



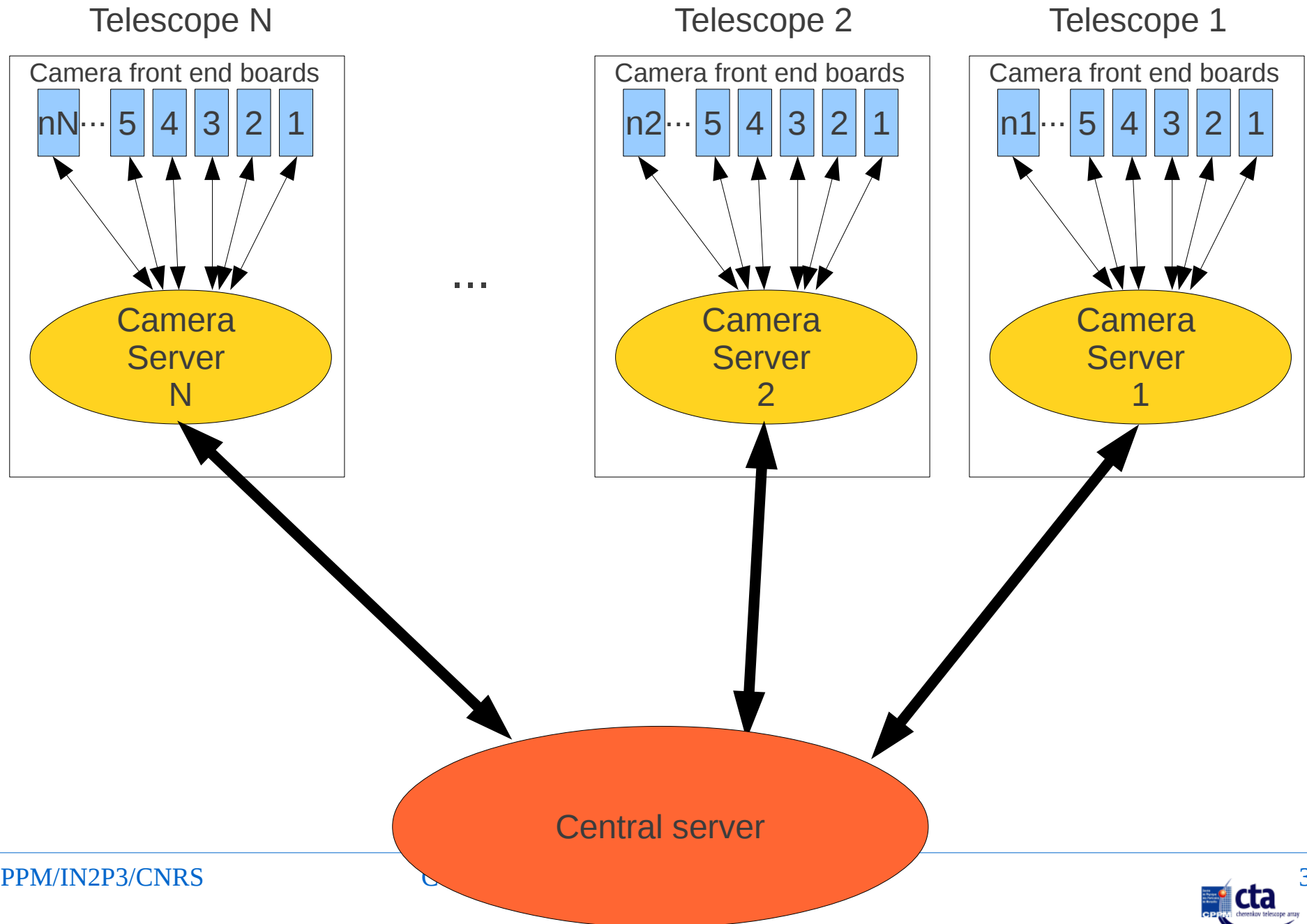
A data acquisition system for the Cerenkov Telescope Array

