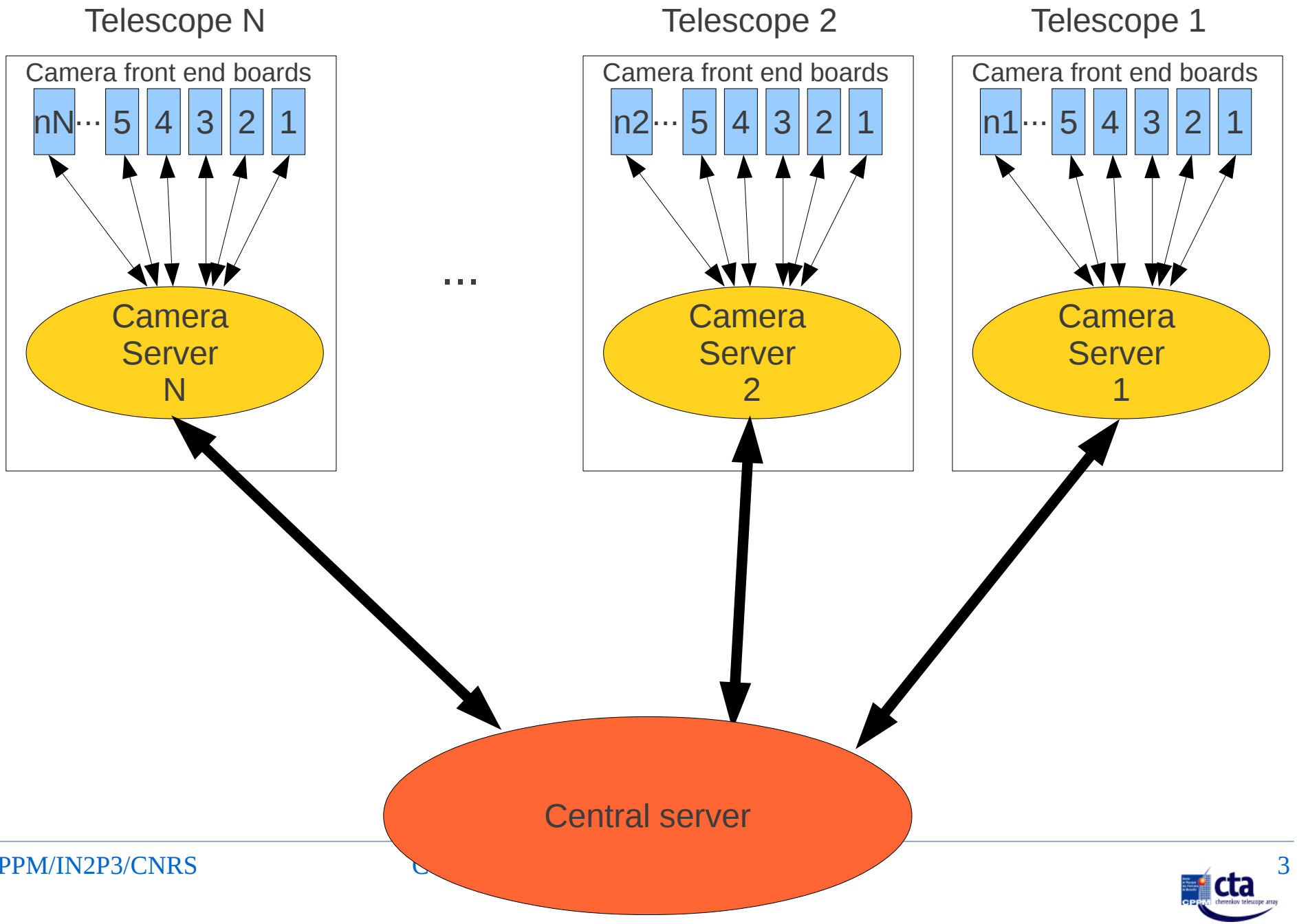
A data acquisition system for the Cerenkov Telescope Array

Julien HOULES, Dirk HOFFMANN
CPPM/IN2P3/CNRS
And the CPPM CTA group

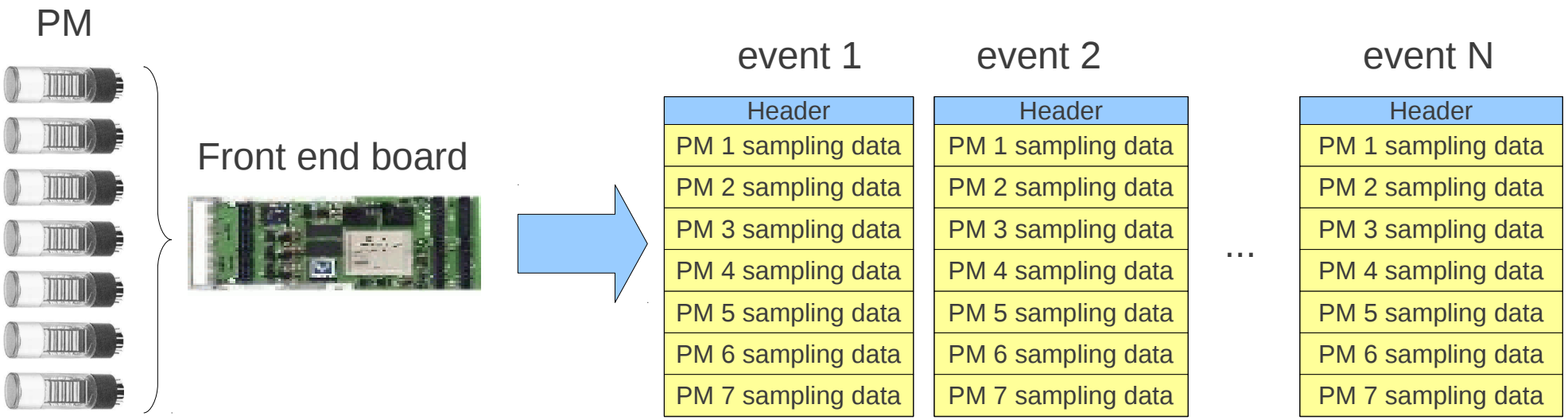
Contact :
houles@cppm.in2p3.fr

Camera server

Global architecture



Camera data flow



Whole Camera ~ 2000 PM -> 300 front end boards

Camera server

- Build event
- L2 trigger on camera server (L1 on front end) :
  - CPU (SSE, AVX...)
  - GPU
- Compress ?
- Send data to central server (array level)

Data flow hypothesis

- ~ 2000 pixels camera

- L1 trigger rate : 10 KHz
- Size of sampling data for 1 PM : 144 bytes (16 bit * 72 samples)
- No data loss (all the L1 events are sent)

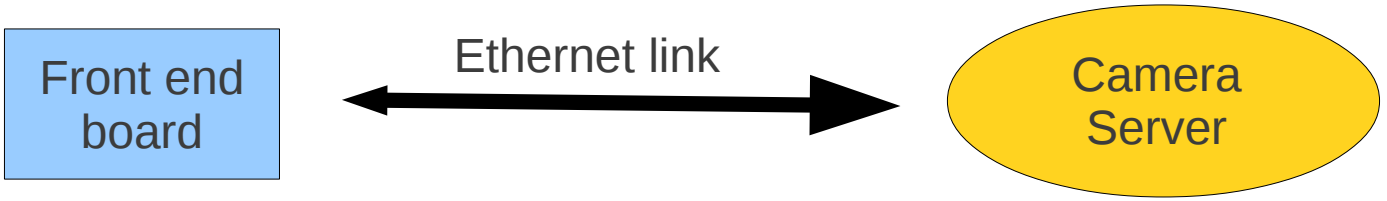
→ Max theorical bandwidth = 10000 * 2000 * 144 = 2.88 GB/s
= 23 Gb/s

- 7 detectors for each front end board : 300 boards/camera

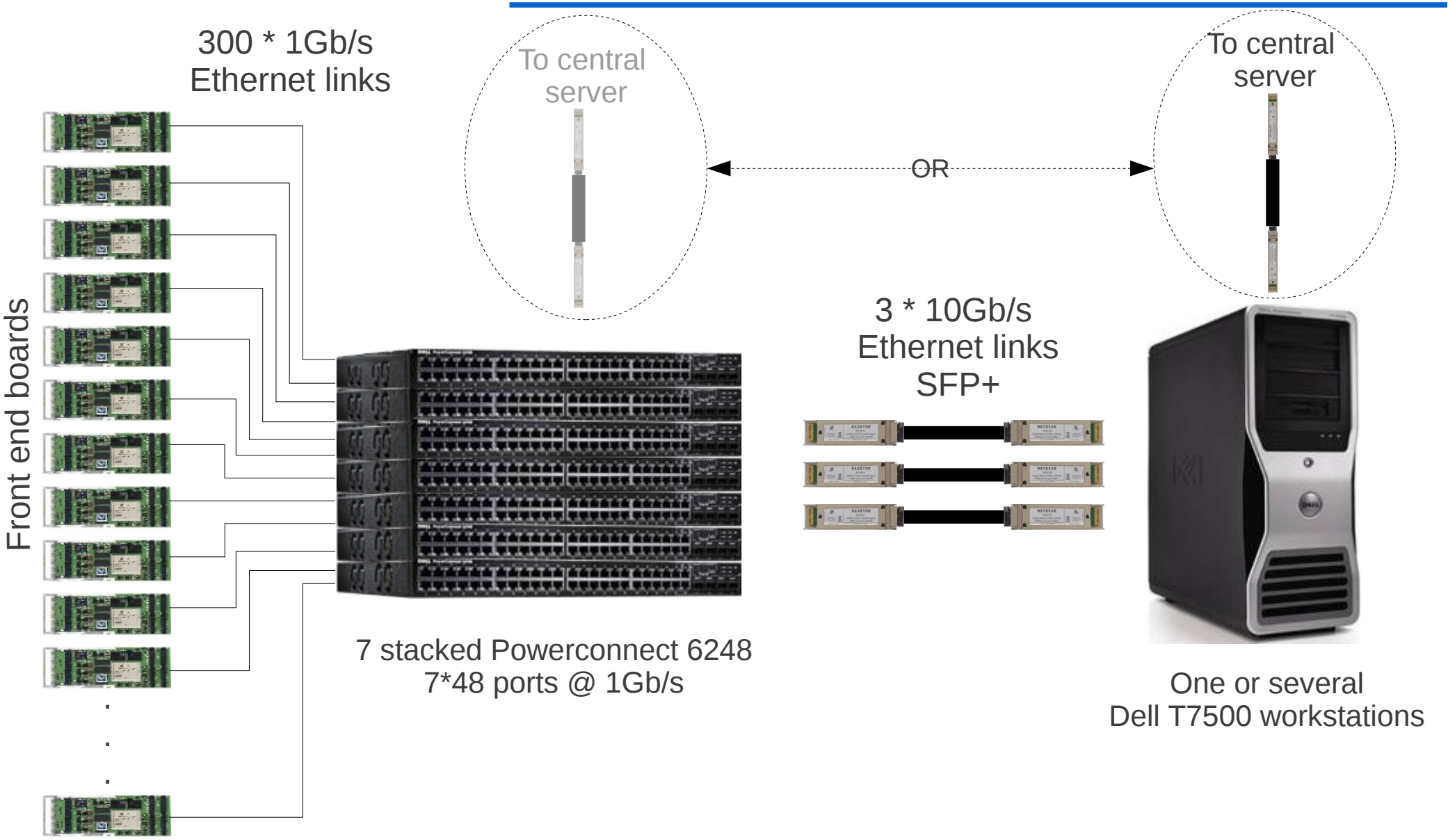
→ Each board generates a flow of 2880/300 = 9.6 MB/s
= 77 Mb/s

