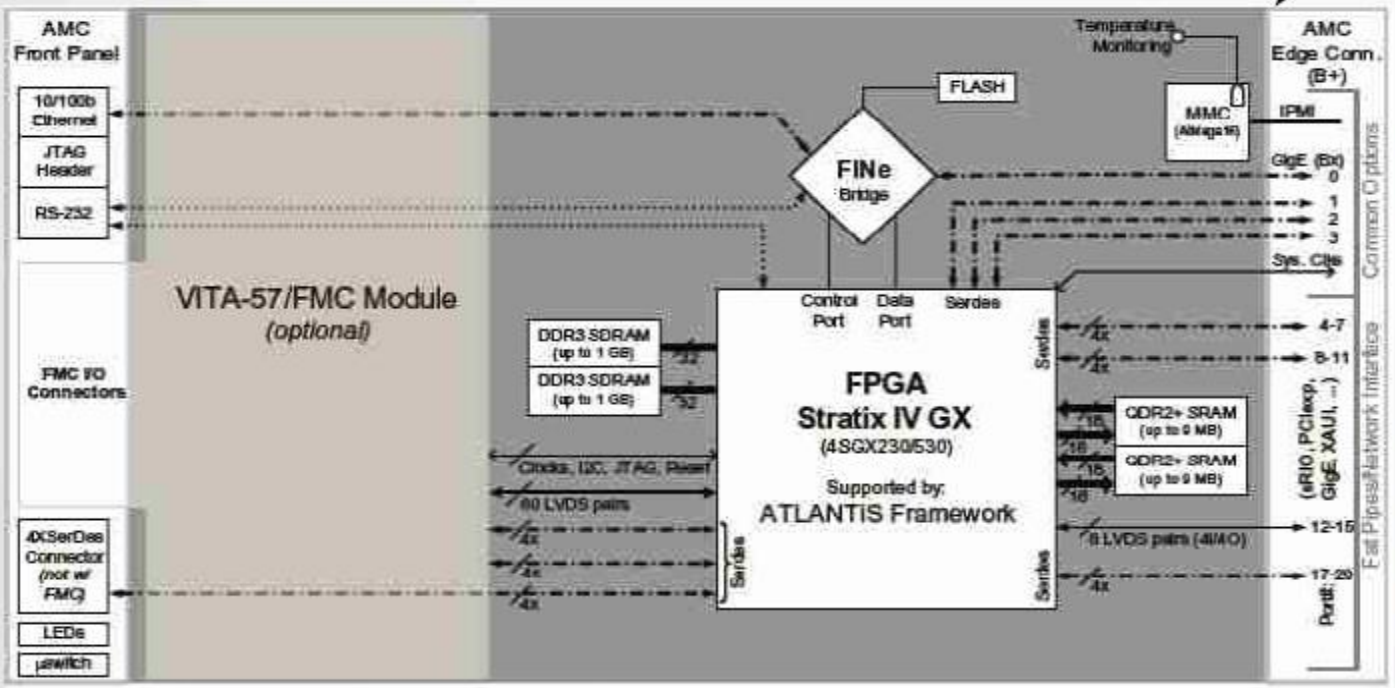
(submitted in fall 2014 available since January 2015)

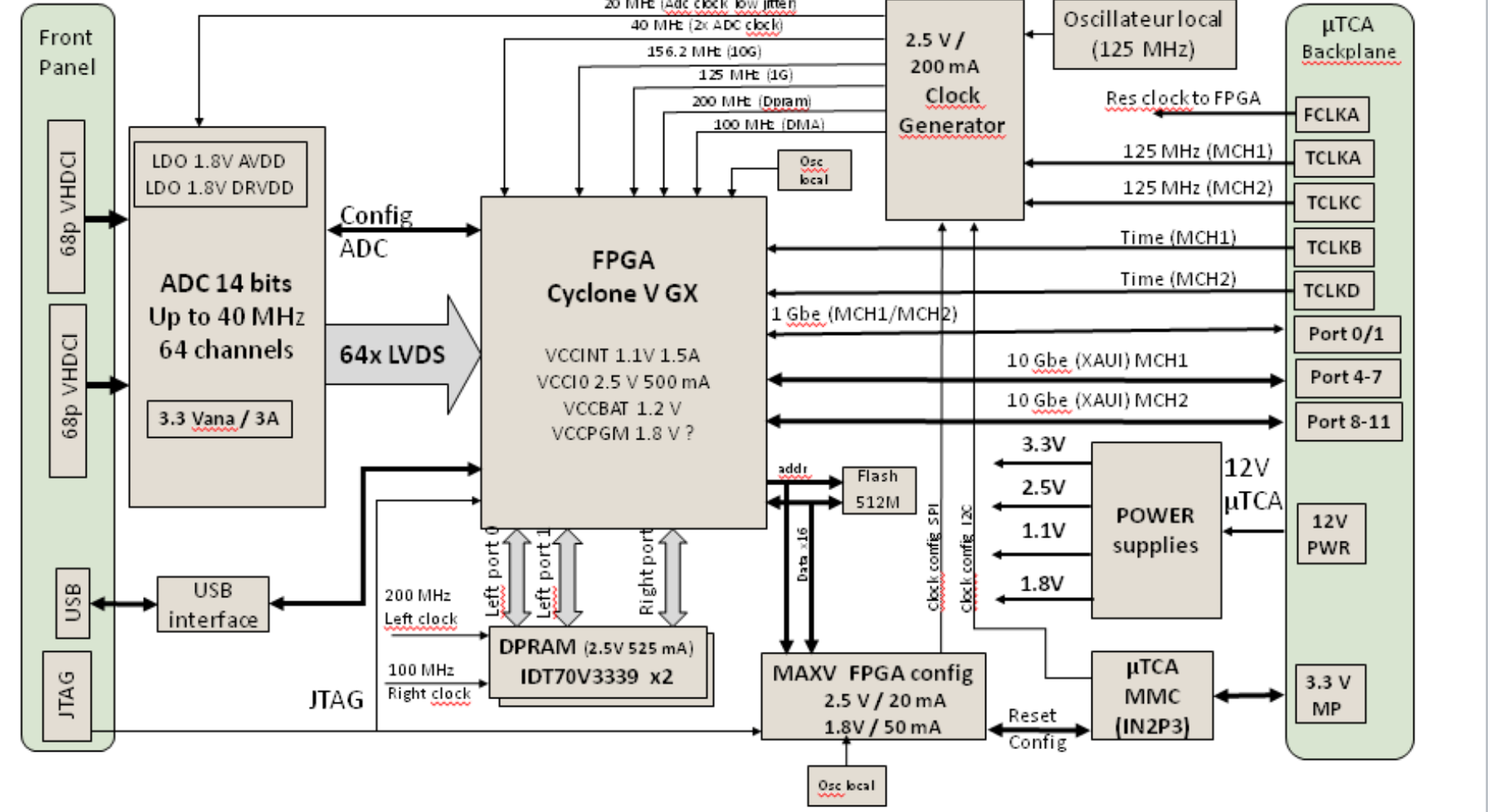
- Based on µTCA AMC S4AM from Bittware
- FMC mezzanine board with 64 ADC channels
- Control through GbE
- Data transmission through 10GbE



AMC card design : work in progress 2015

1- One full model presently developed on a Stratix IV AMC board (Bittware S4AM):
- "recycling" of the NIOS modules
- management of the 64 ADC channels @ full speed
- tests of the 10Gbe (FPGA side & MCH, "public" UDP IP)
- full evaluation of the number of logic elements required for the final AMC card





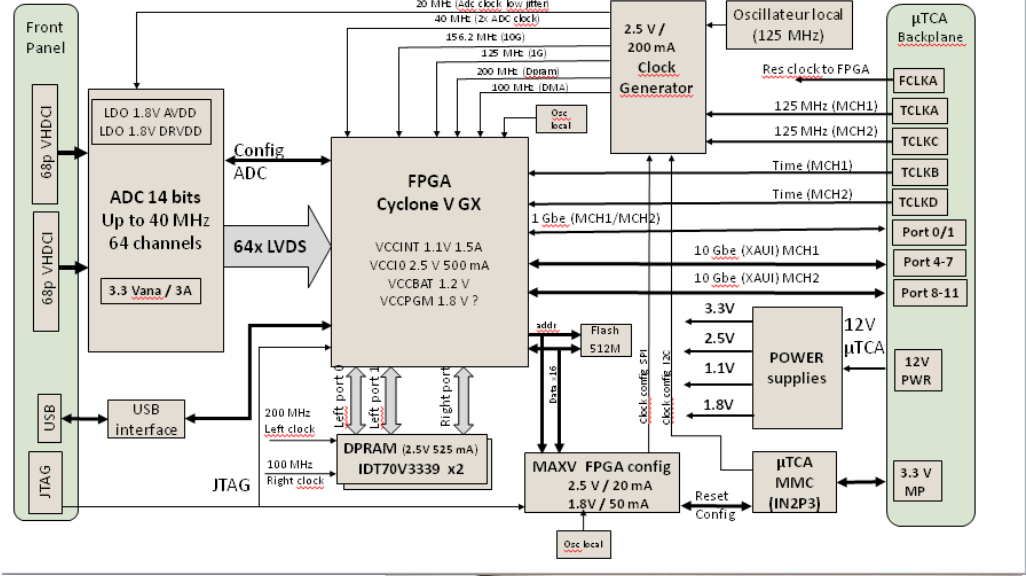
FINAL AMC card:

Production of uTCA AMC cards for the 6x6x6 started at the end of 2015, first batch deployed on 3x1x1

First 64 ch AMC digitization card delivered at the end of July 2016

(2.5-25 MHz, 12 bits, 2V, ADC AD5297, 10 GbE output on uTCA backplane)

uTCA DAQ system:



64 channels AMC digitization cards (2.5-25 MHz, 12 bits output, 10 GbE connectivity