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

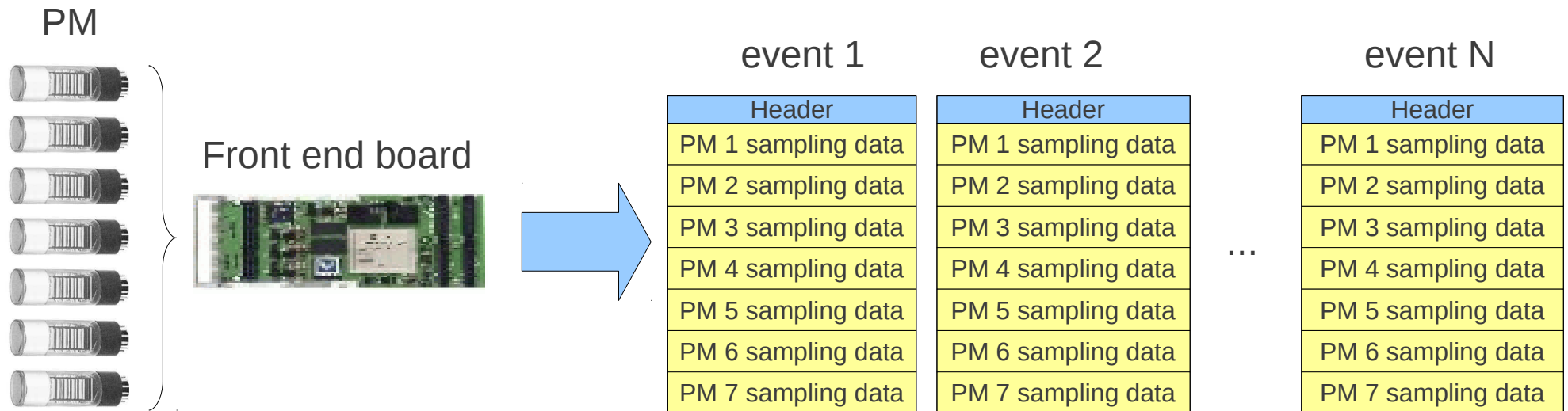
Contact :
houles@cppm.in2p3.fr

Camera server

Global architecture



Camera data flow



Whole Camera ~ 2000 PM -> 300 front end boards

Camera server

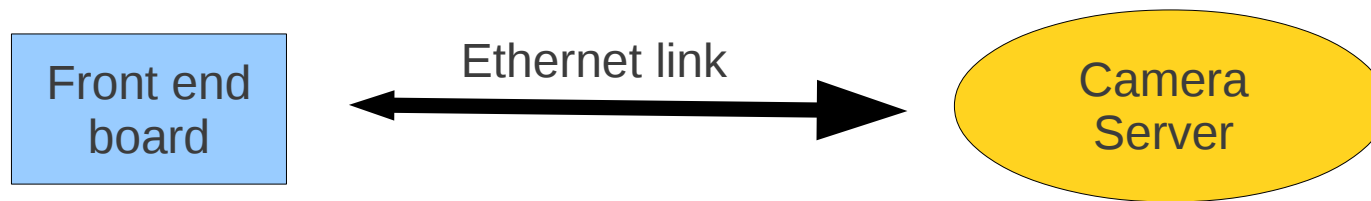
- Build event
- L2 trigger on camera server (L1 on front end) :
 - CPU (SSE, AVX...)
 - GPU
- Compress ?
- Send data to central server (array level)

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 theoretical bandwidth = $10000 * 2000 * 144 = 2.88 \text{ GB/s}$
 $= 23 \text{ Gb/s}$
- 7 detectors for each front end board : 300 boards/camera
 - ➔ Each board generates a flow of $2880/300 = 9.6 \text{ MB/s}$
 $= 77 \text{ Mb/s}$

