protoDUNE DP online network scheme

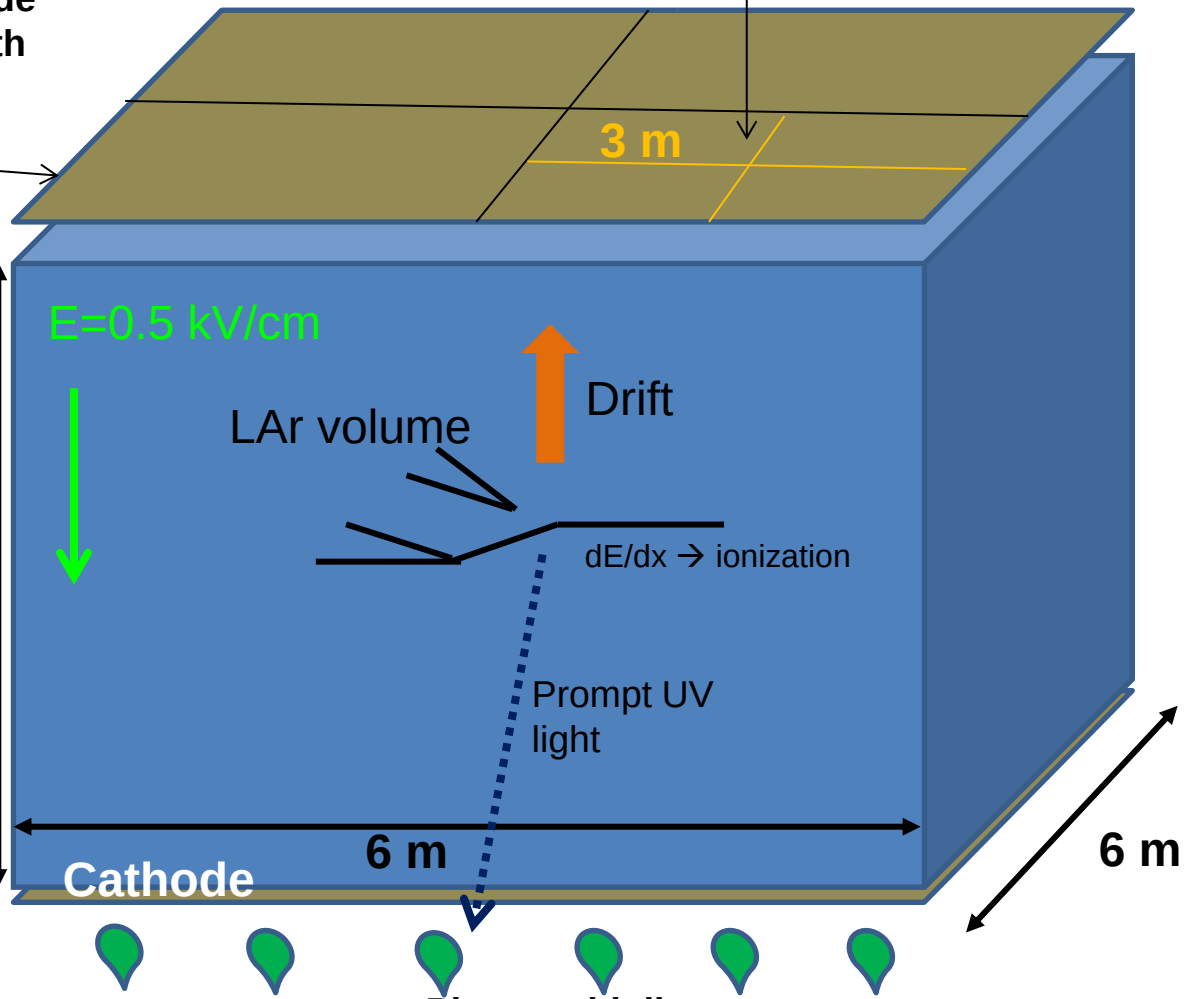
Elisabetta Pennacchio, IT-protoDUNE coordination meeting, 08/03/2017