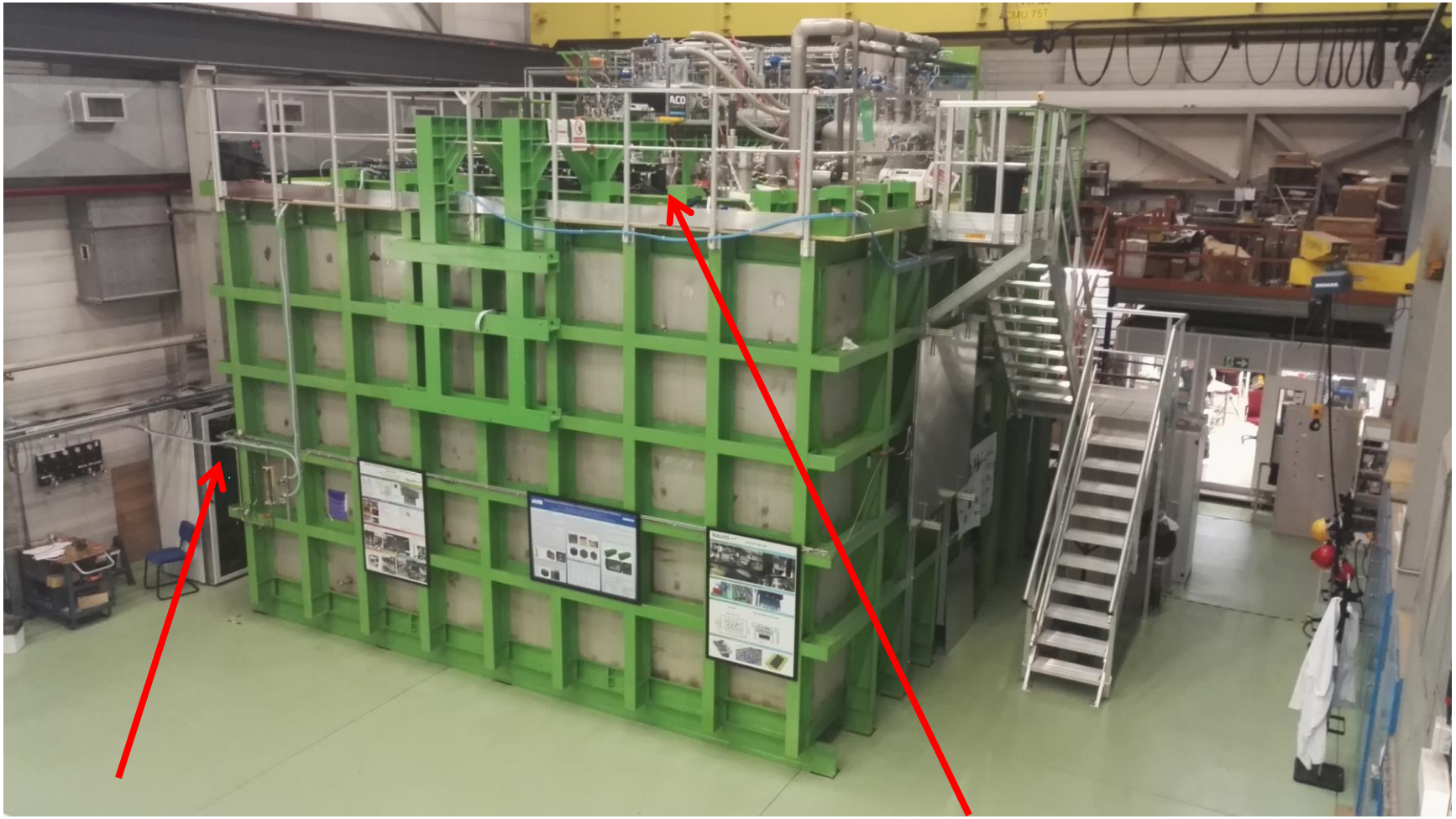
- Demonstrator card with 64 ADC channels built and tested in 2015 for the definition of the final card
- Purchase of main components (ADCs, FPGAs, IDT memories) of the final cards by end of 2015 to equip the entire 6x6x6
  - Final design of digitization PCBs May 2016
  - First assembled cards received in August 2016.

- 20 cards produced by September 2016 to equip the 3x1x1
- Cards production going to be completed with the 2017 budget of remaining 100 FE and uTCA cards for 6x6x6 (main components available)
- The warm flange PCB design for the 6x6x6 is based on an extension of the one of the 3x1x1, routing under finalization

3x1x1 implementation

6x6x6: 12 uTCA crates (120 AMCs, 7680 readout channels)

→ 3x1x1: 4 uTCA crates (20 AMCs, 1280 readout channels)



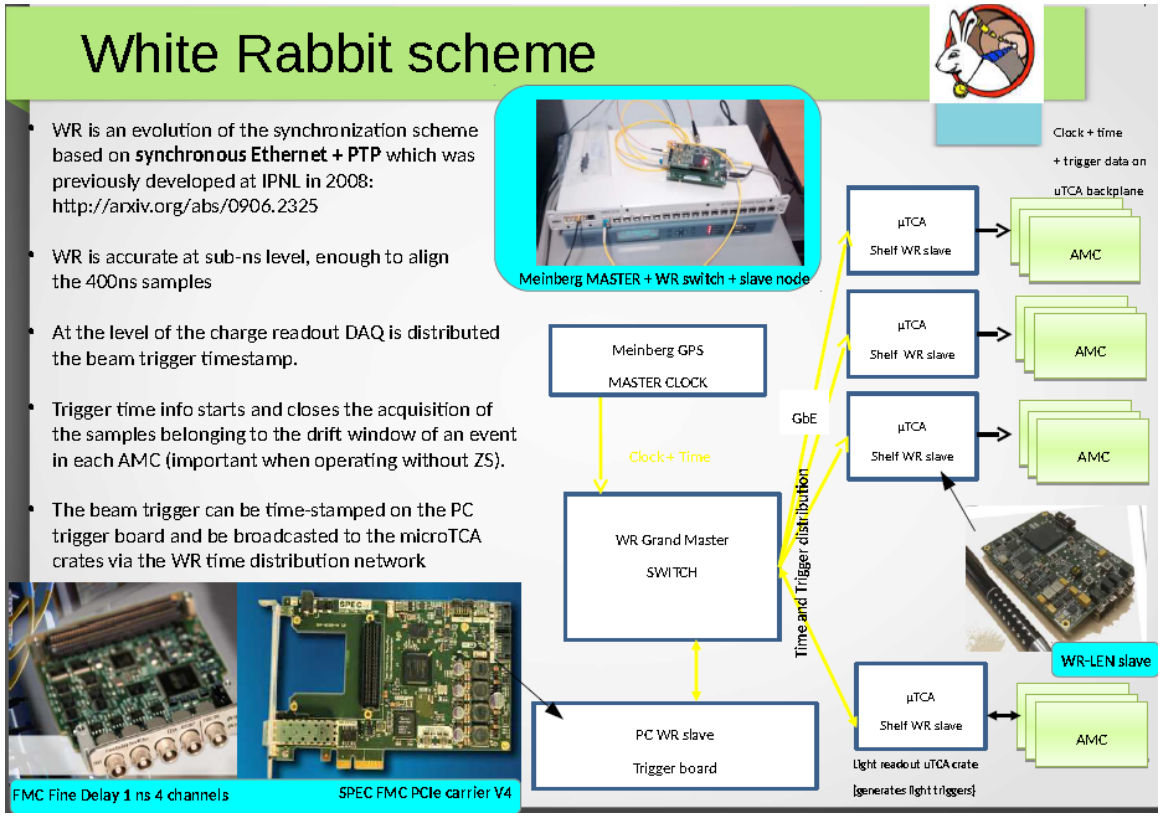
Event builder, network, GPS/White Rabbit GM, WR Trigger PC

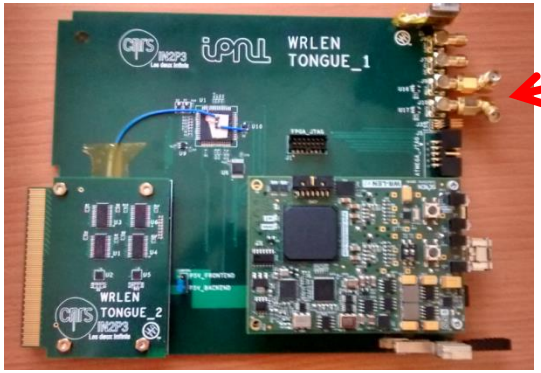
Signal Chimneys and uTCA crates

White Rabbit trigger time-stamping PC (SPEC + FMC-DIO)