https://portal.cta-observatory.org/WG/ACTL/SitePages/Data%20Rates.aspx

Global architecture



Camera infrastructure



Dell Precision T7500

- Two Intel **Xeon X5650** (2.66GHz,6.4GT/s,12MB,6Cores)
- Memory : 24GB (6x4GB) 1333MHz

- **Intel X520** DA2 10GbE Dual Port SFP+ Server Adapter, PCIe x8

- Triple channel (maximum speed reached)
- QPI at 6.4 GT/s (maximum speed on the market)
- Memory DDR3-1333
- 2 full speed full duplex 10 Gb/s links (PCIe x8 Gen 2)
- 1 PCIe x16 slot free (->GPU)
  and 1 PCIe x8 free (-> one more 10 Gbps adapter)
- SFP+ -> Copper or Optical link

~ 3500 euros

Dell Powerconnect 6248

- 48 * 1 Gb/s ports
- Backplane 184 Gb/s
- 2 * 10 Gb/s SFP+ ports included
  2 more 10 Gb/s optional ports
- Up to 12 switches stackable
  -> 576 ports

~ 1500 euros (with 2 * 10 Gb/s)

Event builder

Why a prototype ?

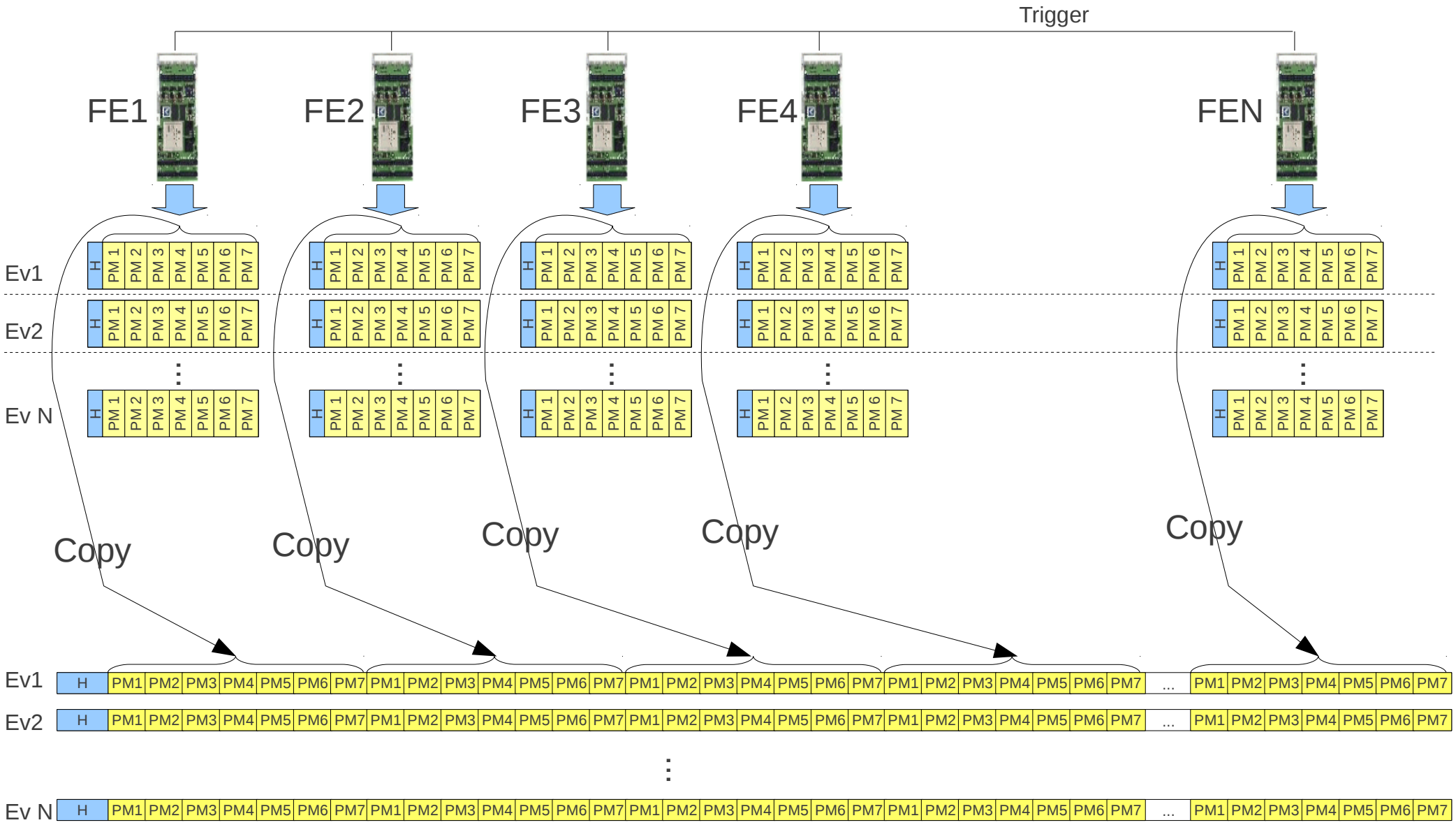
We need a prototype :

- To evaluate the maximum speed reachable
- To test several technologies
- To validate different approaches of the data processing
- To adapt our needs to what we can do

Our first approach

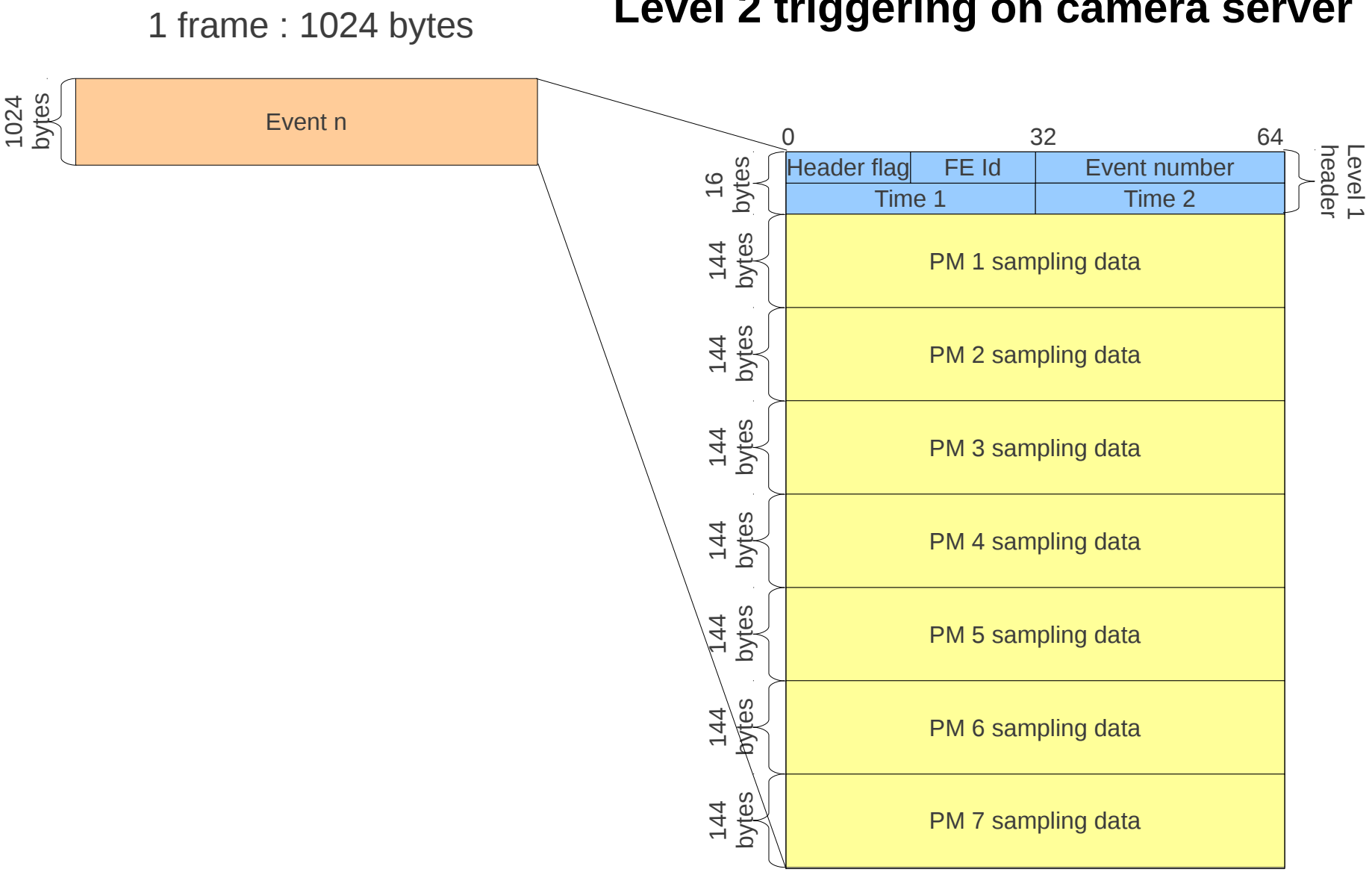
- High modularity to make adaptation to different front end electronics easier
- Multitask approach to divide the flow processing if needed
- Use of a standard Linux distribution but take control on scheduling and memory allocation
- Constrained electronics to reach the best performances (in a first time)