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

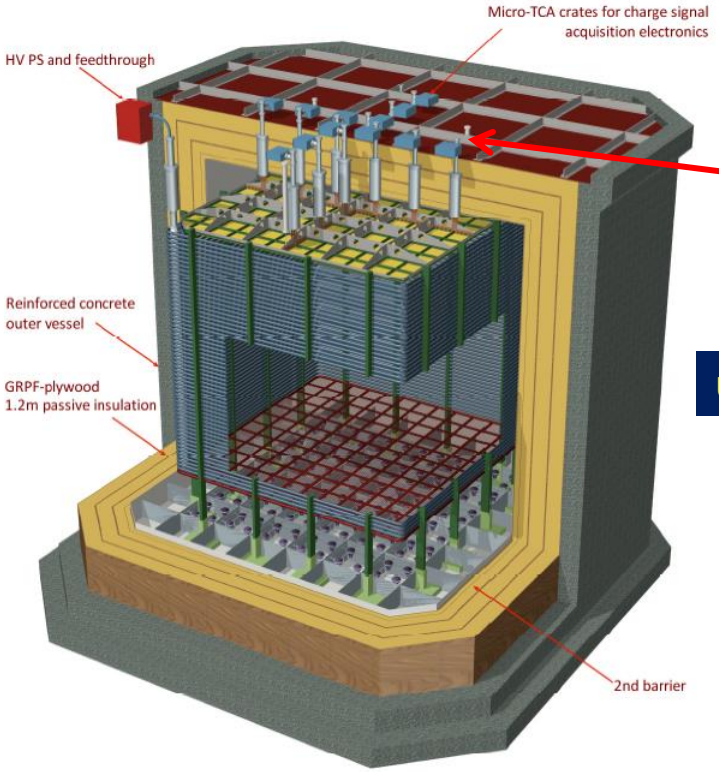
Segmented anode
In gas phase with
double-phase
amplification



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

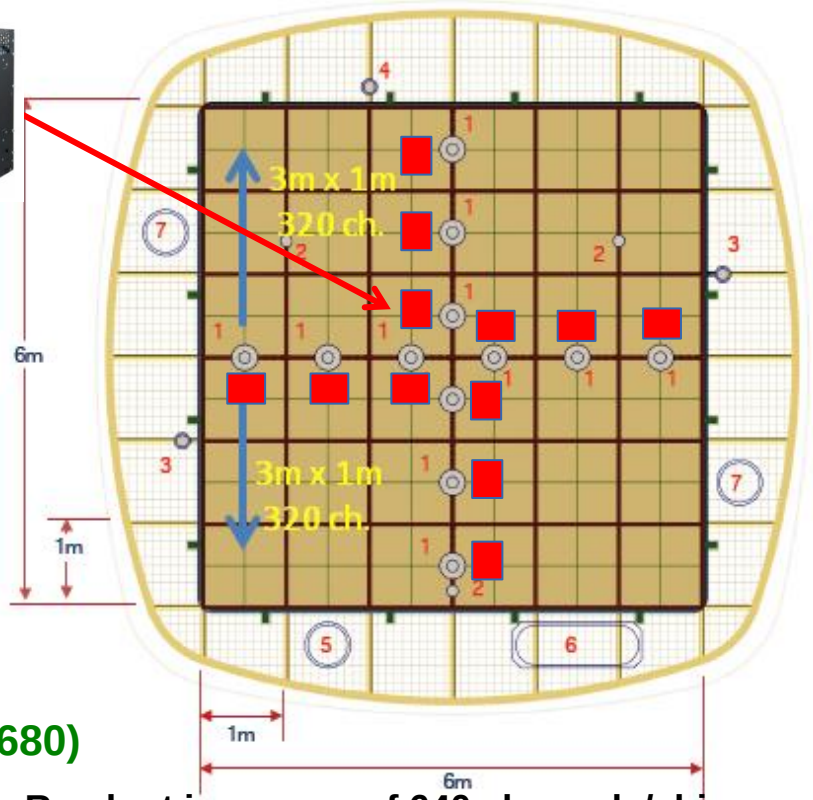
→ 10000 samples
per drift window

uTCA charge readout architecture



uTCA crates

View from anode with signal (1), suspension (2), HV(3), PMT(4), manhole (5), detail insertion (6), clean room IN/OUT (7) nozzles



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

3

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:

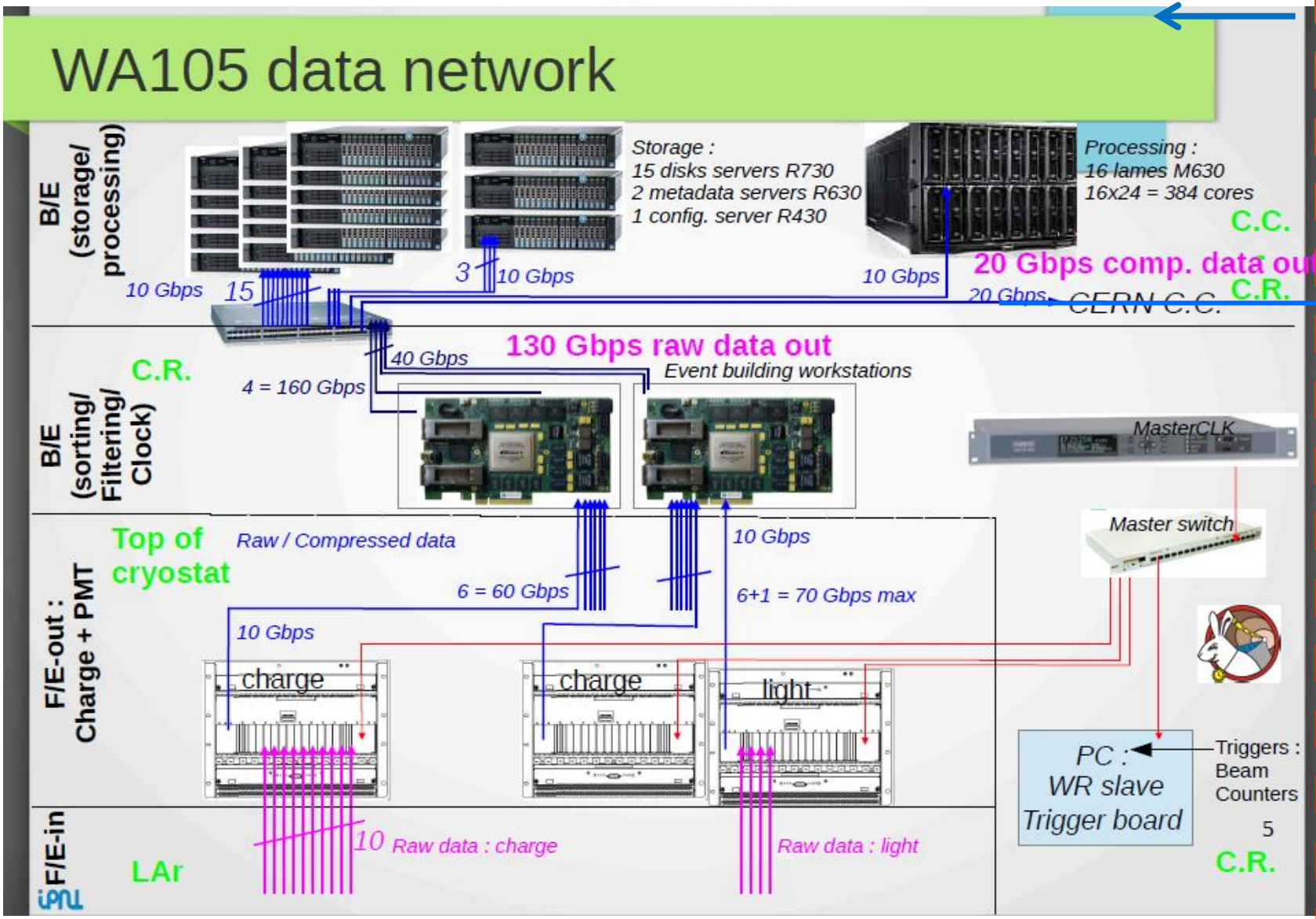
- Completion of event building by connecting the data flows of the two back-end systems
- Fast event reconstruction and disentangling of cosmic rays tracks segments by using also the LRO timing information
- Selection of a subsample of the cosmic ray tracks overlapped to beam events for online purity 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)