White Rabbit Grand-Master

GPS unit

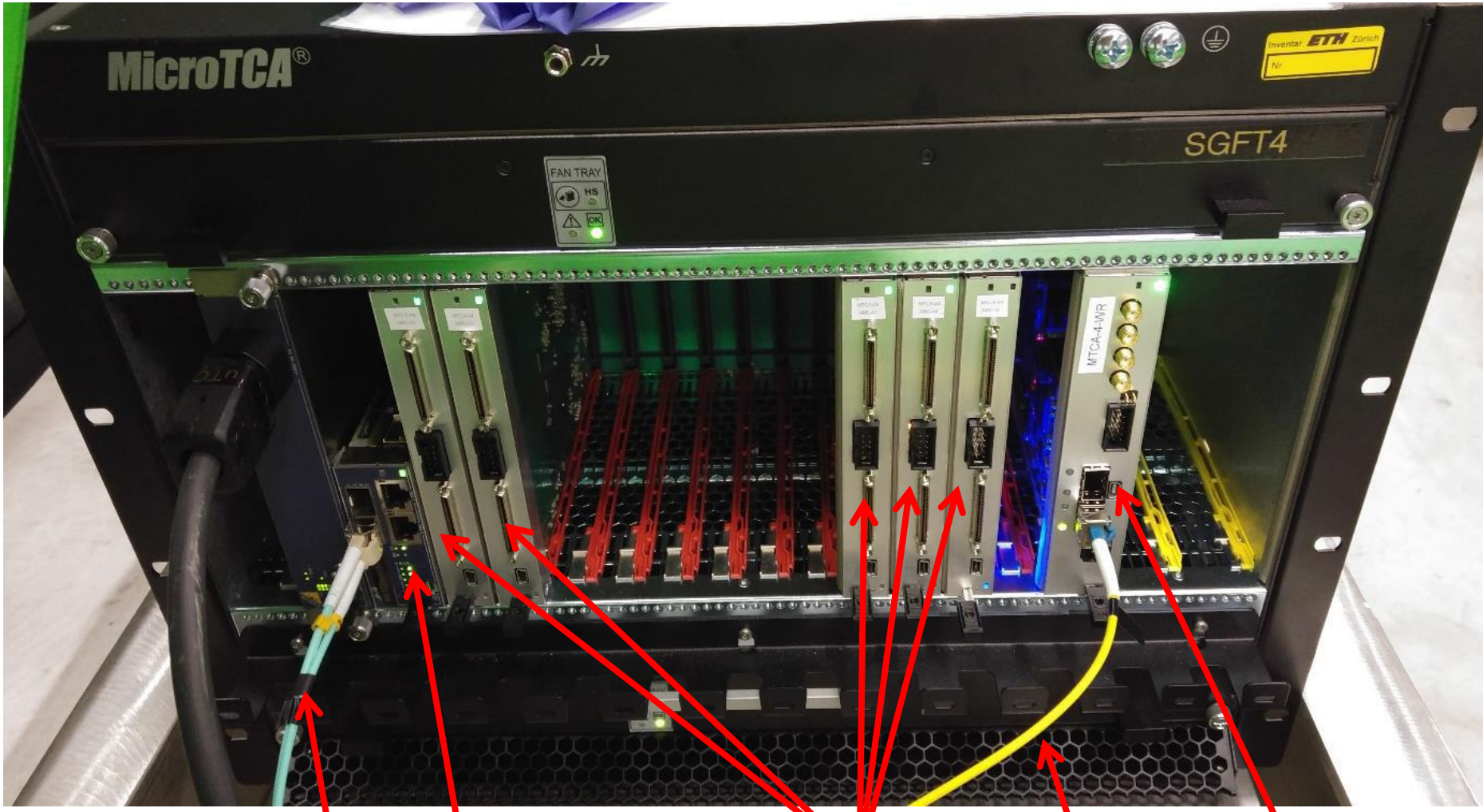




White Rabbit uTCA slave node based on WRLEN developed and produced for entire 6x6x6

Other components of the chain (GPS receiver, WR grandmaster, SPEC+ FMC-DIO + 13 WRLEN) available commercially

How a crates was looking like before VHDCI signals cabling to the warm flange



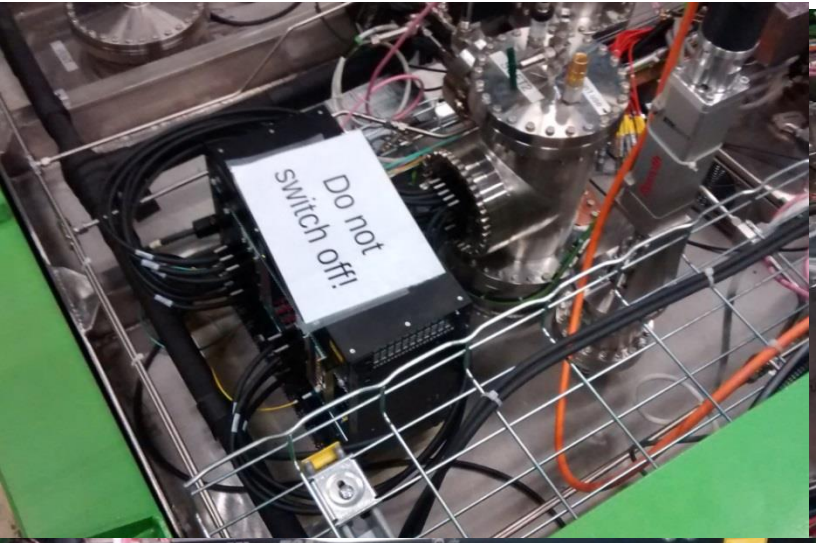
MCH

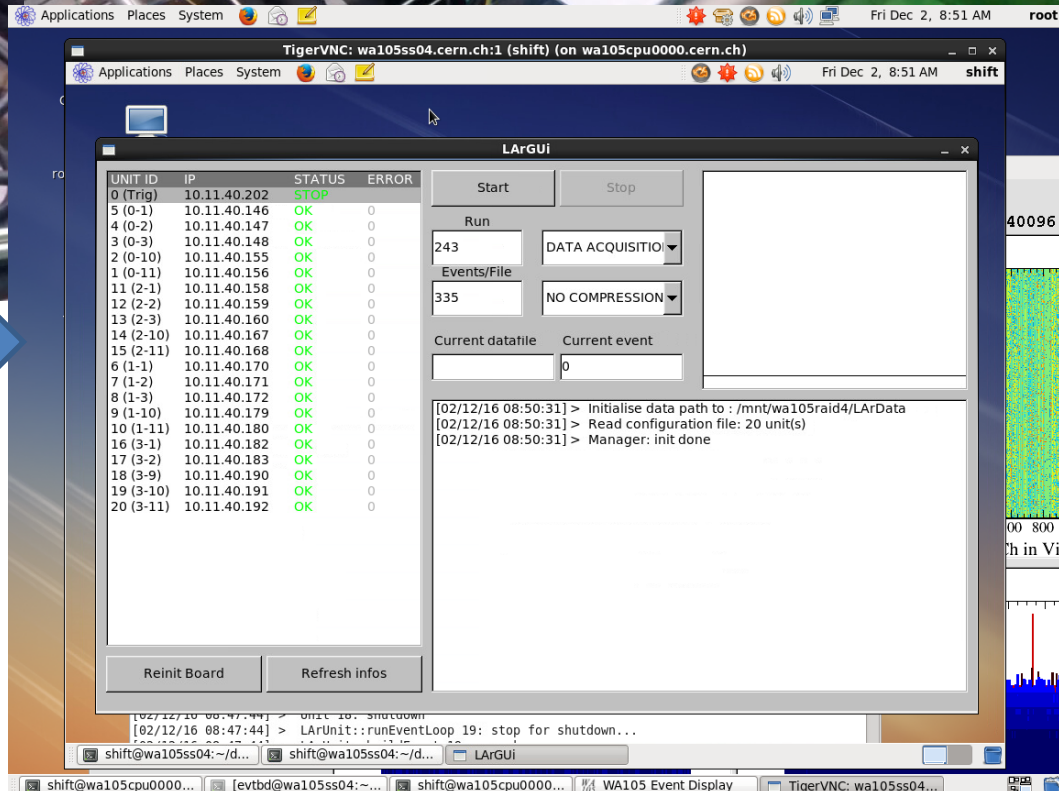
AMC 64 channels digitization cards

WR uTCA slave card node with WRLEN mezzanine

10 Gbit/s data link

White Rabbit optical link





Run control with 20 AMCs

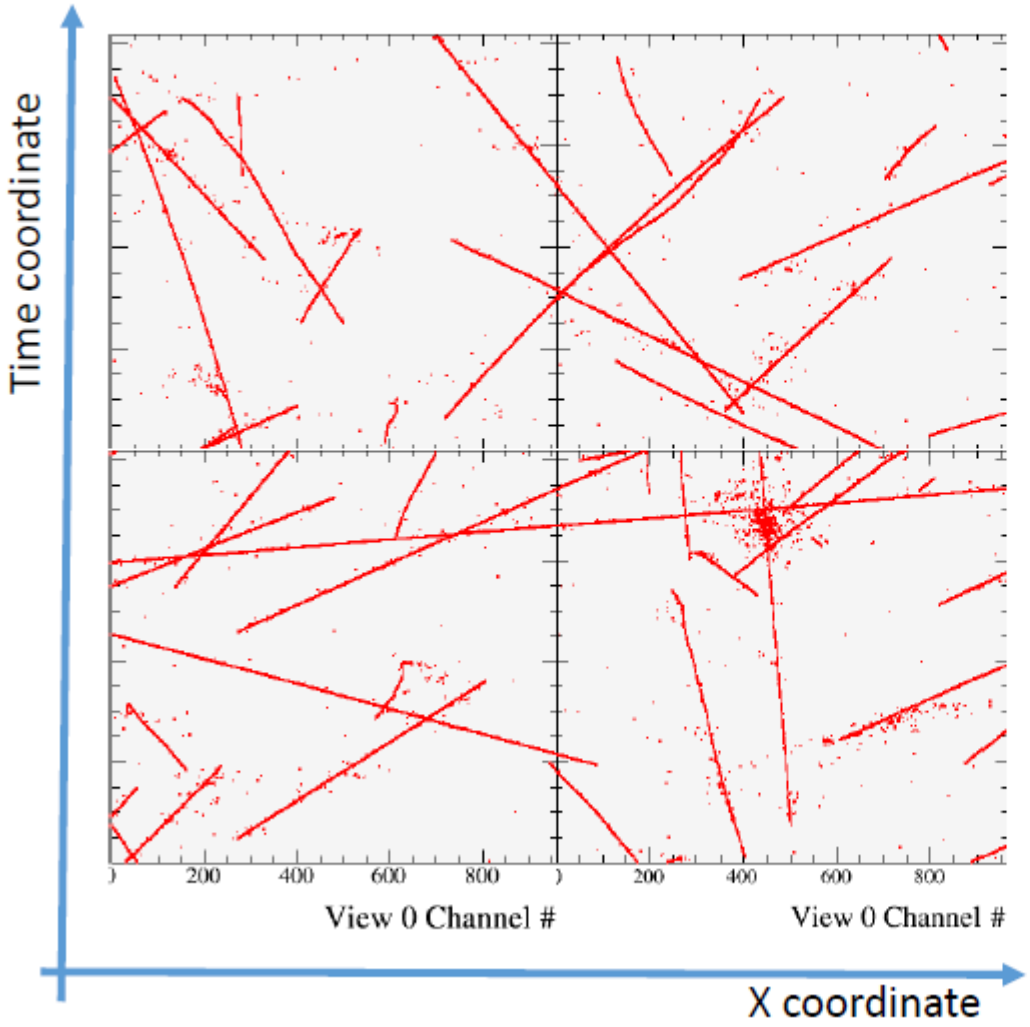
Automatic data processing on online storage/processing farm for purity and gain analysis + data transfer on EOS