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

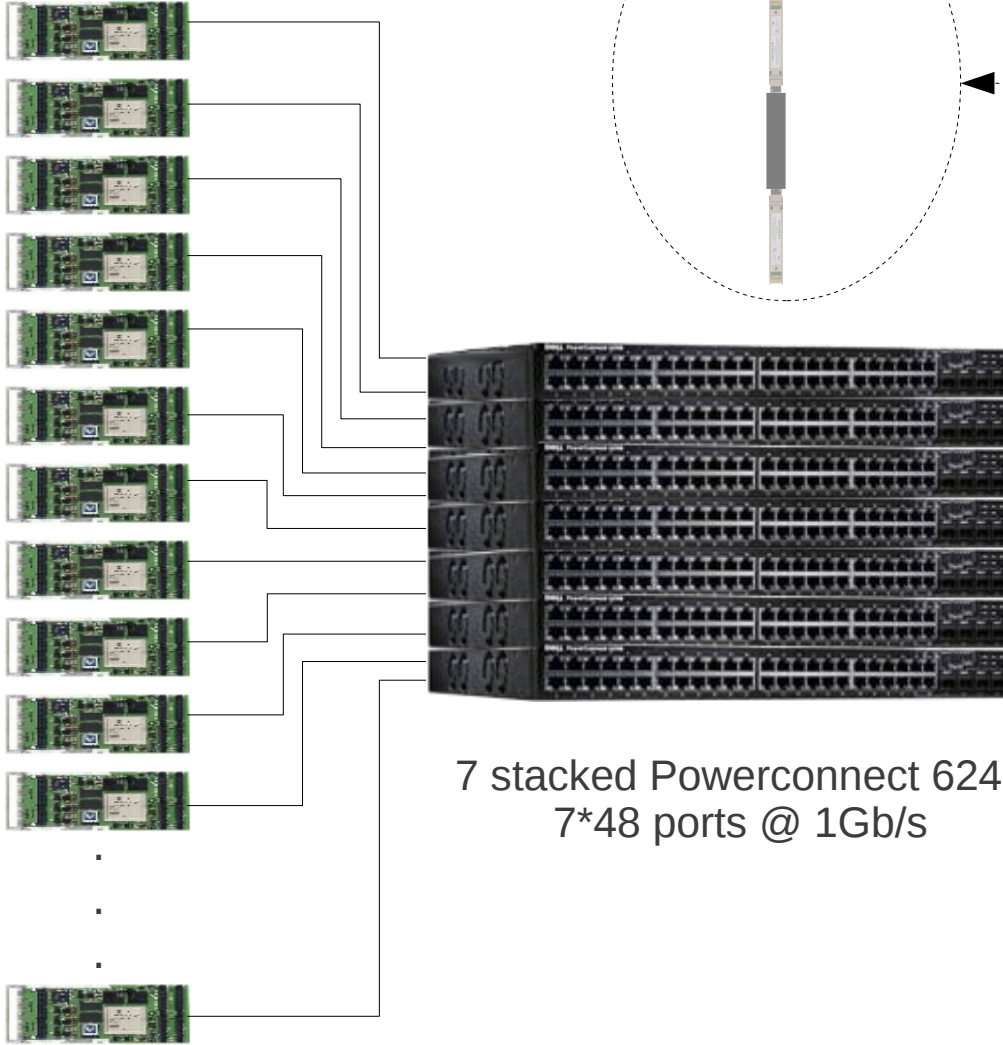
Global architecture



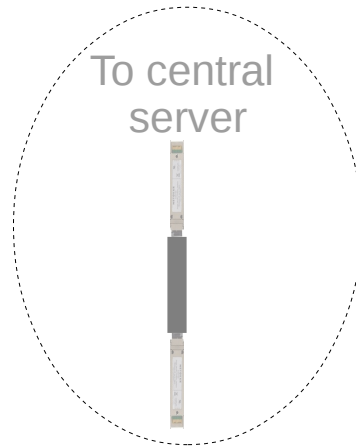
Camera infrastructure

300 * 1Gb/s
Ethernet links

Front end boards

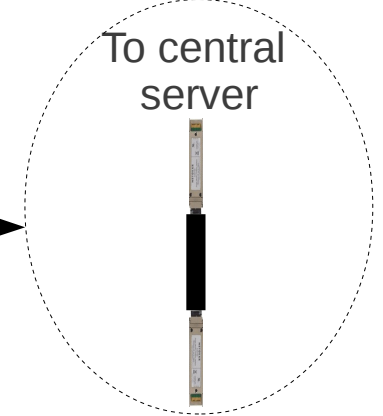


To central
server

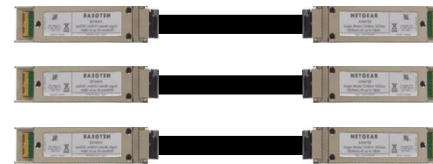


OR

To central
server



3 * 10Gb/s
Ethernet links
SFP+



One or several
Dell T7500 workstations

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 ?