Event builder



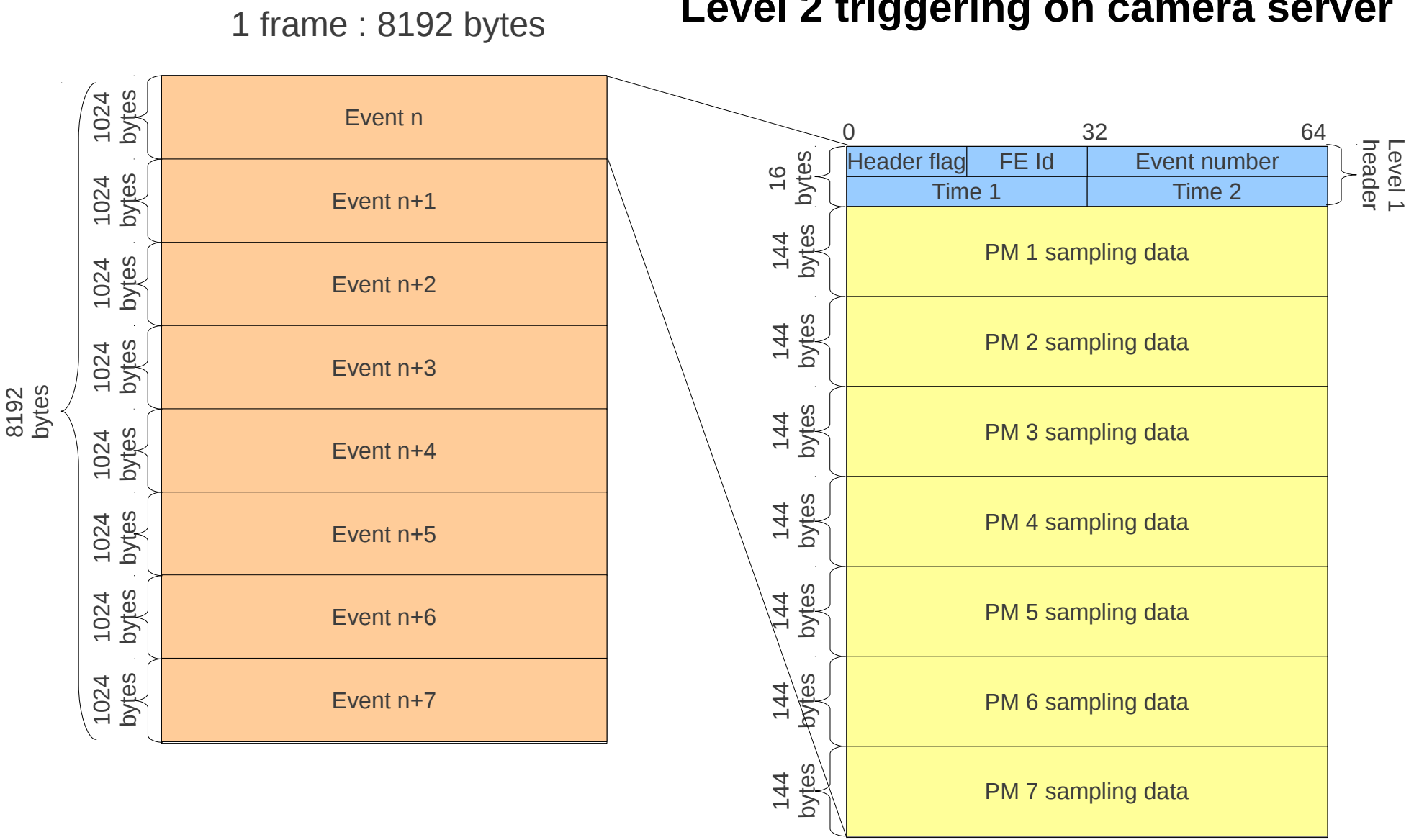
Data format : regular frame

Level 2 triggering on camera server

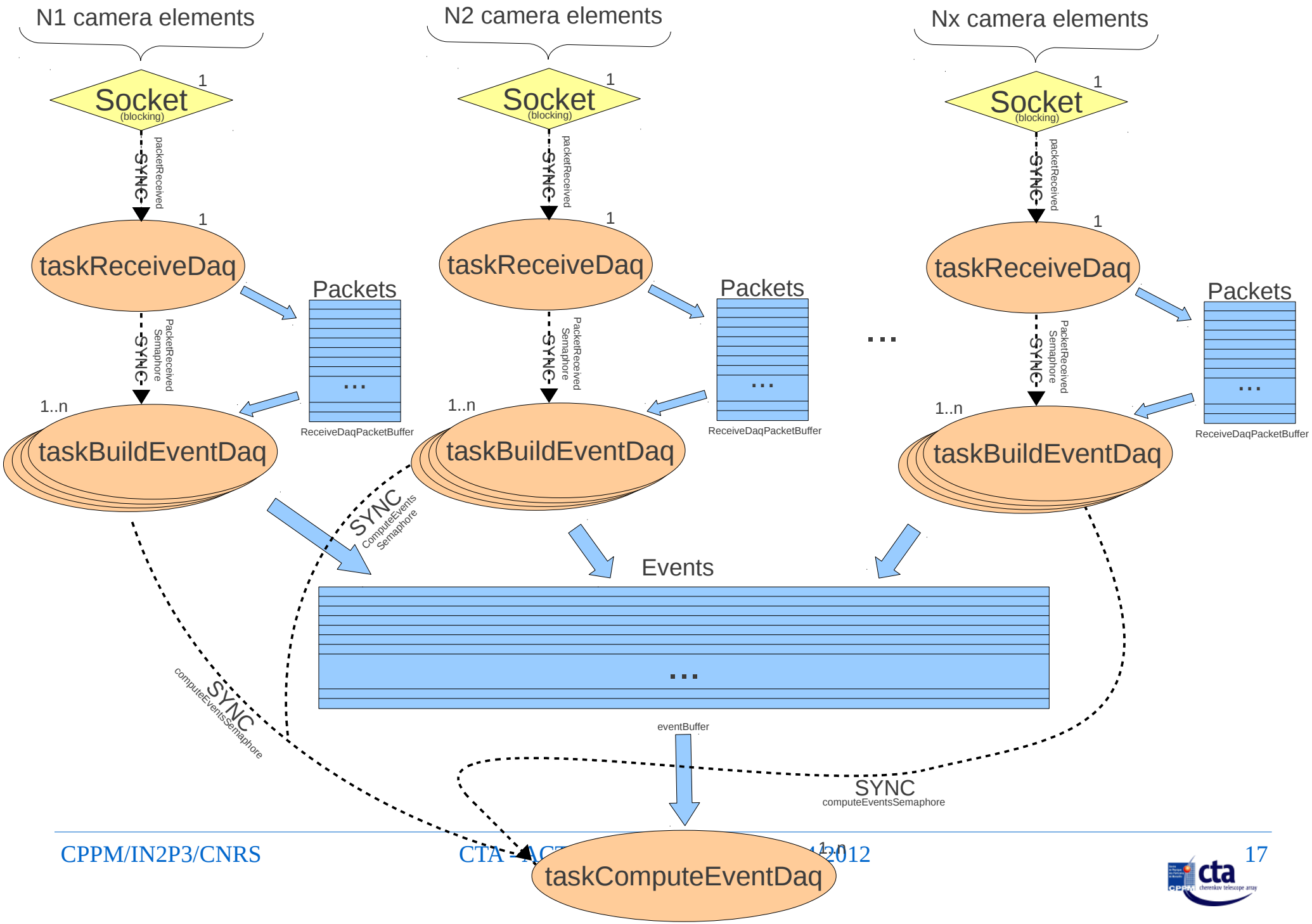


Data format : jumbo frame

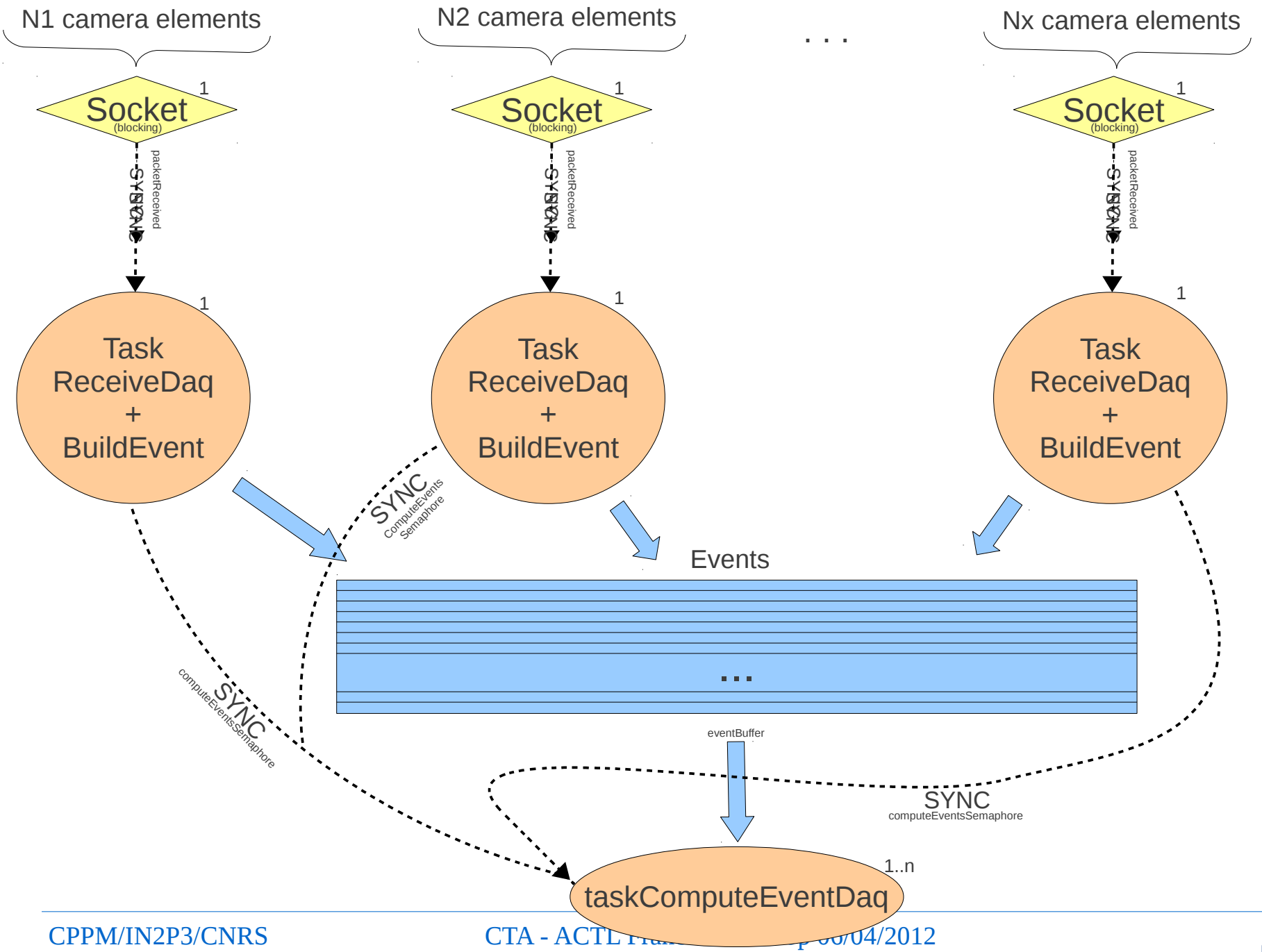
Level 2 triggering on camera server



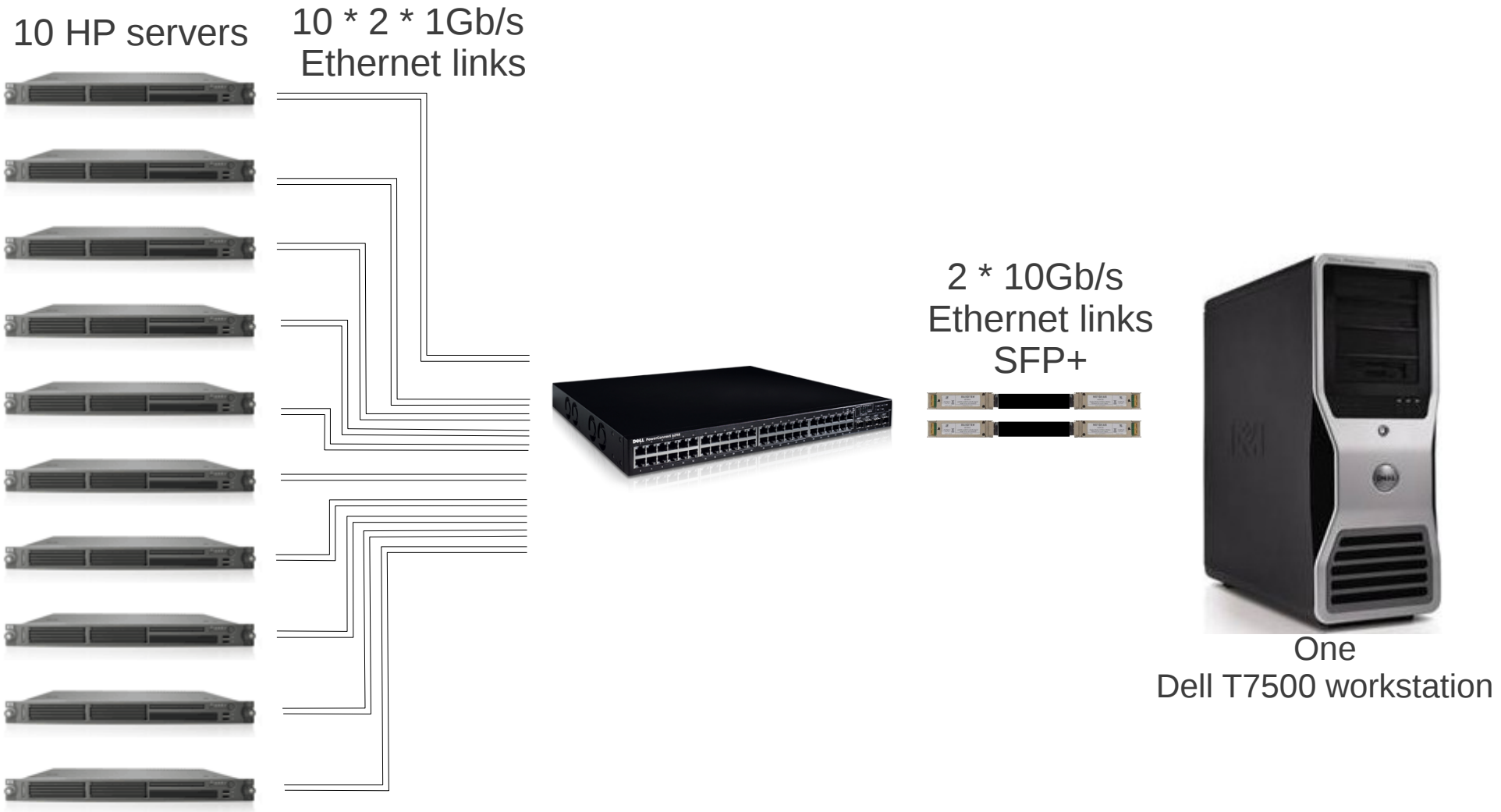