Stable system, noise conditions at warm 1.5-1.7 ADC counts RMS

Typical event signature for ground surface Liquid Ar TPC operation

For each beam trigger we can have on average 70 cosmics overlapped on the drift window after the trigger (these cosmics may have interacted with the detector in the 4 ms before the trigger and in the 4 ms after the trigger → chopped tracks, “belt conveyor” effect

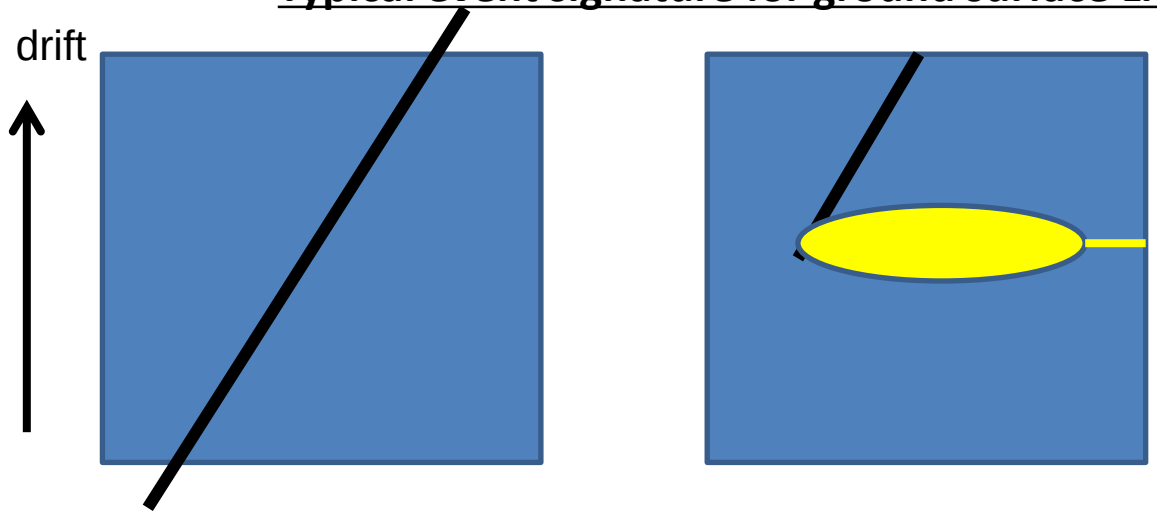
In-spill cosmics in charge data



Example of cosmics only event (in one of the views)

- Red points are reconstructed hits
- TPC is readout in 4 3x3m^2 modules
- After track reconstruction:
  - Attempt to correlate found tracks with light data
  - Remove CR background from beam event
  - Select a subsample of long tracks for calibration purposes

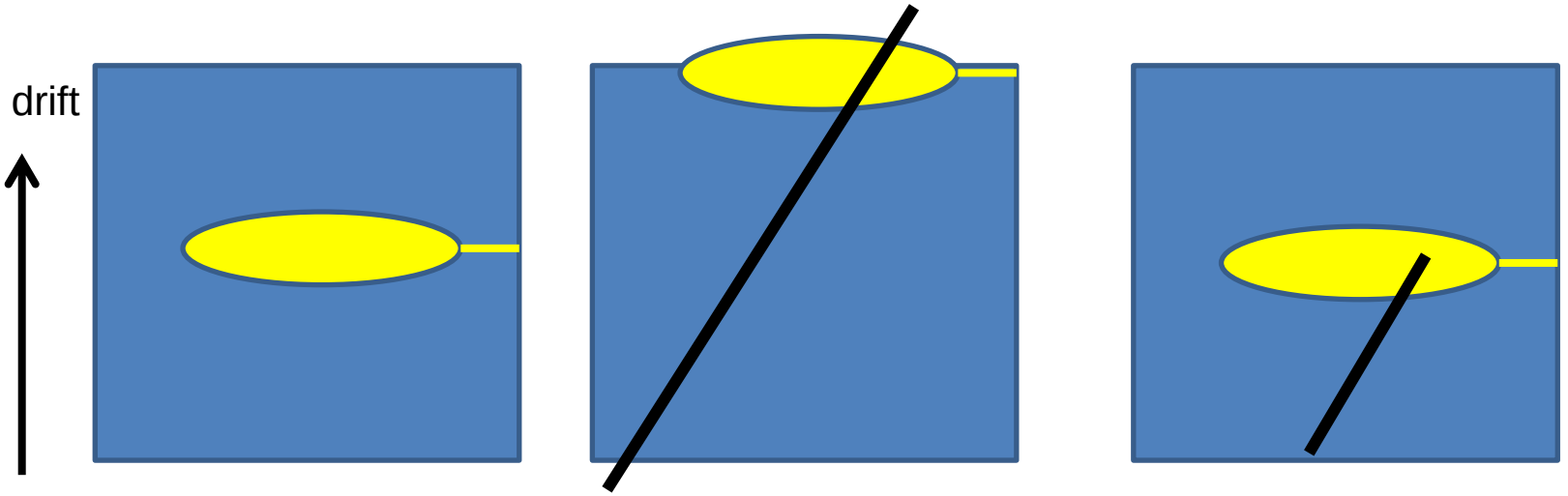
Typical event signature for ground surface Liquid Ar TPC operation



t=beam trigger - 2 ms

t=beam trigger → reconstructed event

The « belt conveyor » effect +- 4 ms around the beam trigger time



t=beam trigger

t=beam trigger + 2 ms

reconstructed event

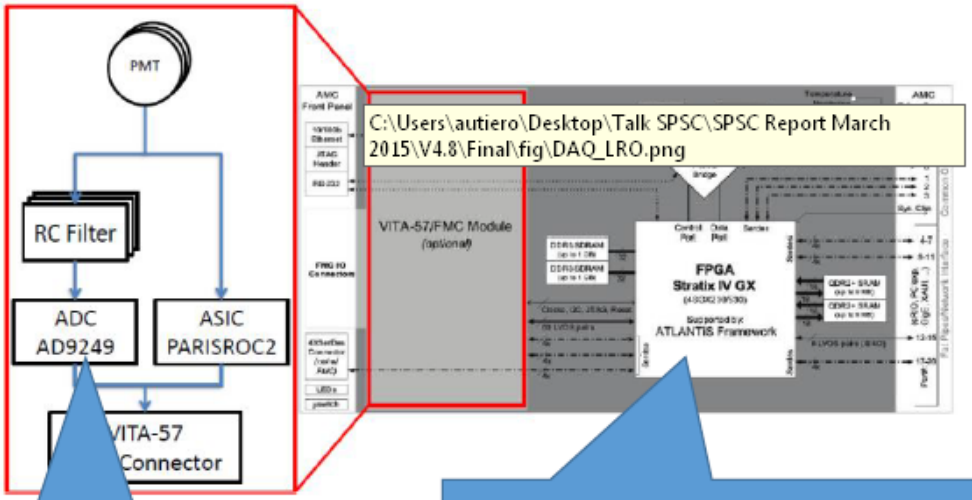
→ During spills it is needed a continuous digitization of the light in the +-4 ms around the trigger time (the light signal is instantaneous and keeps memory of the real arrival time of the cosmics)

→ Sampling can be coarse up to 400 ns just to correlate to charge readout

Light readout electronics

Two modes of acquisition:

- External beam trigger to acquire ±4ms around the spill
- Internal trigger from PARISROC2 ASIC to acquire short time segments



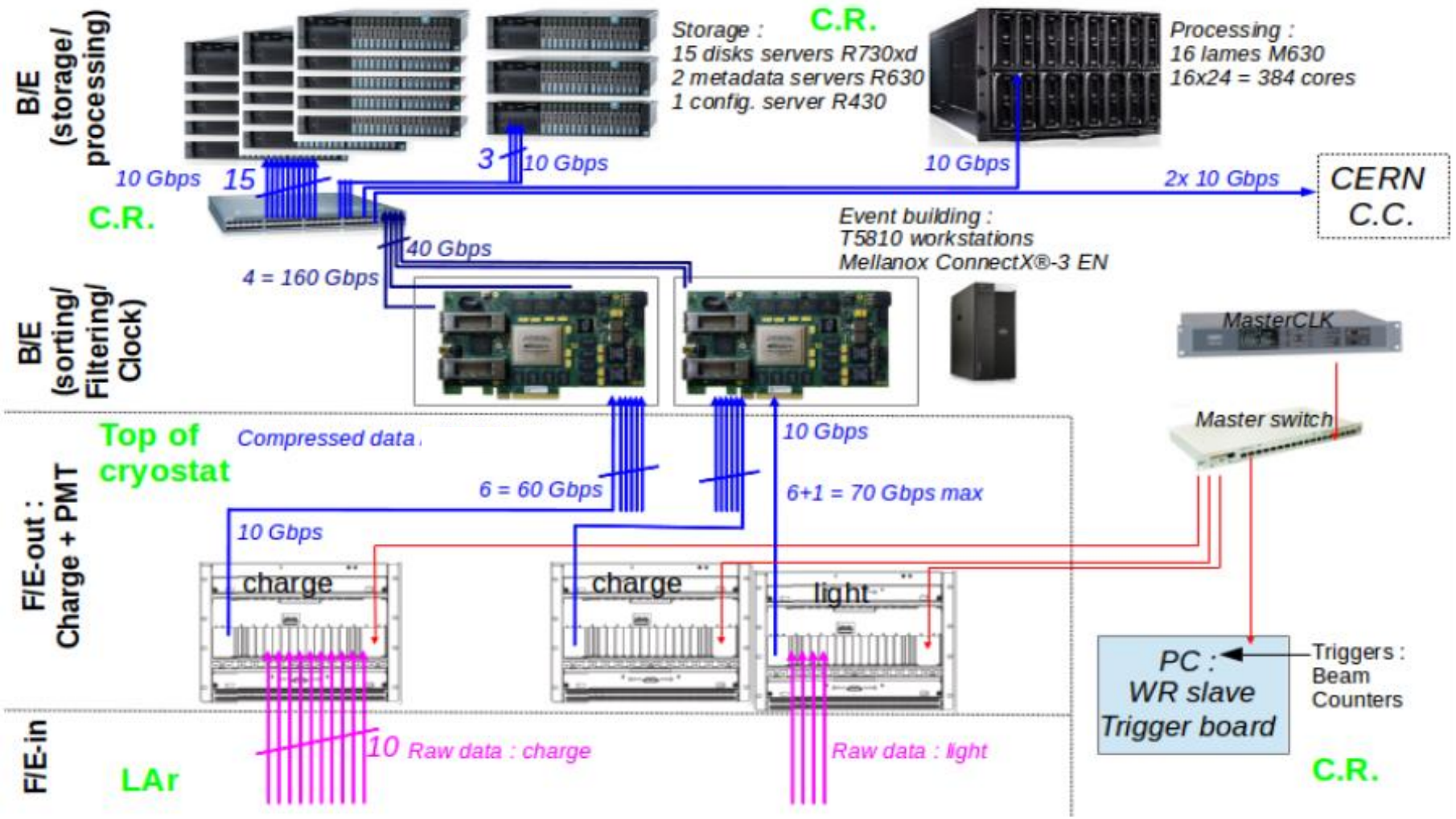
Digitizer: nominally runs at 40 MHz, 14 bits