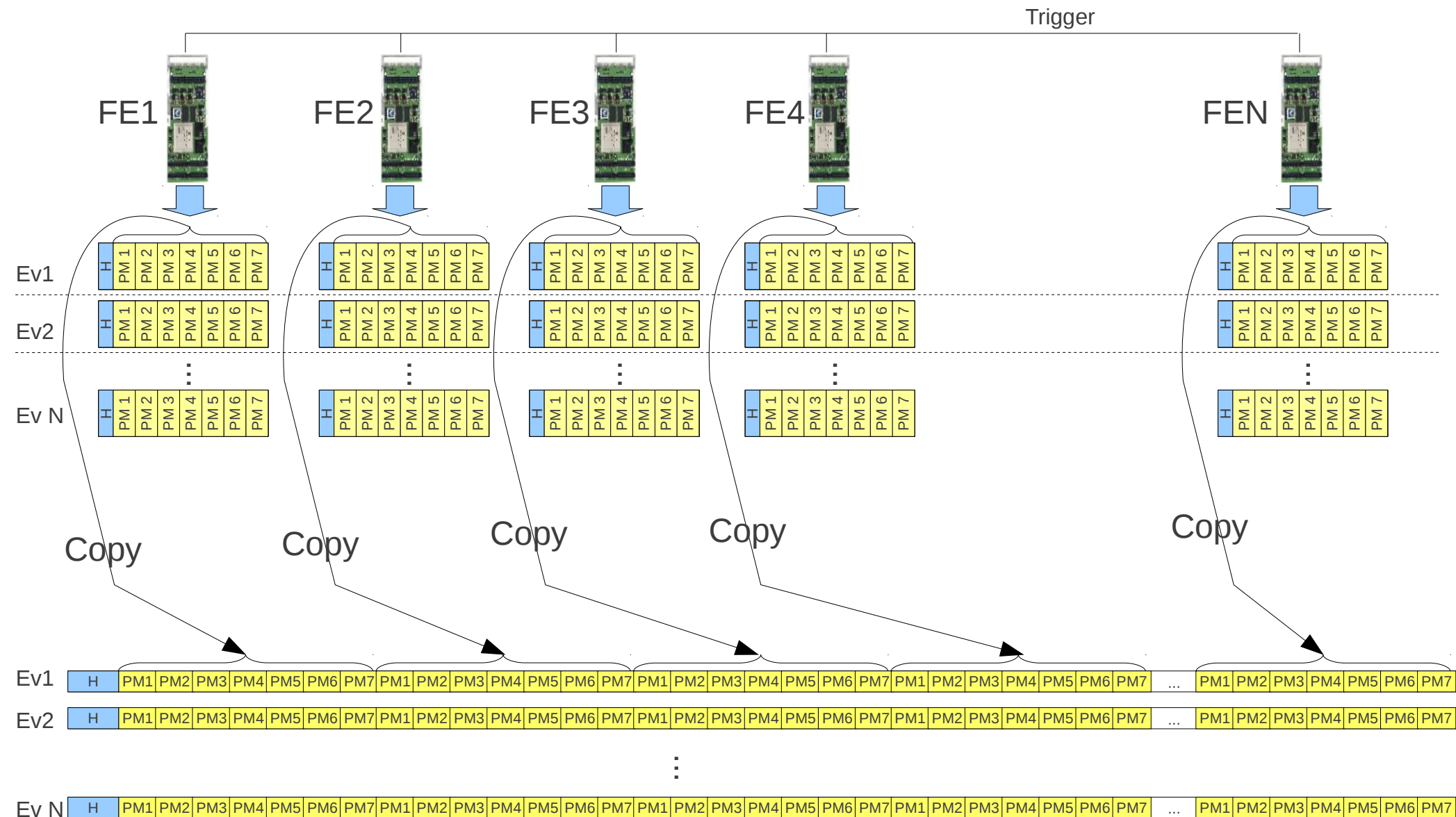
We need a prototype :