Software overview : 1st architecture



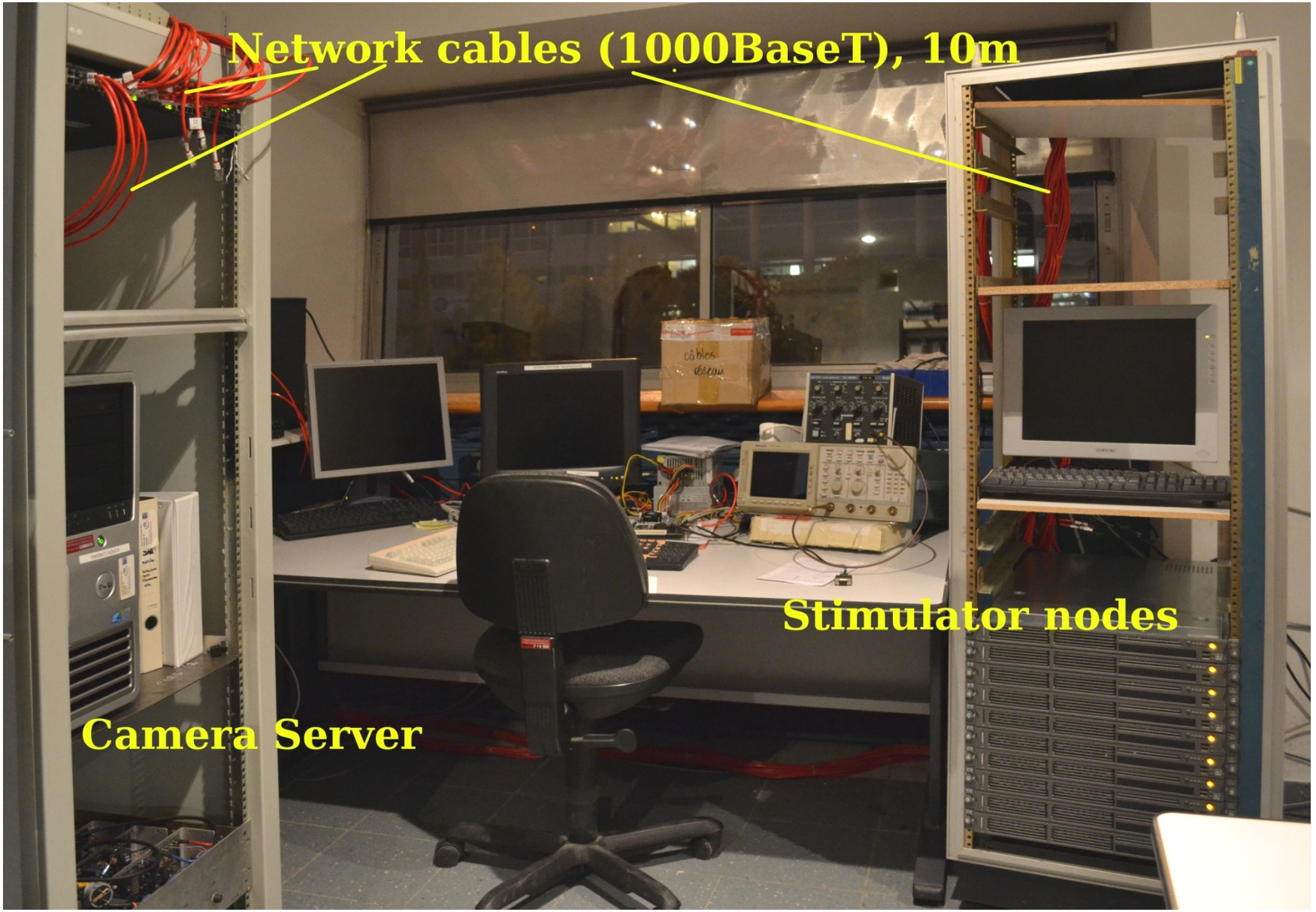
Software overview : 2nd architecture



Stimulation configuration



Stimulation room



First results : event builder

150 nodes (15 per HP server) sending data to interface 1
150 nodes (15 per HP server) sending data to interface 2
} 300 nodes simulated