Digitizer data is buffered in 1G memory buffer connected to FPGA Averaged for a multiple of 40MHz (to reduce the data volume)

→ Sum 16 samples at 40MHz to get an effective 2.5 MHz sampling like for the charge readout

The LRO card has to know spill/out of spill
Out of spill it can define self-triggering light triggers when “n” PMTs are over a certain threshold and transmit its time-stamp over the WR

Online processing and storage facility: internal bandwidth 20 GB/s, 1 PB storage, 384 cores: key element for online analysis (removal of cosmics, purity, gain, events filtering)



C.R. stands for Counting Room

Data size

- Data are expected to be taken without zero skipping and exploiting loss-less compression and the system has been designed to support up to 100 Hz of beam triggers without zero-skipping and no compression

- 7680 channels, 10k samples in a drift windows of 4ms→ 146.8MB/events, No zero skipping

- Beam rate: 100Hz

- Data flow= 14.3 GB/s (without compression), 1.43 GB/s (with compression)

Huffman lossless compression can reduce the non-zero-skipped charge readout data volume by at least a factor 10 (S/N for double phase ~100:1, small noise fluctuations in absolute ADC counts)

- Light readout does not change in a significant way this picture (<0.5 GB/s)

→ Integrated internal local DAQ bandwidth on the “20 GB/s scale” in order to have a robust safety factor for concurrent read/write

Local data buffer ~ 1000TB (no zero skipping, no compression), also used for local processing

- 100 M triggers expected to be taken in 120 days of beam time in 2018

- If totally stored in non-zero-skipped, lossless compression format
  (assuming Huffman, factor 10 compression: 15MB/event) → 2.4 PB
  + cosmic runs and technical tests

- Requested link from online-storage to CERN computing division at 20 Gbps, compatible with 100 Hz non-zero-skipped, Huffman compressed (factor 10) data flow.

- This link would allow to transfer the entire beam triggers data volume with a typical occupancy of less or equal than 80%.

- The availability of a large local buffer allows as well to release the disk cashing requirements at the other end of the data link at the computing division being consistent with a dilution of the beam data transfer over the periods during which the experiment is not having beam time.

Online storage/processing farm motivation :

SPSC report, April 2016

6x6x6

The local bandwidth of 20 GB/s also allows comfortable concurrent reading and writing access to the compressed data on the local storage system for online analysis. Data transfer to the IT division should happen by clustering the events in files having dimensions of a few gigabytes. This file size is needed for an efficient storage on the Castor system at the computing center. The online storage facility has also the task of buffering the events and formatting them for transfer on this typical file size.

In addition to the storage buffers requirement described above, the online storage processing farm allows for the following functionalities:

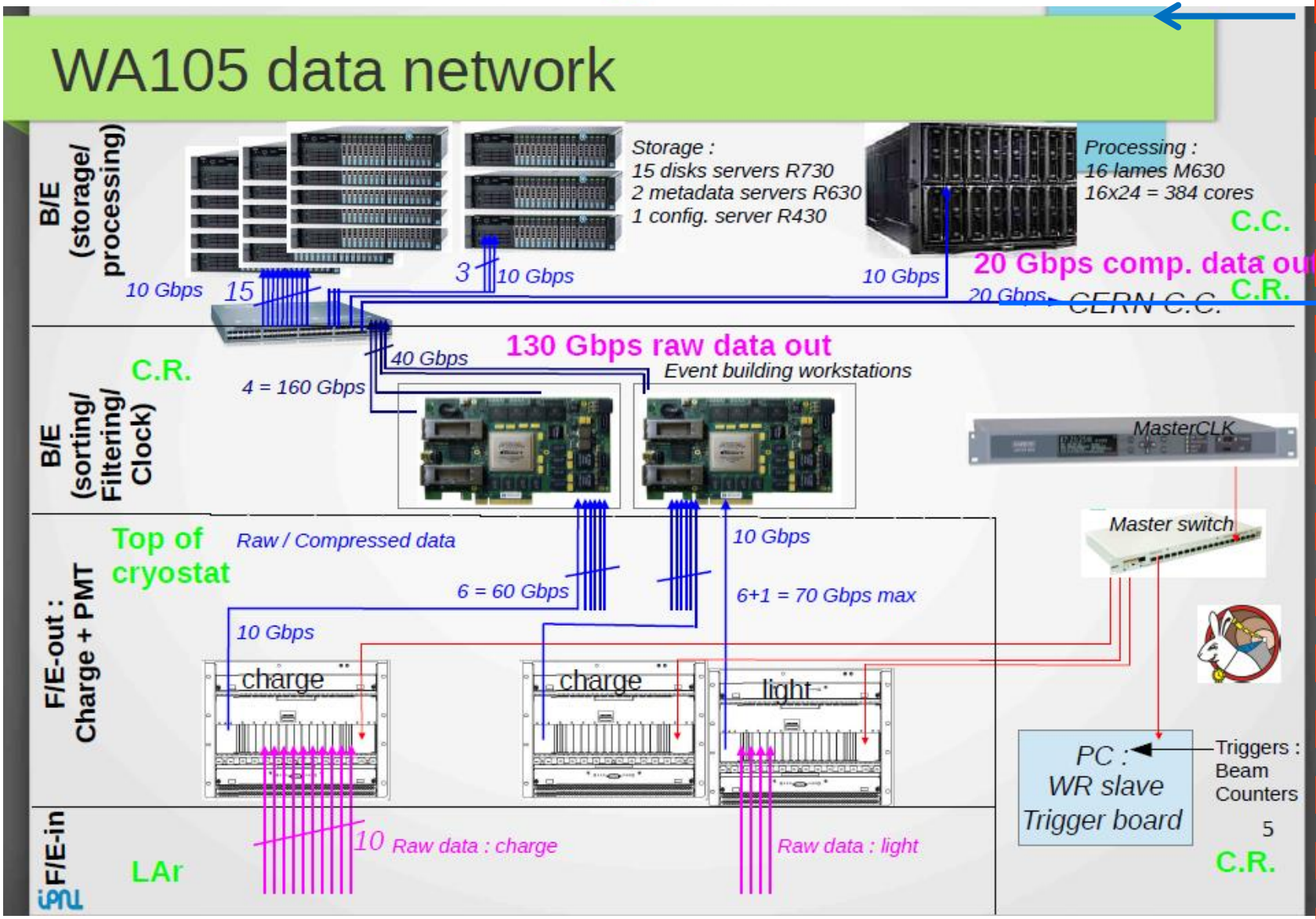
- Completion of event building by connecting the data flows of the two back-end systems
- Fast event reconstruction and disentangling of cosmic rays tracks segments by using also the LRO timing information
- Selection of a subsample of the cosmic ray tracks overlapped to beam events for online purity analysis and detector gain monitoring
- General online data quality checks
- Events filtering and formatting for final storage

Online processing and storage facility: internal bandwidth 20 GB/s, 1 PB storage, 384 cores: key element for online analysis (removal of cosmics, purity, gain, events filtering)

online

offline

WA105 data network



Network back-end FPGA board for online processing and event building

Bittware S5-PCIe-HQ processing card :

- High density Altera Stratix V GX/GS FPGA
- PCIe x8 interface supporting Gen1, Gen2, or Gen3
- Dual QSFP+ cages for 40GigE or 10GigE direct to the FPGA for lowest possible latency

⇒ up-to 8 x 10Gbe links w/o data loss

- Up to 16 GBytes DDR3 SDRAM
- Up to 72 MBytes QDRII/II+
- Two SATA connectors

- Time-stamping support
- Utility I/O includes:

USB 2.0, RS-232, and JTAG

- Hosted in standard PCs
- OpenCL supported

- Applications : video compression, image processing etc.