- 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)

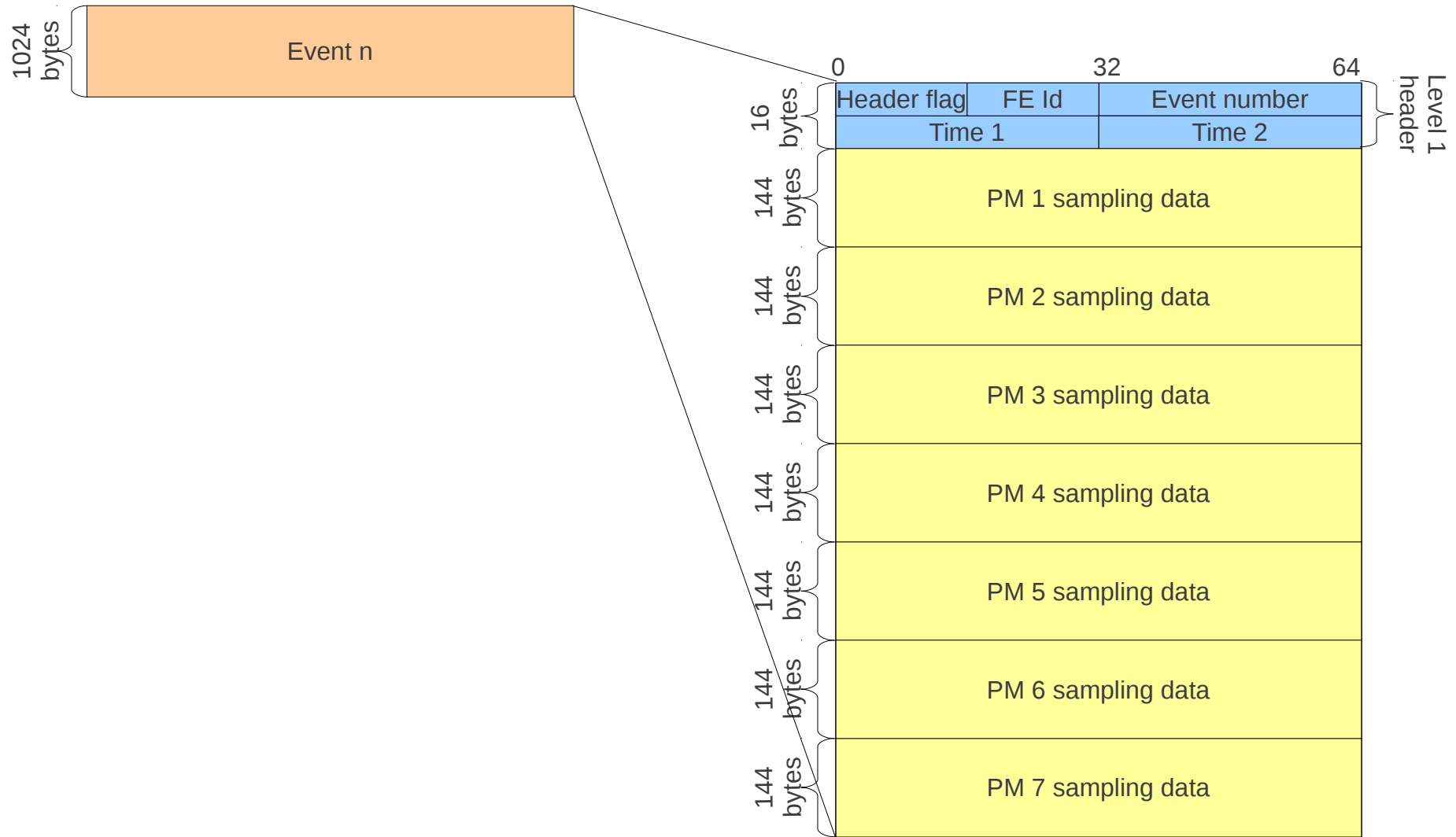
Event builder