Tests of the event building with varying packet size:

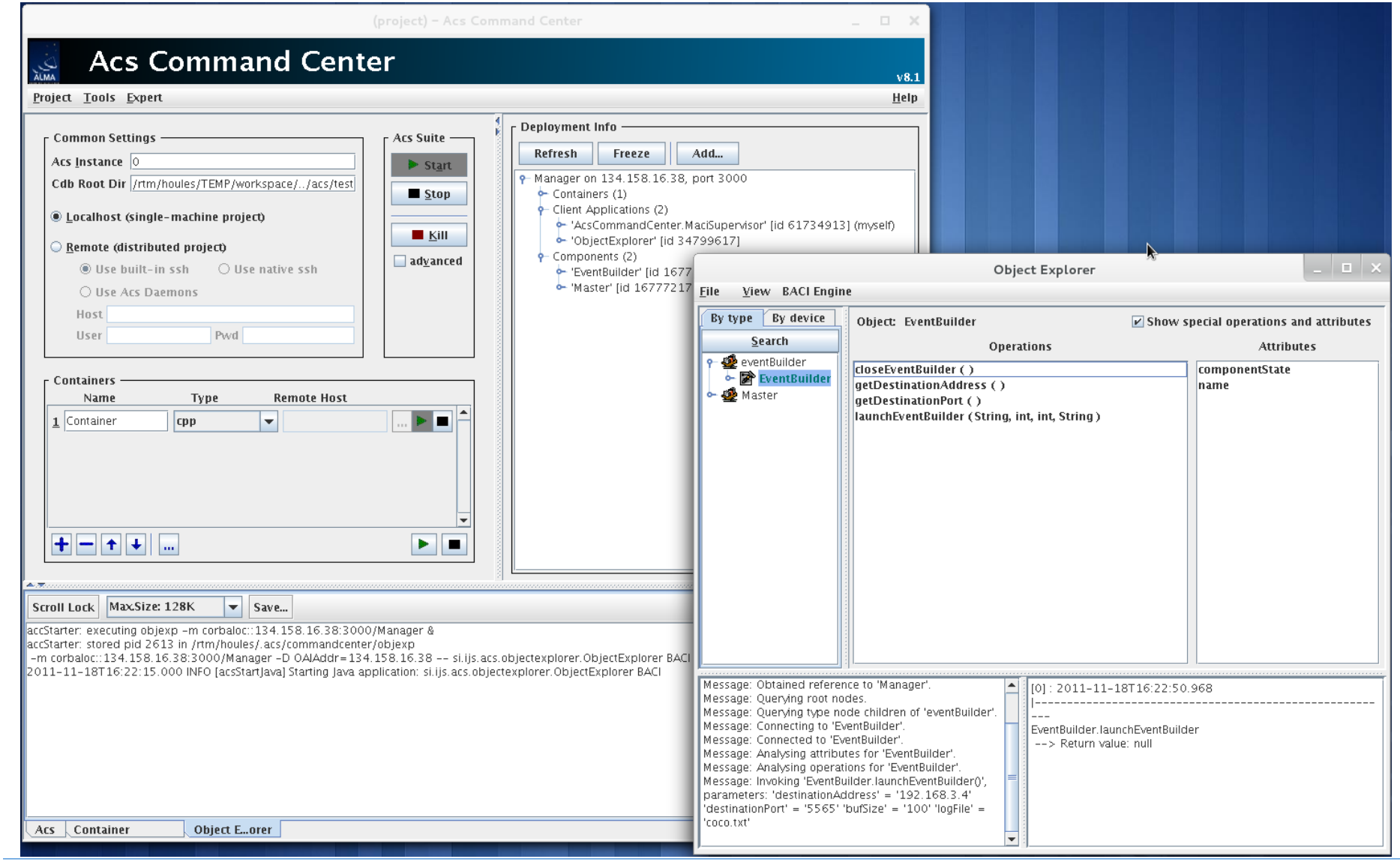
1st Achitecture	2nd Achitecture
Jumbo frames (8192 bytes) : **19,2 Gb/s** (2,4 GB/s ~ 8000 evts/s) with no loss CPU usage : 300 % (3 cores/12)	Jumbo frames (8192 bytes) : **19,2 Gb/s** (2,4 GB/s ~ 8000 evts/s) with no loss CPU usage : 160 % (1.6 cores/12)
Regular frames (1024 bytes) : 6,5 Gb/s (0,82 GB/s) with no loss CPU usage : 300 % (3 cores/12)	Regular frames (1024 bytes) : 8 Gb/s (1 GB/s) with no loss CPU usage : 170 % (1.7 cores/12)

→ **Results obtained with standard libraries/drivers (Linux).**

→ Test of a direct I/O solution to improve small frames reception in progress

Integration in ACS

The basic functions of the Event Builder are available from the ACS interface



Stimulator

Need for a real stimulator

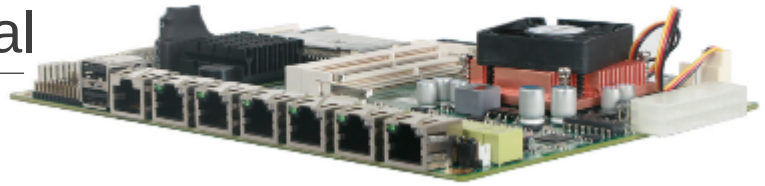
Need a stimulator to make :

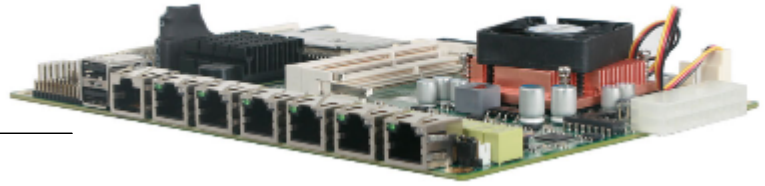
- timing measurements on software
- real time validation
- algorithms validation
- trigger validation
- latency measurements on network
- front end boards and stimulator mix
- validate the complete acquisition chain

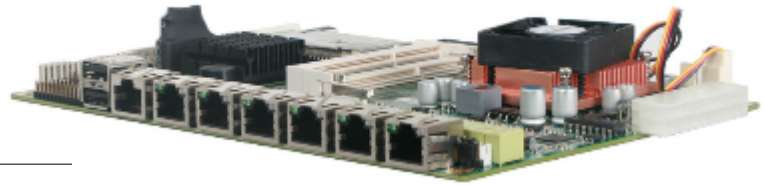
Testing configuration

Sync Signal

EVOC NET-1820

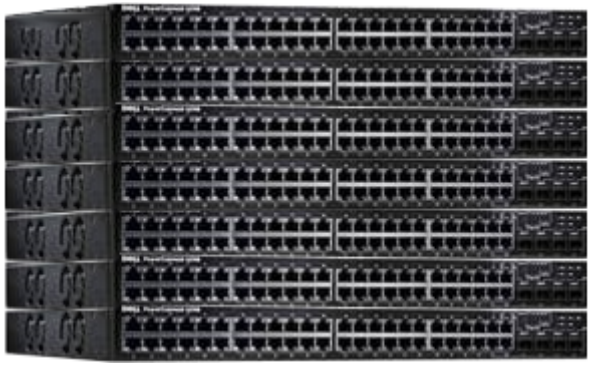


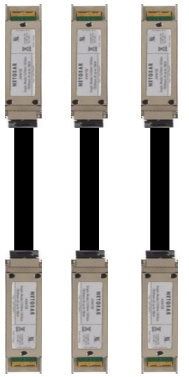




1Gb/s links

Powerconnect 6248 stack



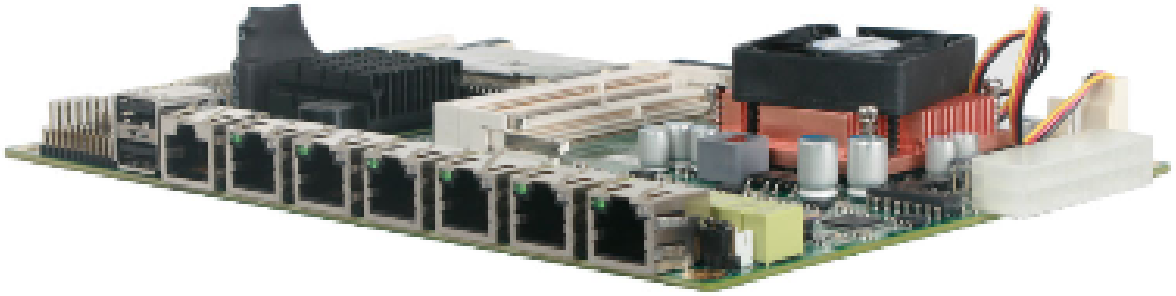


10 Gb/s links

To camera server

EVOC NET-1820

Most promising candidate *(50 € / port)*



Intel Atom D525 dual core processor 1.8GHz
4.0 GB RAM
6 x Intel 82574L Giga LAN
(supports 9K frames and boot on LAN)
8-bit Digital I/O interface
1 x Parallel port, Serial port

~ 300 € each → ~ 15000 € for 300 ports without switches

Measured throughput @ CPPM: ~ 2,4 Gb/s (400 Mb/s per port) → can easily be improved.

Future

Future work

- Test of a direct I/O solution to improve regular frames reception (in progress)
- Perform precise measurements on performances
- Improve the event builder

- Build a full-size stimulator

- Design a L2 trig (CPU ? GPU ?)