Purchased in September 2013

FPGA board for online processing 2015

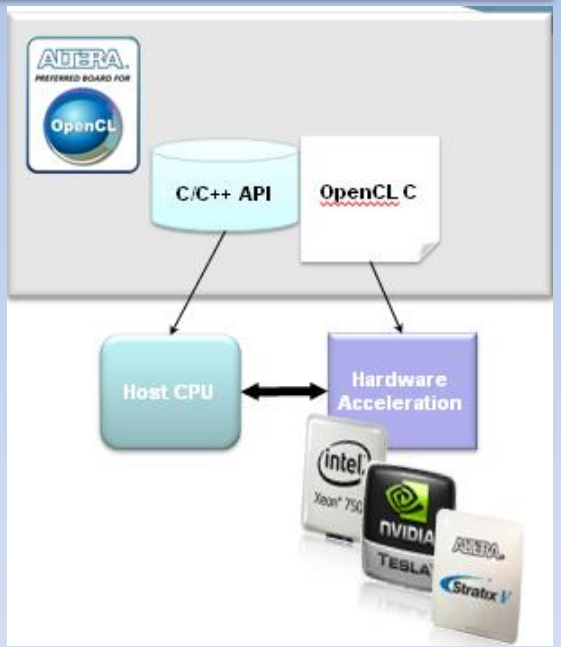
Bittware S5-PCIe-HQ processing card : upgrading from S5PHQ-D5 to S5PHQ-D8

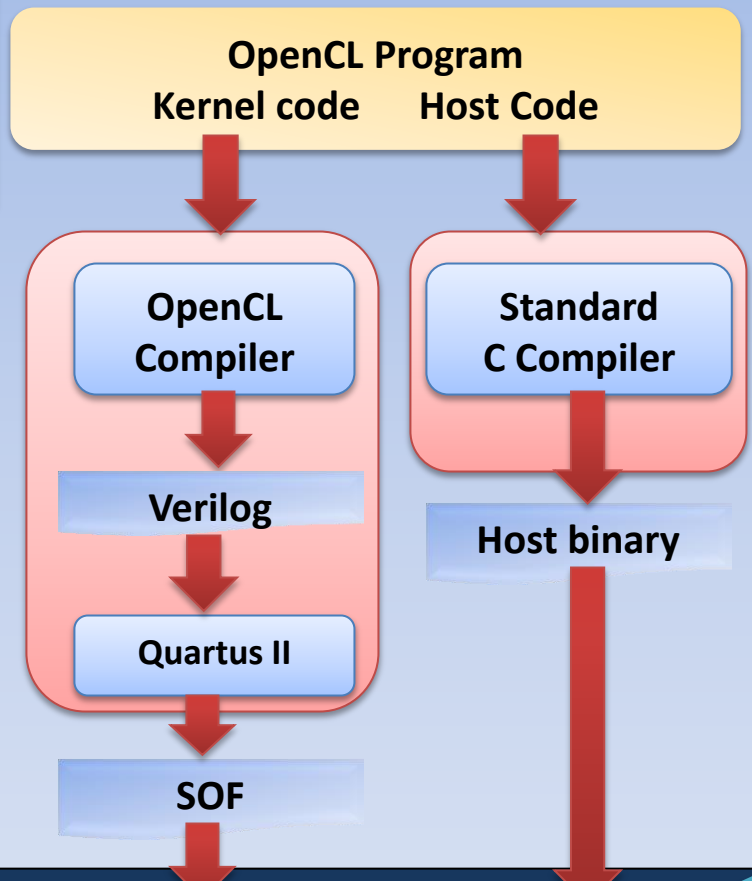
- Ruggedization Options: Commercial (0C to 50C)
- S5 Family, HardIP, Size: Stratix V GSED8
- S5GXBSpeed: 2
- DDR3BankA&B: 4GB 800MHz
- QDR2 Bank A&B&C&D Size: 18MB
- QDR2 Bank A&B&C&D Type/Speed: QDR2+ 550MHz
- Anemone: None
- Oscillator speed : Std
- Front Panel Options:
  No SMA
- Heatsink: FPGA Fansink

What is OpenCL ?

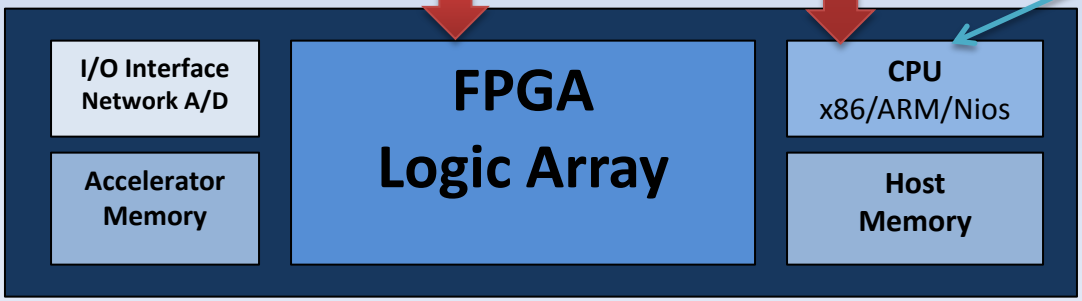
Community for network applications starting up

```
__kernel void
sum(__global float *a,
    __global float *b,
    __global float *y)
{
  int gid = get_global_id(0);
  y[gid] = a[gid] + b[gid];
}
```






```
main() {
   read_data( … );
   maninpulate( … );
   clEnqueueWriteBuffer( … );
   clEnqueueNDRange(…,sum,…);
   clEnqueueReadBuffer( … );
   display_result( … );
}
```



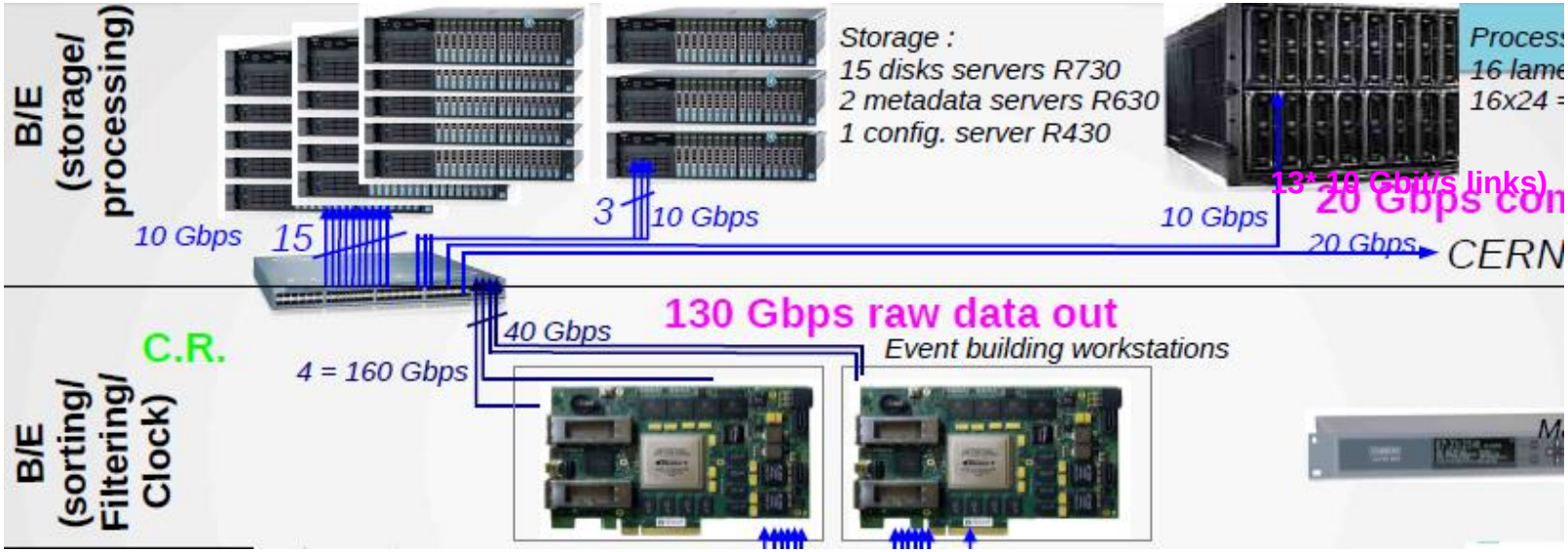
Host CPU separate or embedded

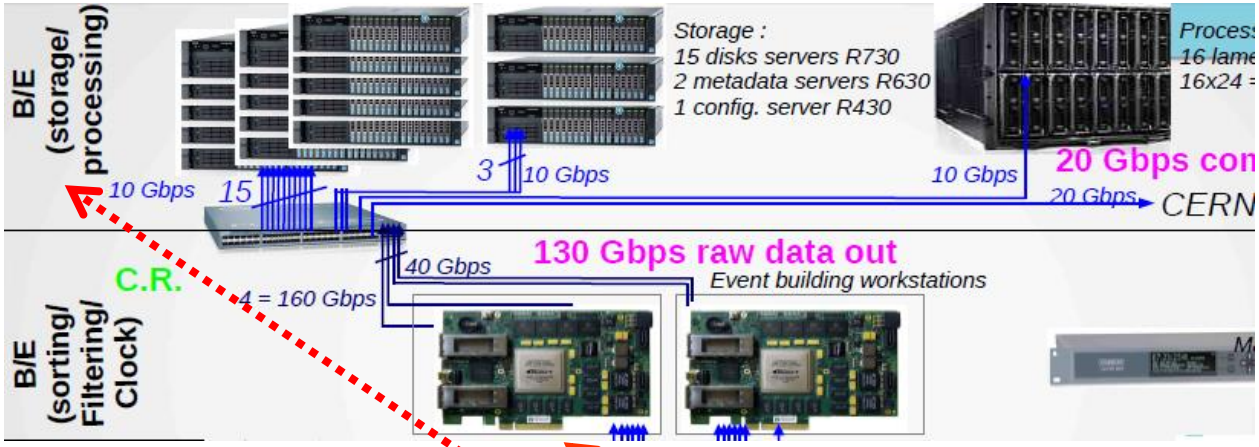
A data flow of 12 Gbps (compressed) has to be treated by the online storage farm

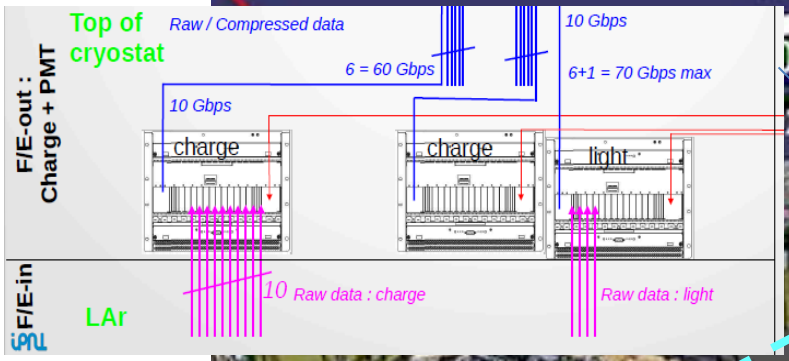
Data storage is distributed on 15 servers R730xd, each one including 16 disks of 6PB

The system also includes 2 MetaData Servers (MDS, DELL R630) +1 configuration Server (DELL 430)