Data format : regular frame

1 frame : 1024 bytes

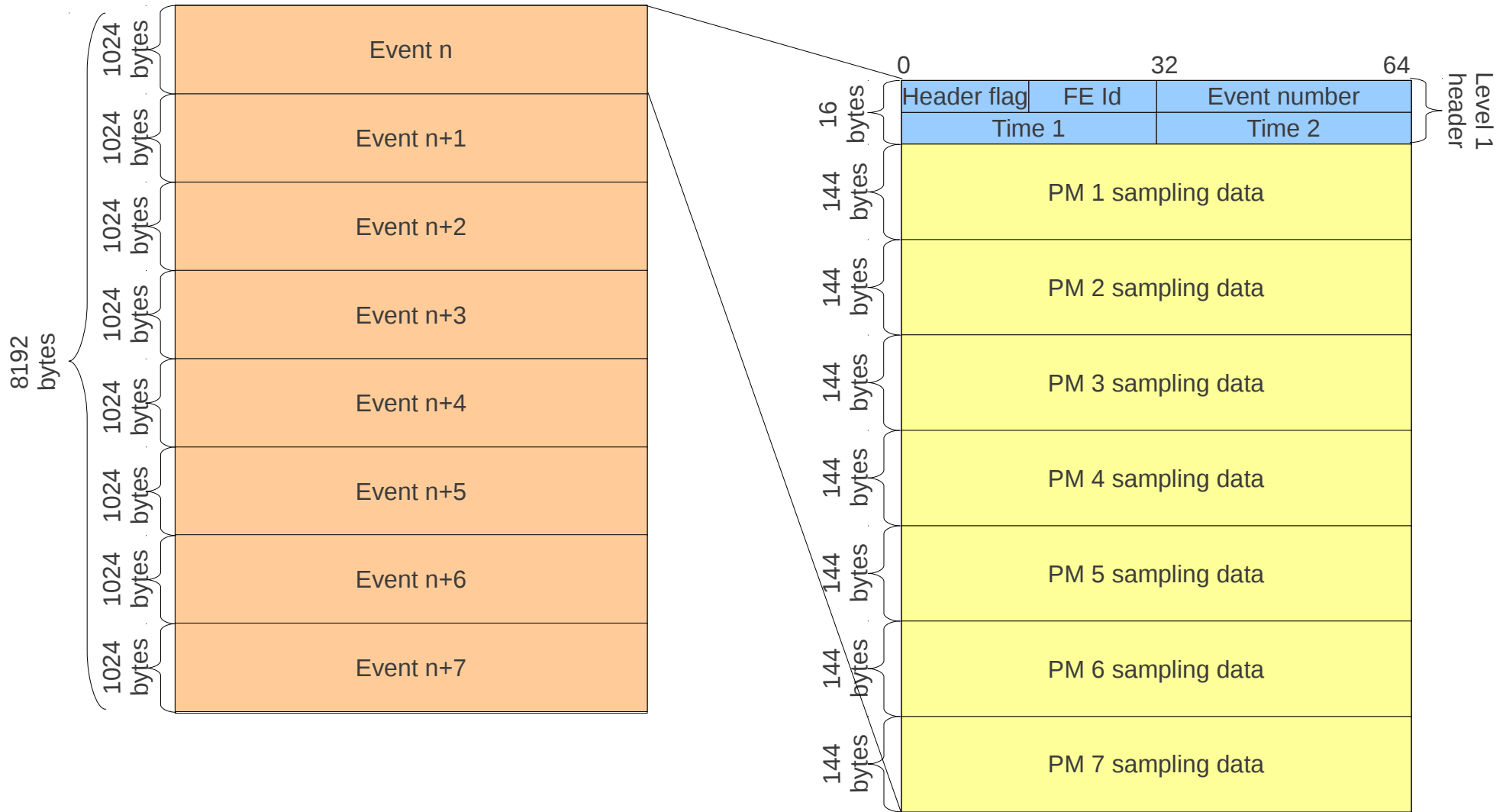
Level 2 triggering