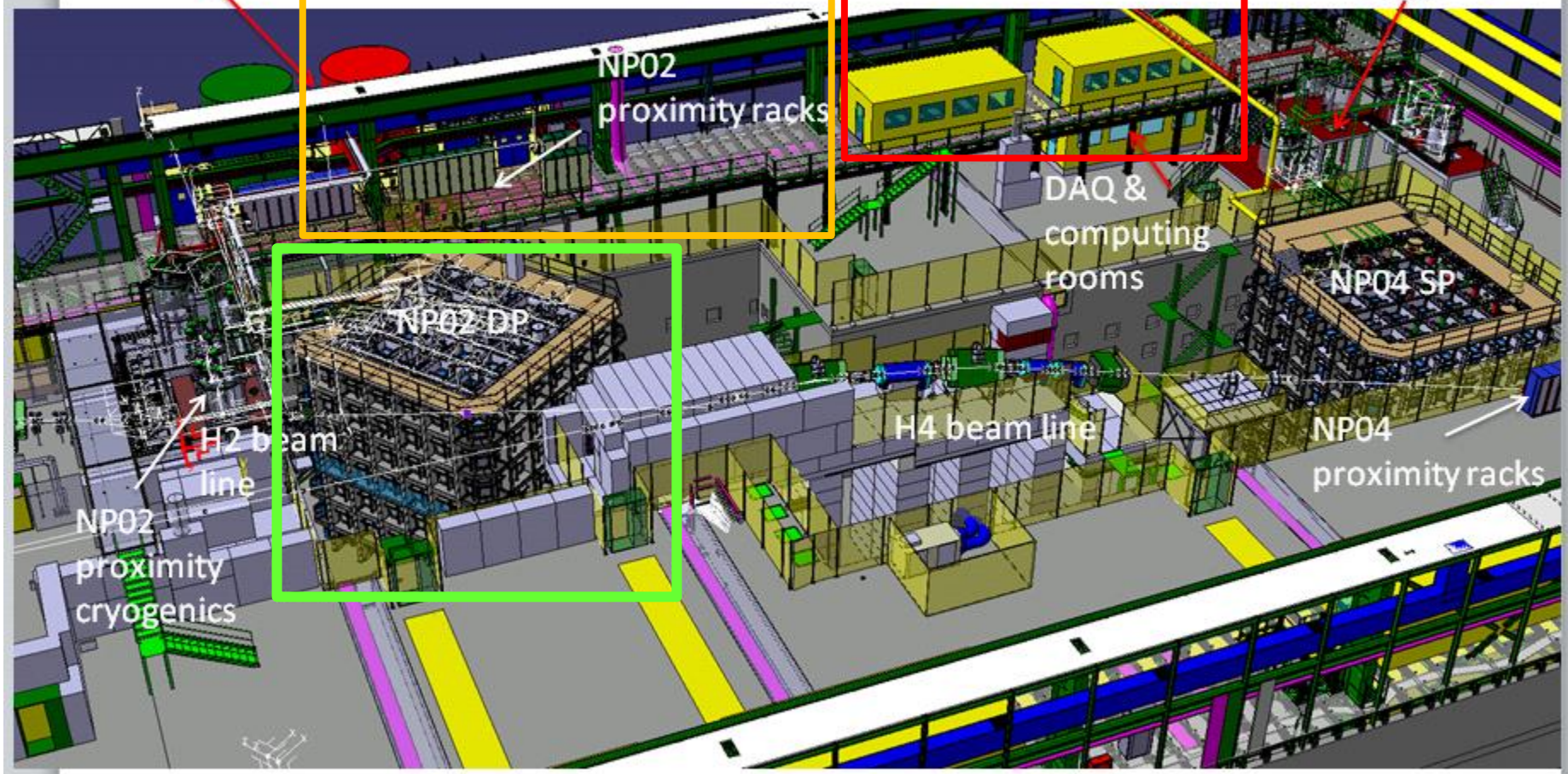
online

offline

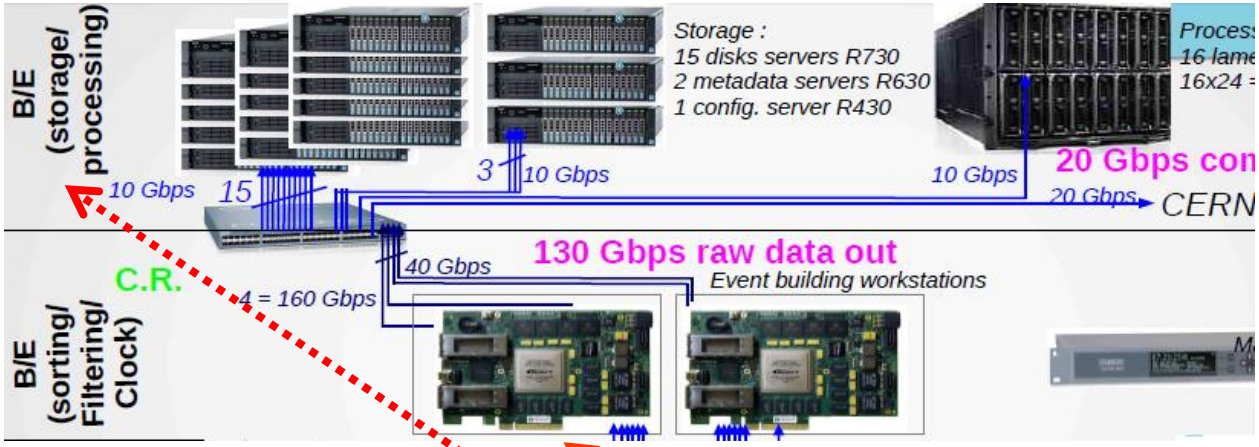
WA105 data network

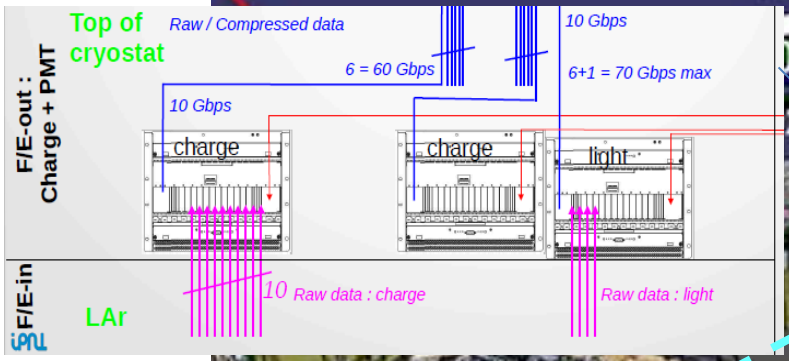


EHN1-1 Layout



CATIA, integration model


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