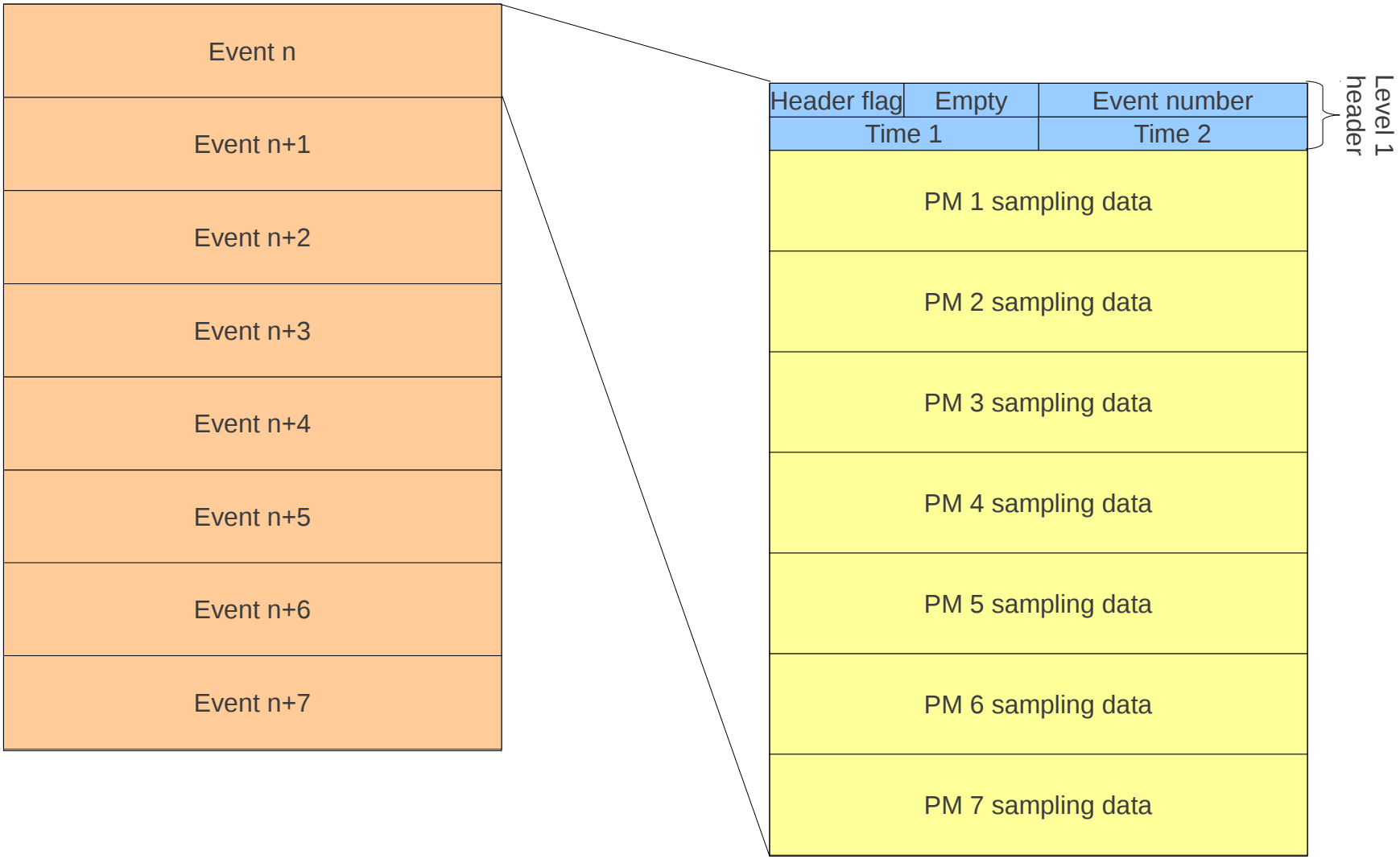
- Work with slow control and array server communications teams

- Full ACS integration
- Make the software reliable enough for production stage

Interface definition

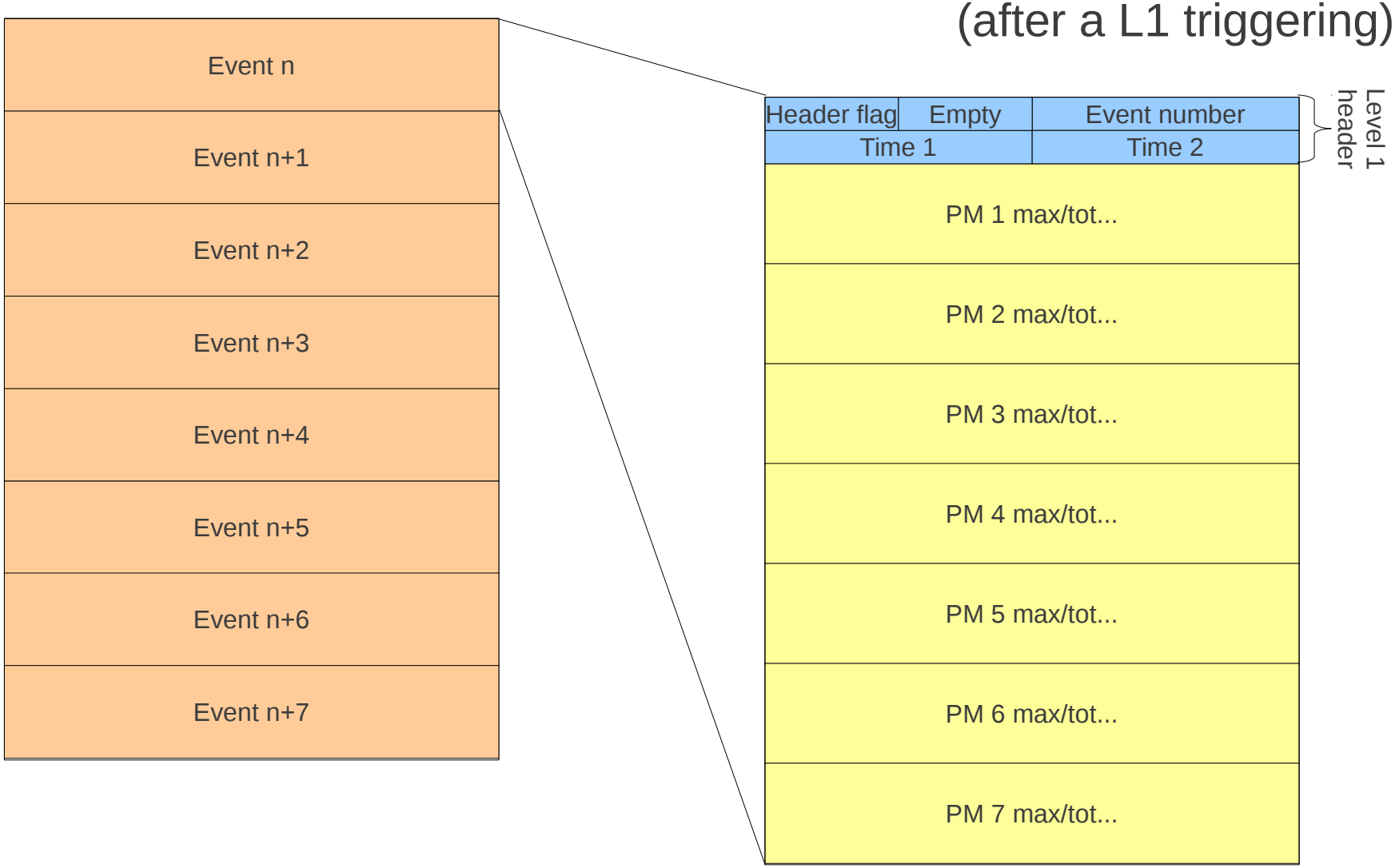
Data format : type 1.0

The front end electronics transmit all events (after a L1 triggering)



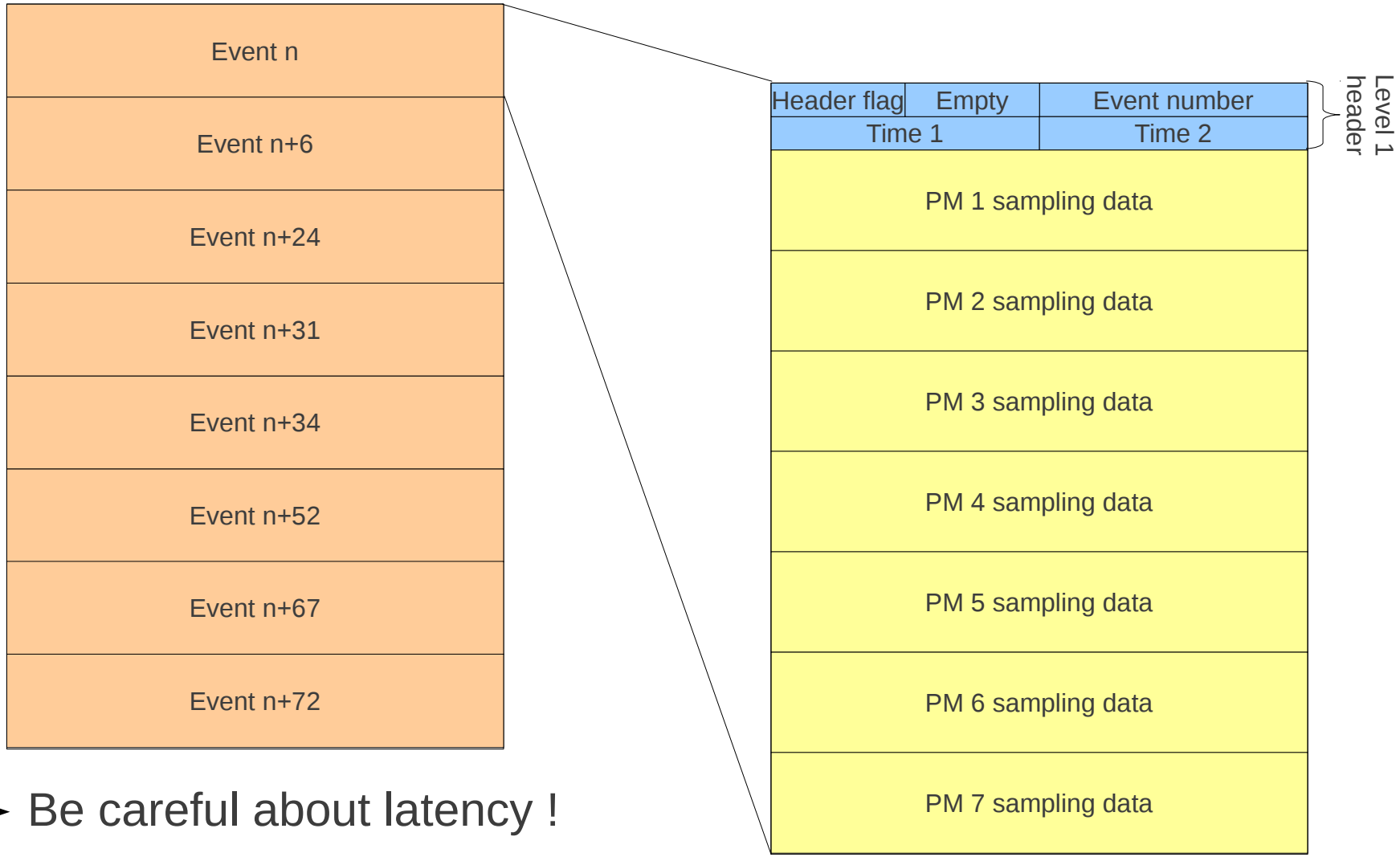
Data format : type 1.1

The front end electronics transmit a single value for each PM for all events (after a L1 triggering)