on camera server



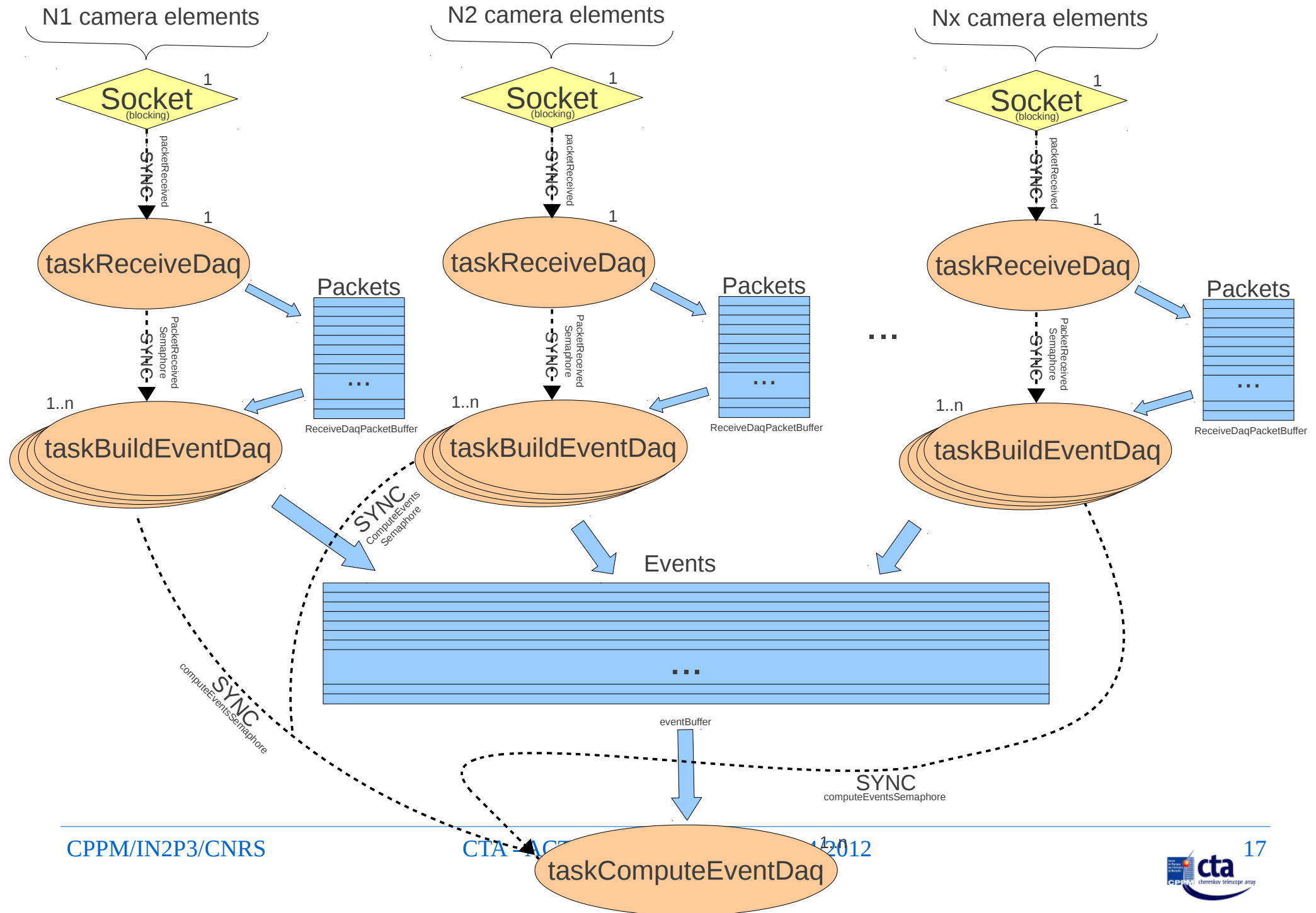
Data format : jumbo frame

1 frame : 8192 bytes