4 = 160 Gbps
Event building workstations
B/E (sorting/ Filtering/ Clock)


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

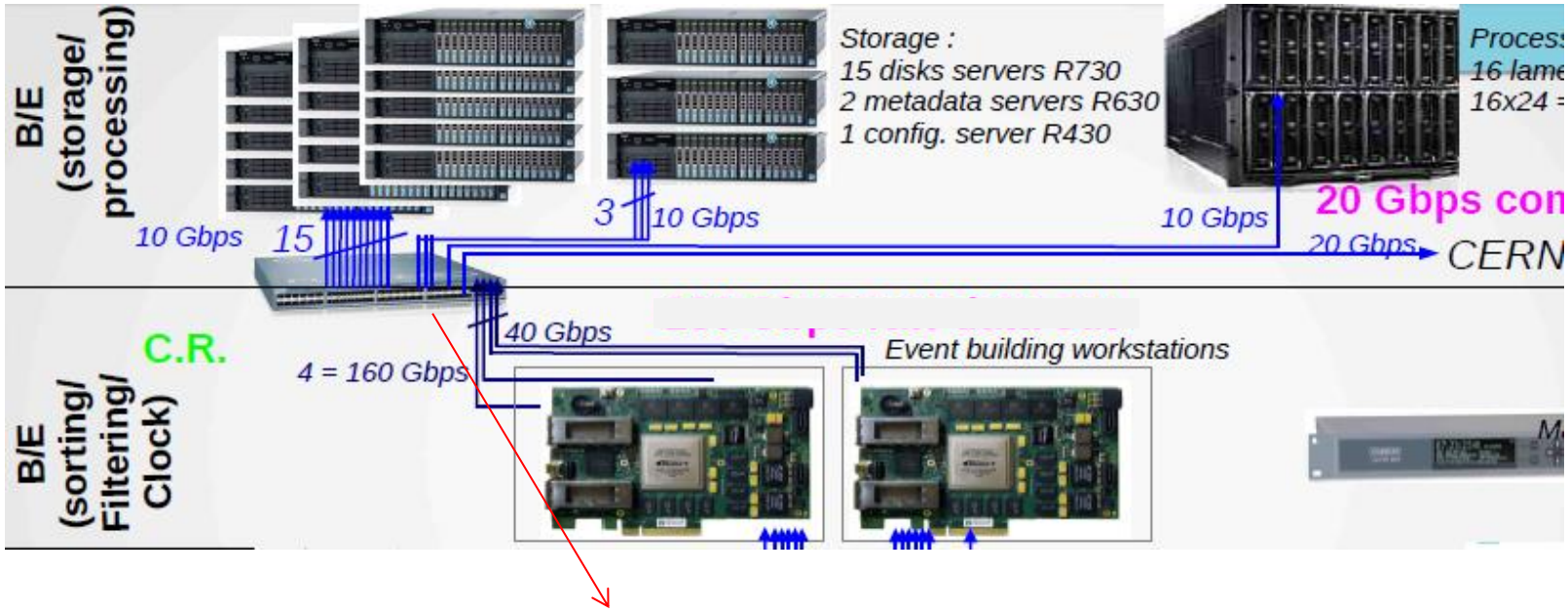
2x 10 Gbps
CERN C.C.
ximity racks
DAQ & computing rooms
NP02 DP
beam
H4 beam line