Data format : type 2.0

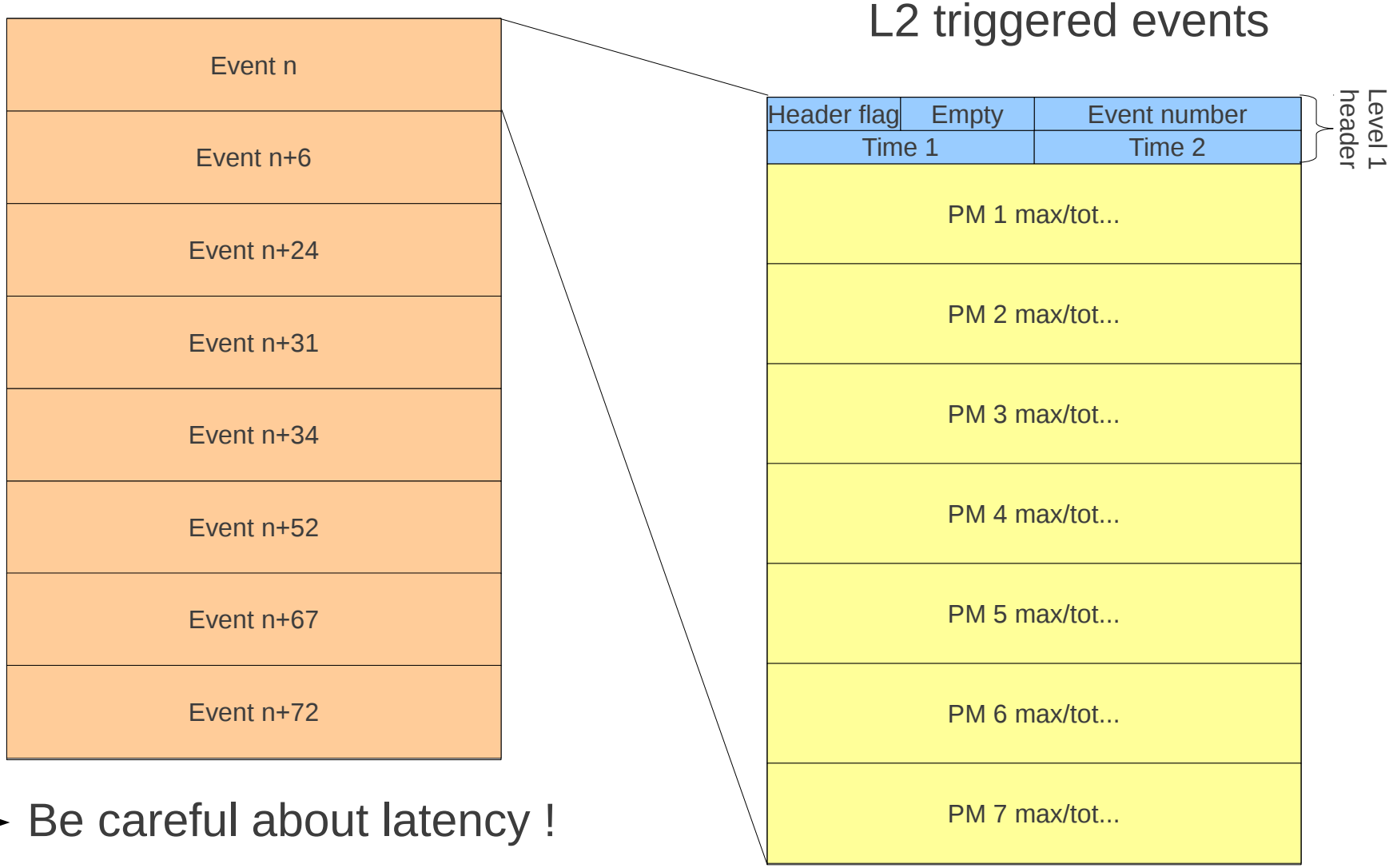
The front end electronics only transmit L2 triggered events



→ Be careful about latency !

Data format : type 2.1

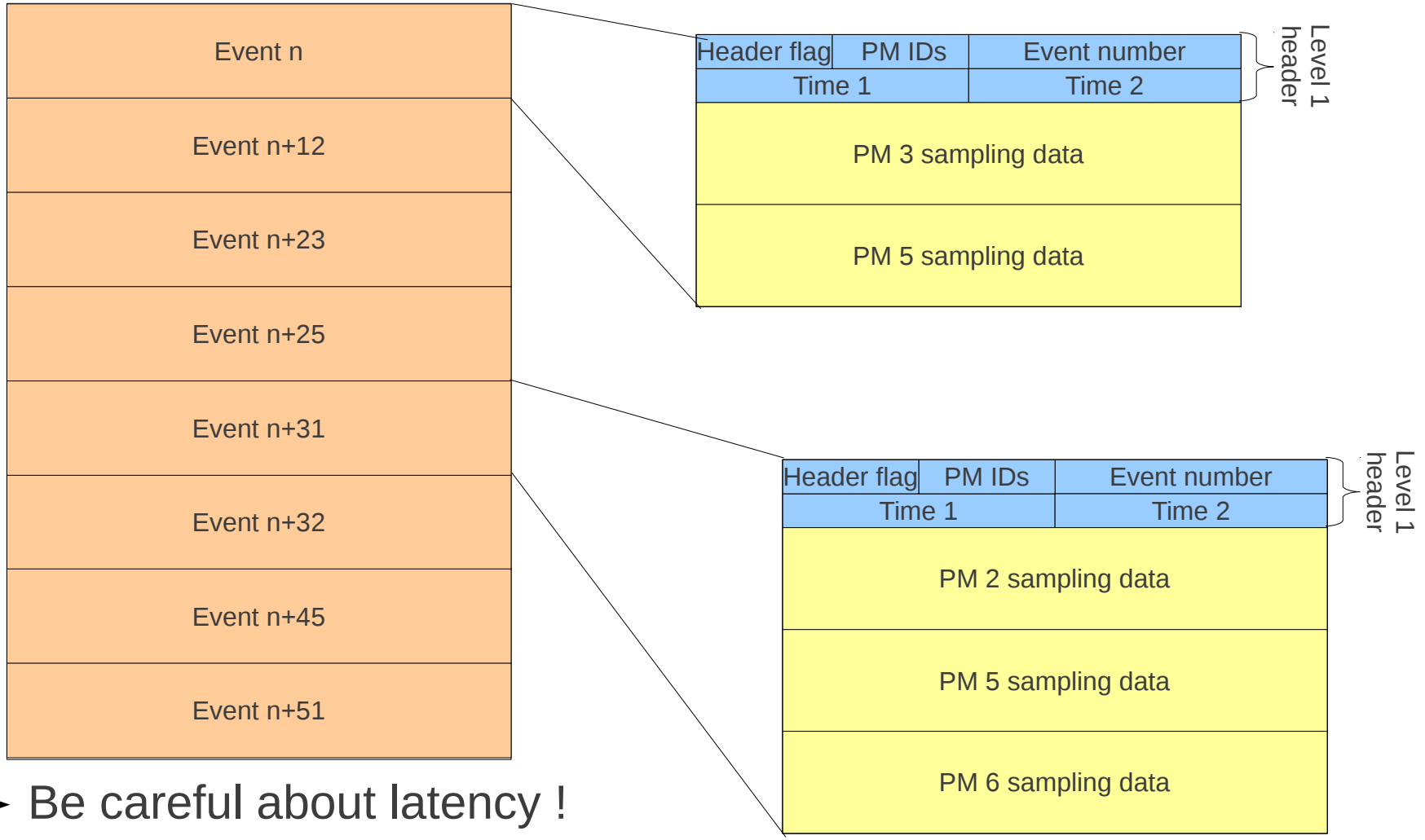
The front end electronics transmit a single value for each PM in L2 triggered events



→ Be careful about latency !