The local storage is based on EOS




B/E (storage/ processing)
Storage :
15 disks servers R730
2 metadata servers R630
1 config. server R430
10 Gbps
15
3
10 Gbps
10 Gbps
20 Gbps
CERN
C.R.
40 Gbps
130 Gbps raw data out
Event building workstations
4 = 160 Gbps
B/E (sorting/ Filtering/ Clock)


Top of cryostat
Raw / Compressed data
10 Gbps
6 = 60 Gbps
6+1 = 70 Gbps max
10 Gbps
charge
charge
light
F/E-out : Charge + PMT
F/E-in
LAr
10
Raw data : charge
Raw data : light

2x 10 Gbps
CERN C.C.
ximity racks
DAQ & computing rooms
NP02 DP
H4 beam line
beam

- First design of online storage/processing DAQ back-end farm performed in 2016 (1PB, 300 cores, 20Gb/s data flow),

DELL-based solution : configuration

storage servers :
* 15 R730XD (storage servers) including :
* 16 disks 6To
* 32Go RAM
* 2 disks system RAID 1, 300 Go 10k
* 1 network card Intel X540 double port 10 GB
* 4 years extended guarantee (D+1 intervention)
* 2 processors Intel Xeon E5-2609 v3
* raid H730P
* Rails with management arm
* double power supply

metadata servers (MDS) :
* 2 R630 (metadata servers), including :
* 2 disks 200 Go SSD SAS Mix Use MLC 12Gb/s
* 2 processors Intel Xeon E5-2630 v3
* 32Go DDR4
* RAID H730p
* network : Intel X540 2 ports 10 Gb
* 4 years extended guarantee (D+1 intervention)
* Rails with management arm
* double power supply

configuration server :
* 1 R430 (configuration server)
* 1 processor E5-2603 v3
* RAID H730
* 2 hard disks 500 Go Nearline SAS 6 Gbps 7,2k
* 16 Go DDR4
* Rails with management arm
* double power supply

Offline computing farm: 16*24 = 384 cores
* 1 blade center PowerEdge M1000e with 16 lames M630, each including :
* 128Go DDR4
* 2 processors Intel Xeon E5-2670 v3
* 4 years extended guarantee (D+1 intervention)
* 2 hard disks 500 Go SATA 7200 Tpm
* netwok Intel X540 10 Gb

Switch Force10, S4820T (see next slide) :
* 48 x 10GbaseT ports
* 4 x 40G QSFP+ ports
* 1 x AC PSU
* 2 fans

- Smaller test scale system already installed and operative for 3x1x1

- Tests to finalise the architecture of final online storage/processing facility.

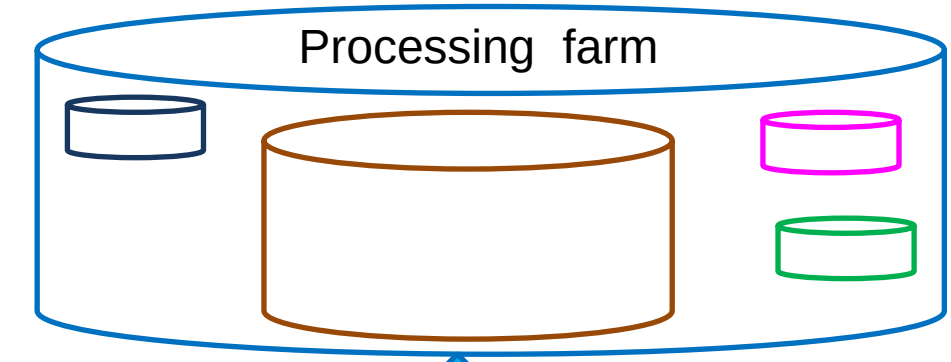
- Thanks to CERN/IT support

Extensive tests during the last year have show the validity of use a local EOS implementation as high performance distributed file system to build up a storage backend at the 20 GB/s level

Online storage and data processing system also fully operative since the beginning of December

- EOS storage file system/metadata server
- Batch system: Torque
- Files transfer to EOS for users analysis access on lxplus
- Scripting and software developed for automatic files handling, storage, dispatching to batch workers and analysis
- Online analysis software for purity/gain determination, storing of results on EOS

Beyond use for 3x1x1 online monitoring (which has a rather modest data flow) this system is a prototype/test bench in order to study the design and perform the development of the final online data storage and processing system foreseen for the 6x6x6 via mock data challenges at high rate with both simulated and real data



Background processes

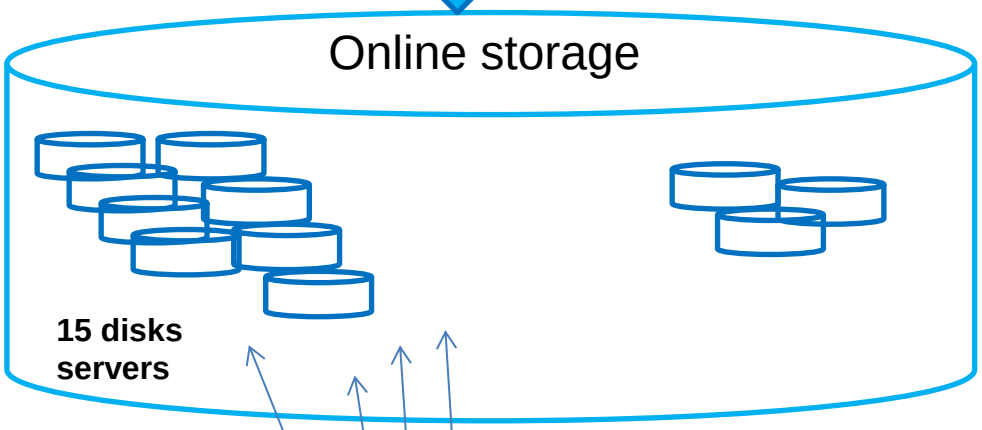
Data processing

Monitoring tasks

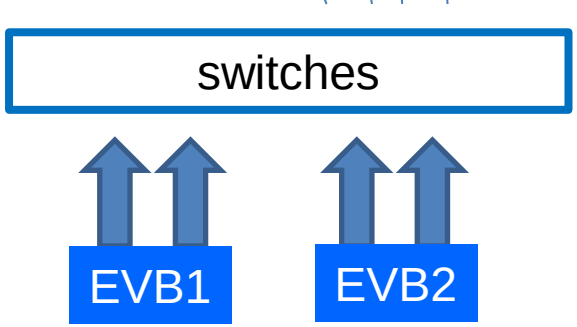
Data transfer

shifter

run event display and monitor online analysis results



CERN IT



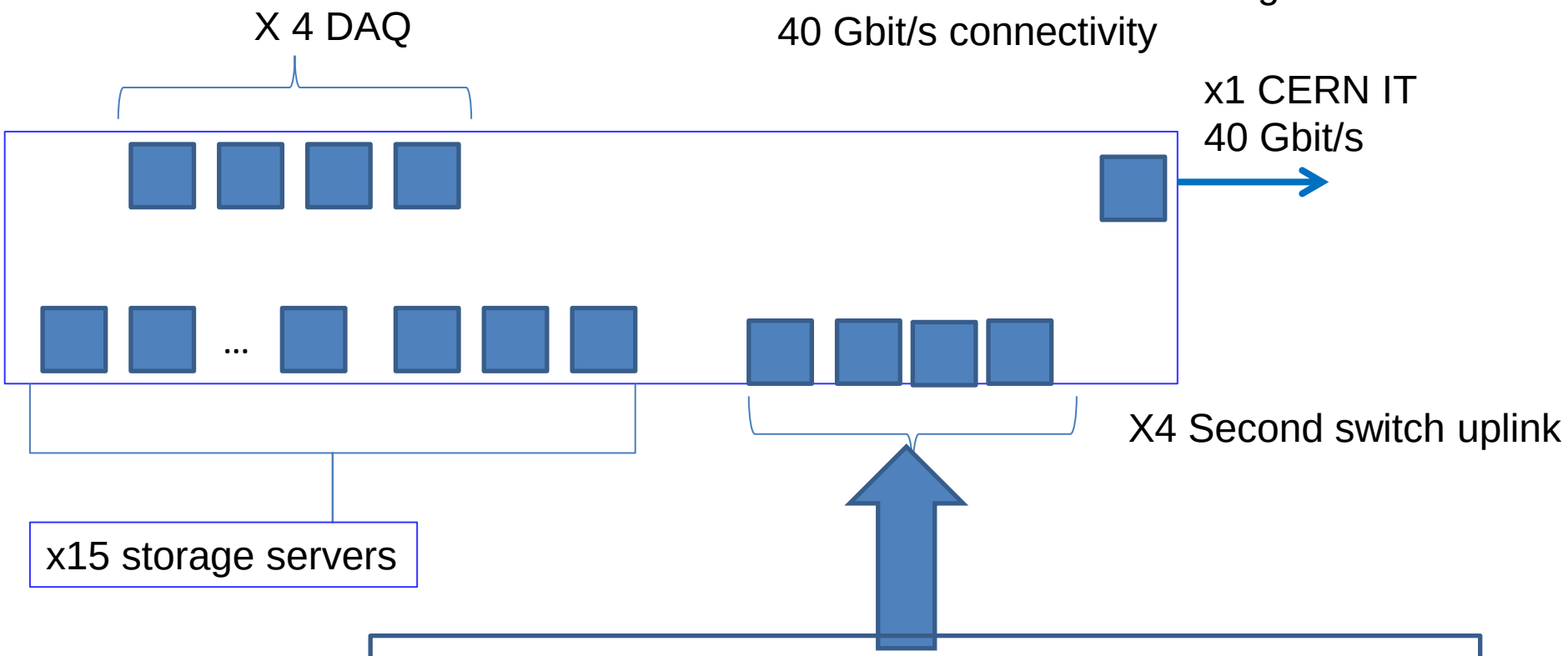
Private network, with a local bandwidth of 20GB/s