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

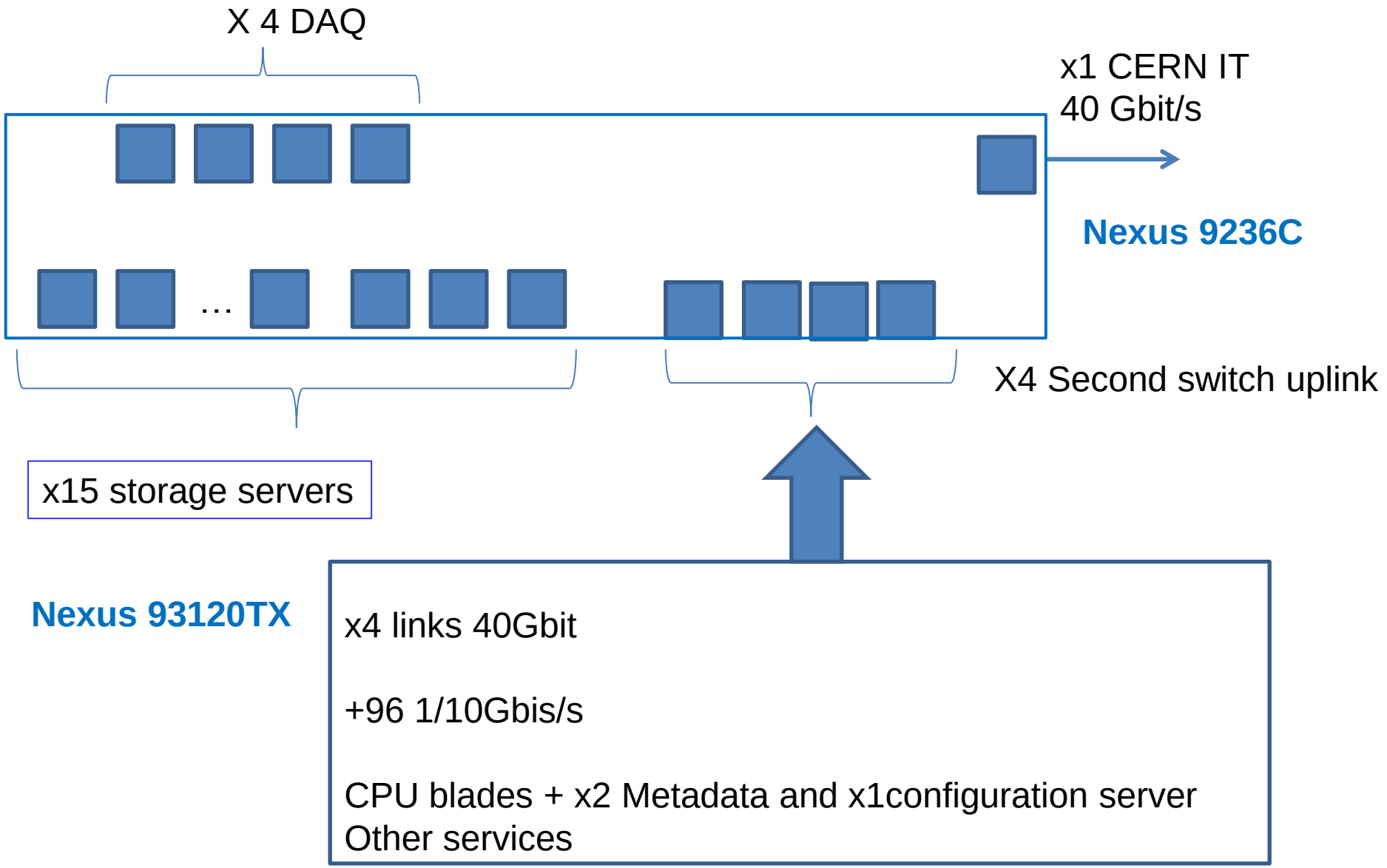
The system also includes 2 MetaData Servers (MDS, DELL R630) +1 configuration Server (DELL 430)