Data format : type 3.0

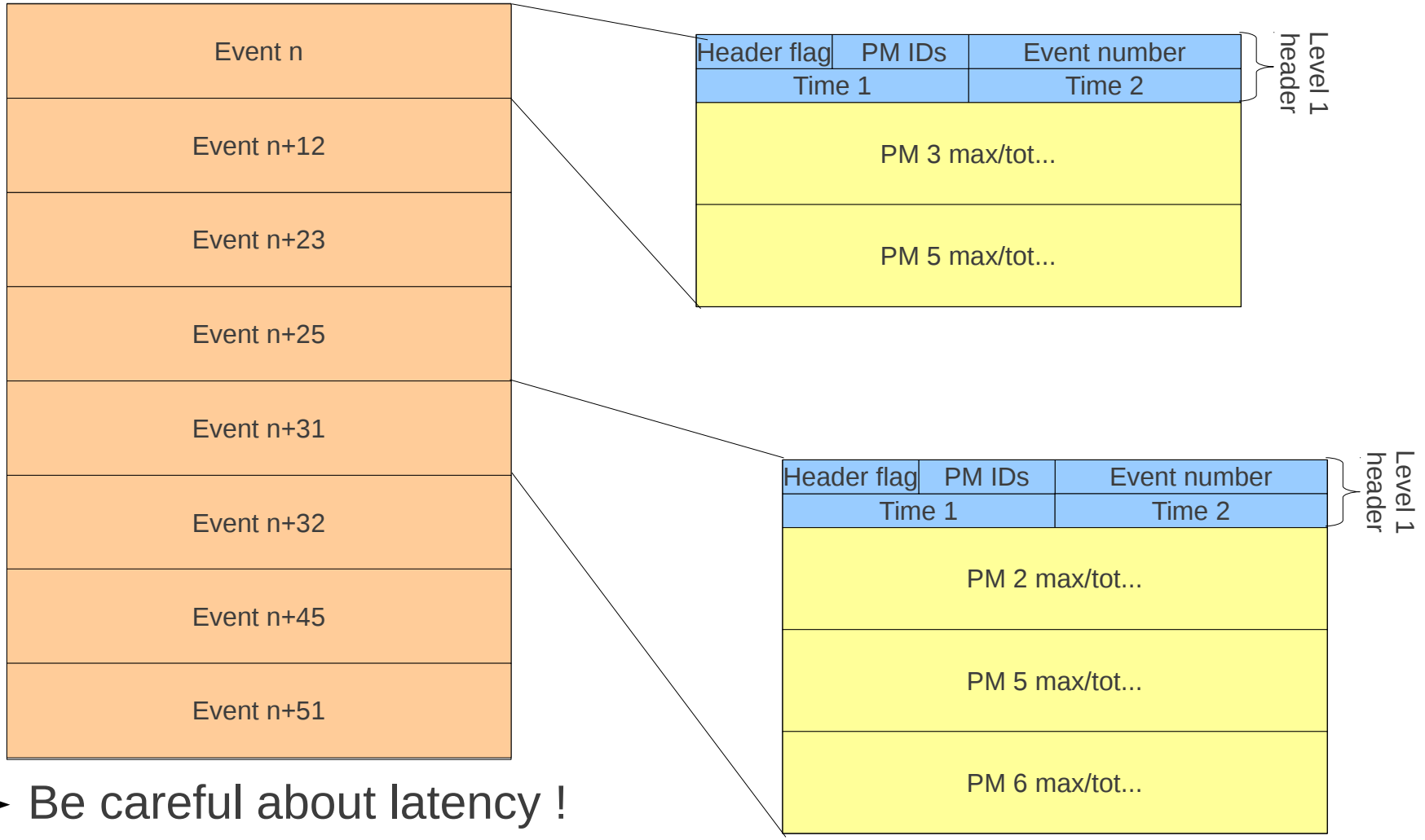
The front end electronics only transmit triggering PM in L2 triggered events



→ Be careful about latency !

Data format : type 3.1

The front end electronics transmit a single value for triggered PM in L2 triggered events

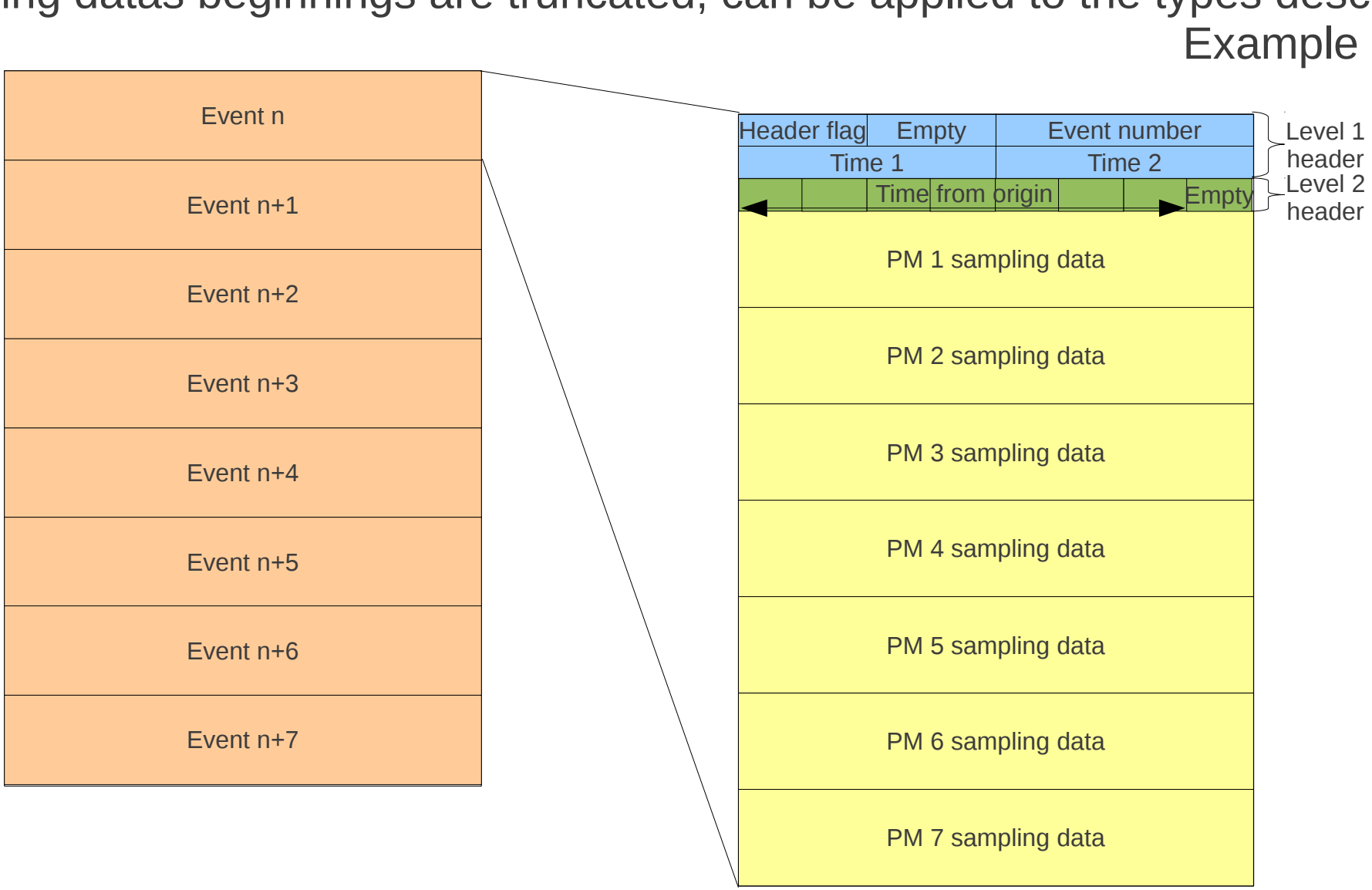


→ Be careful about latency !

Data format : type X.Y.1

sampling datas beginnings are truncated, can be applied to the types describ

Example 1.0

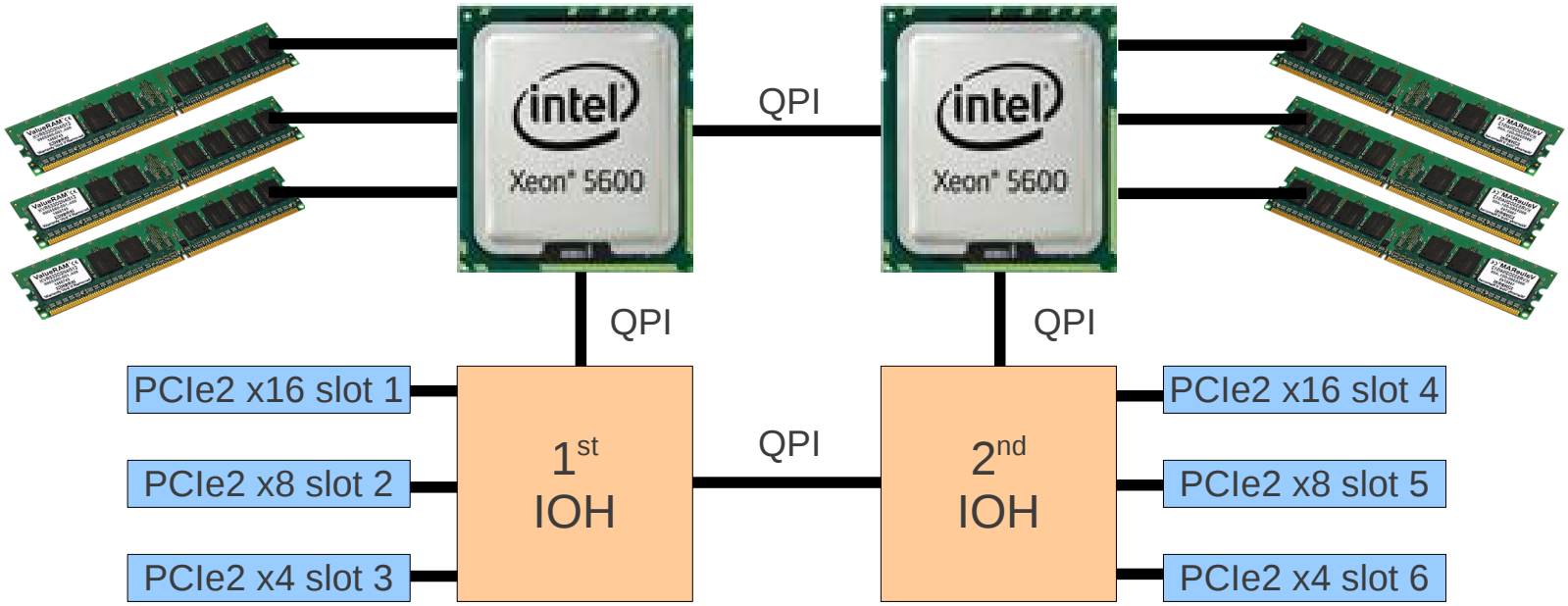


Data format : discussion

The formats exposed are just a proposition and must be discussed with all the concerned teams

Backup

Non Uniform Memory Access



QPI @ 6.4 GT/s bandwidth < DDR3-1333 memory bandwidth