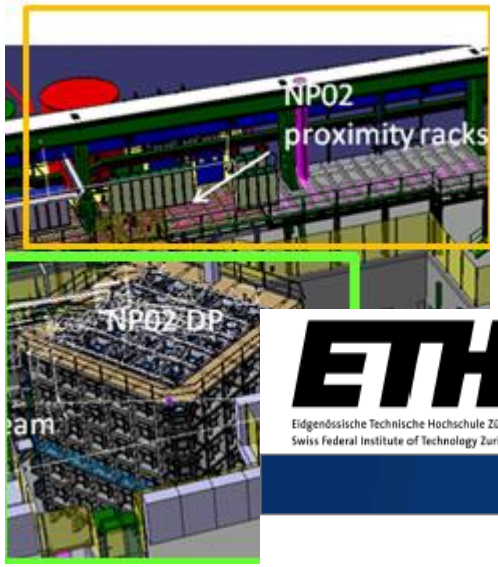
to allow the transfer of compressed data to local storage system and concurrent reading and writing access for online analysis

CERN GENERAL PURPOSE NETWORK for access to machines

x24 40Gbits/s links

Further evolution of the network backbone architecture for 6x6x6 taking into account 40 Gbit/s connectivity



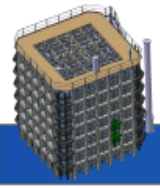


Slow Control

ETH Eidgenössische Technische Hochschule Zürich Swiss Federal Institute of Technology Zurich

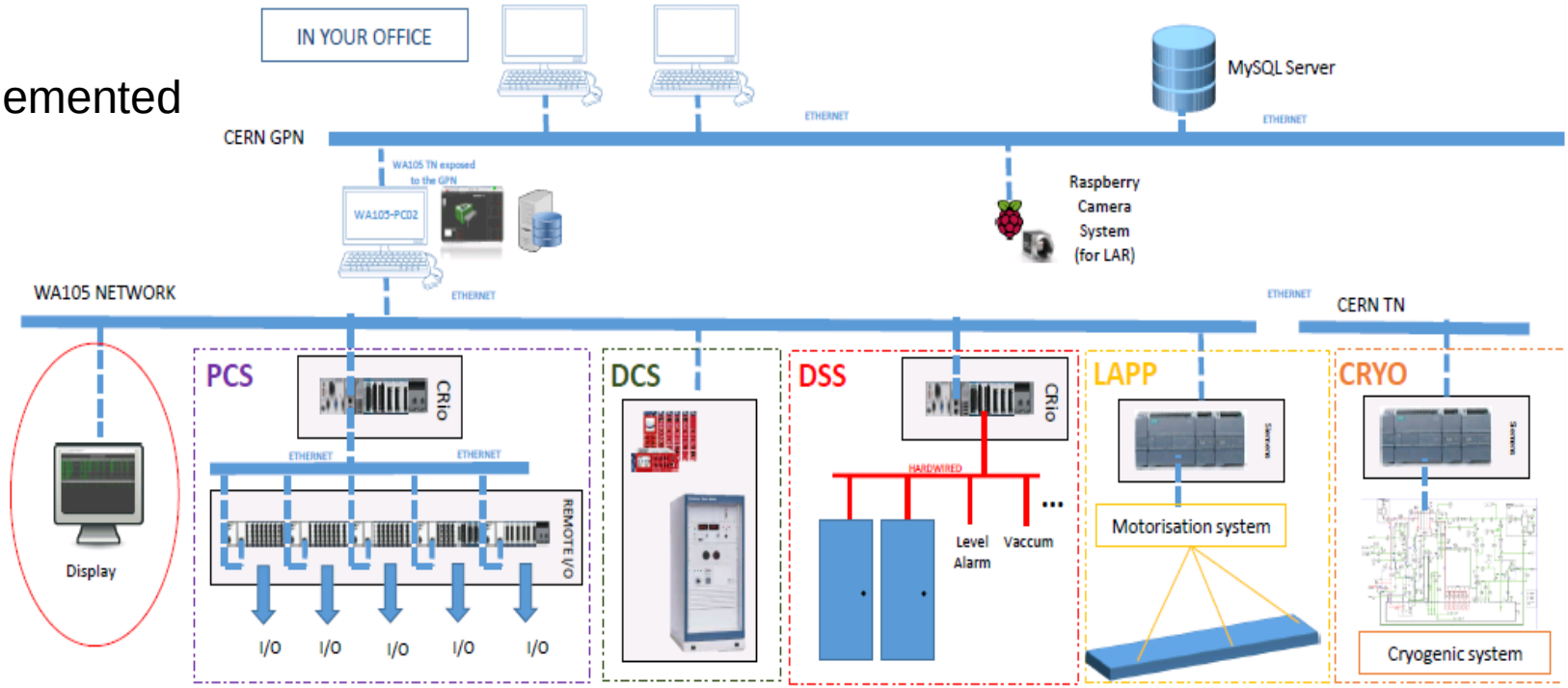
CERN | EP-DT Detector Technologies

WA105



NP-02: ProtoDUNE Double Phase Prototype

Scheme implemented on 3x1x1



Analog FE electronics
Digital FE electronics and DAQ

(Production, QA, installation)

Cryogenic FE electronics :

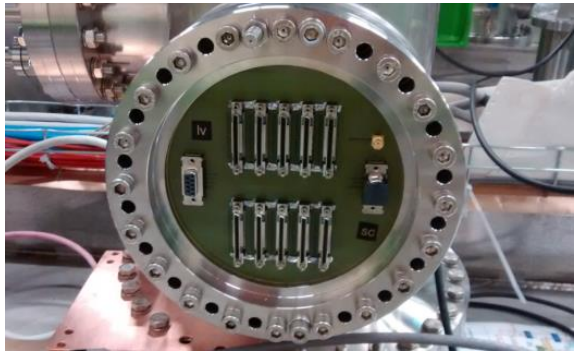
Dual-slope ASICs final version. Full production for 6x6x6 produced and purchased (700 chips) in March 2016

FE-cards designed in 2016 together with chimneys warm flanges PCBs

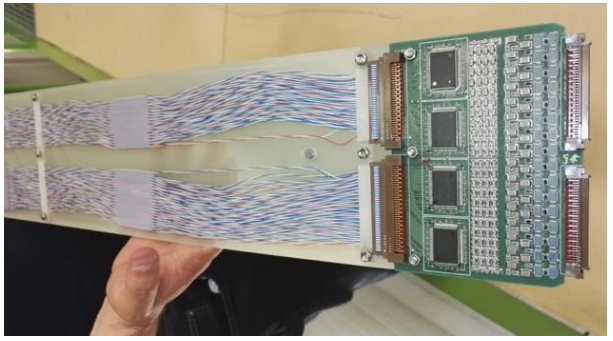
20 FE cards (1280 channels) produced and installed on 3x1x1 pilot detector at CERN

All cards pretested and calibrated in laboratory for all channels before mounting on detector **(no dead channels)**

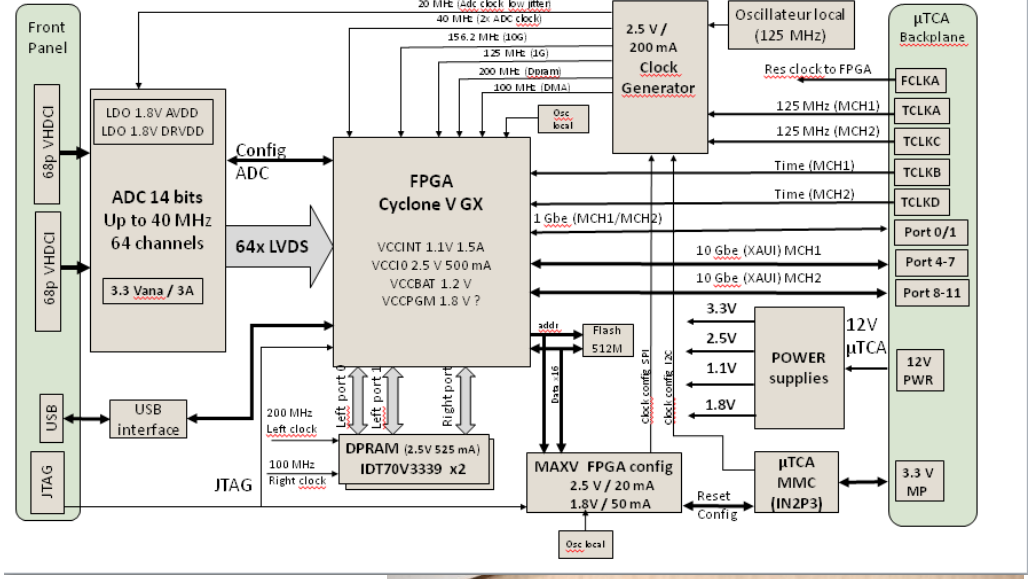
→ Remaining production of 100 cards for 6x6x6 being launched on 2017 budget, with similar QA procedure as for 3x1x1 prebatch
→ Installation in October 2017



Warm flange PCB,
Extension of 3x1x1 design
(x2 more channels)
Routing in progress,
production by July



uTCA DAQ system:



64 channels AMC digitization cards (2.5-25 MHz, 12 bits output, 10 GbE connectivity

- Demonstrator card with 64 ADC channels built and tested in 2015 for the definition of the final card

- Purchase of main components (ADCs, FPGAs, IDT memories) of the final cards by end of 2015 to equip the entire 6x6x6

- 20 cards produced by September 2016 to equip the 3x1x1

- Complete QA procedure with laboratory tests for all channels (groups of 10 cards), no malfunctioning channels

- Cards production going to be completed with the 2017 budget of remaining 100 FE and uTCA cards for 6x6x6 (main components available), similar QA and for pre-batch of 3x1x1. Installation in November 2017

Online Processing:

- Extensive tests during the last year have show the validity of use a local EOS implementation as high performance distributed file system to build up a storage backend at the 20 GB/s level

- This is already implemented at the level of the 3x1x1 which provided also very good experience for all the software and system management developments)

- Procurement of computing material (disk servers, CPU and network architecture) for 6x6x6 online computing and storage facility already started in collaboration with the Neutrino Platform people

- Very strong collaboration and support from the IT division

- Very good infrastructure with counting rooms racks and cooling power (about 350 kW), final 40 Gbit/s connectivity to IT for each one of the protoDUNE detectors

- Other more critical components like the DAQ backend machines, metadata server etc, being finalized). All the backend hardware should be finalized for procurement in the next two months.

→ Looking forward to the start of the FE-DAQ installation in November 2017 !