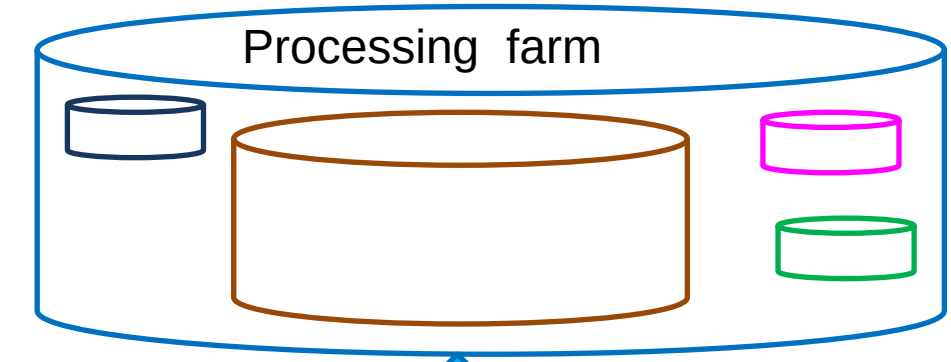
The local storage is based on EOS (already implemented on the 3x1x1 test farm)



see next page

x24 40Gbits/s links (40 Gbits cards for local connectivity of servers became available at low cost)





Background processes

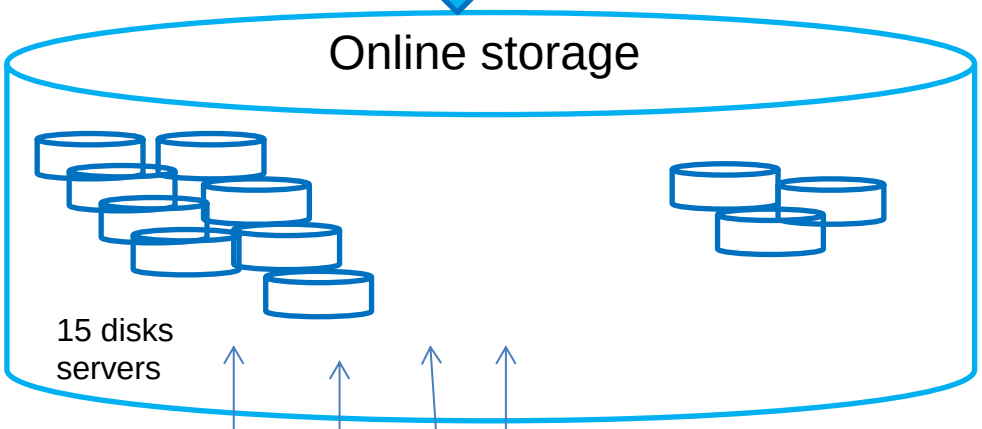
Data processing

Monitoring tasks

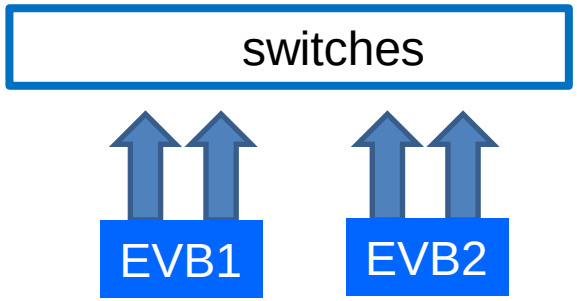
Data transfer

shifter

run event display and monitor online analysis results



CERN IT



private network(s), 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 ?