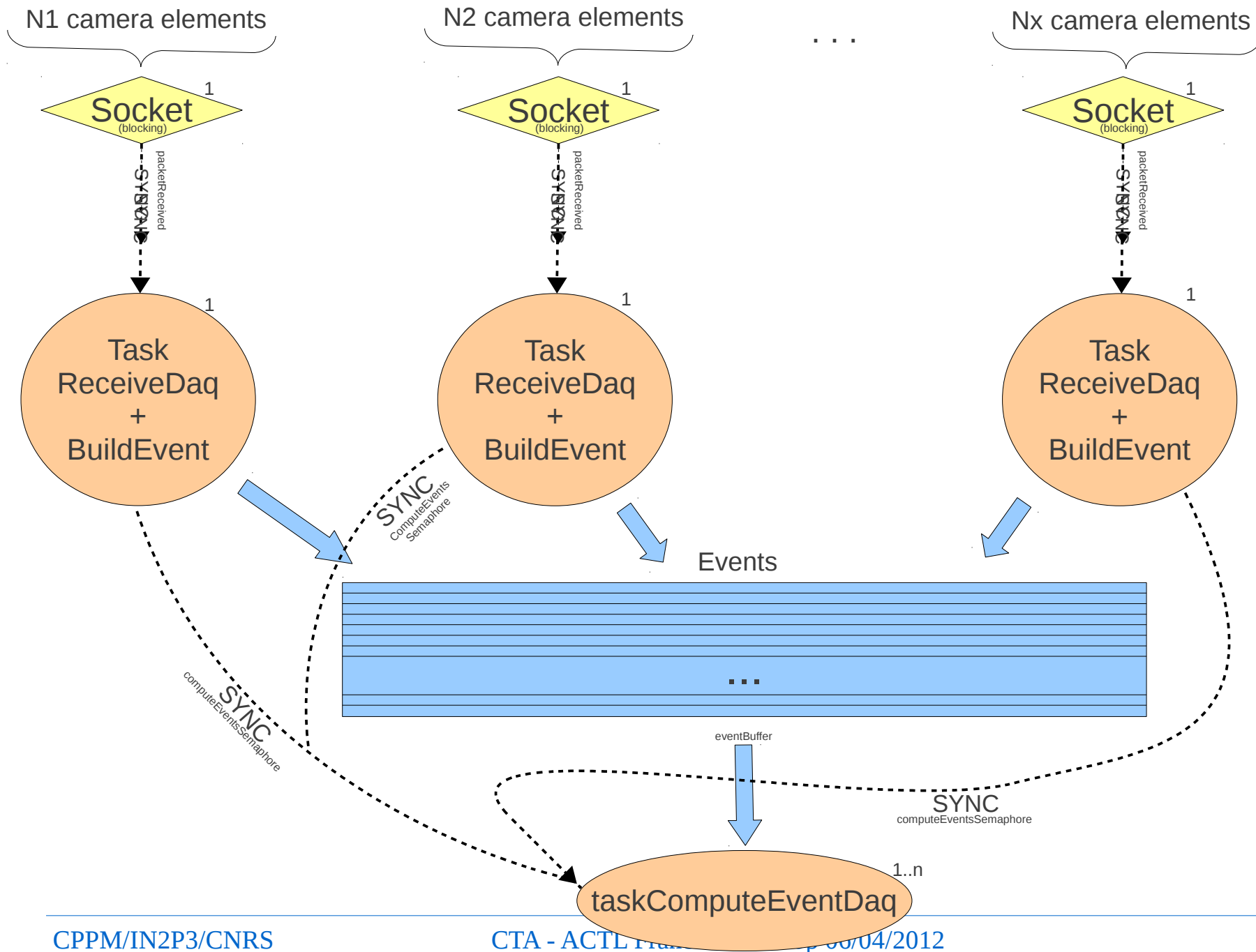
Level 2 triggering on camera server



Software overview : 1st architecture



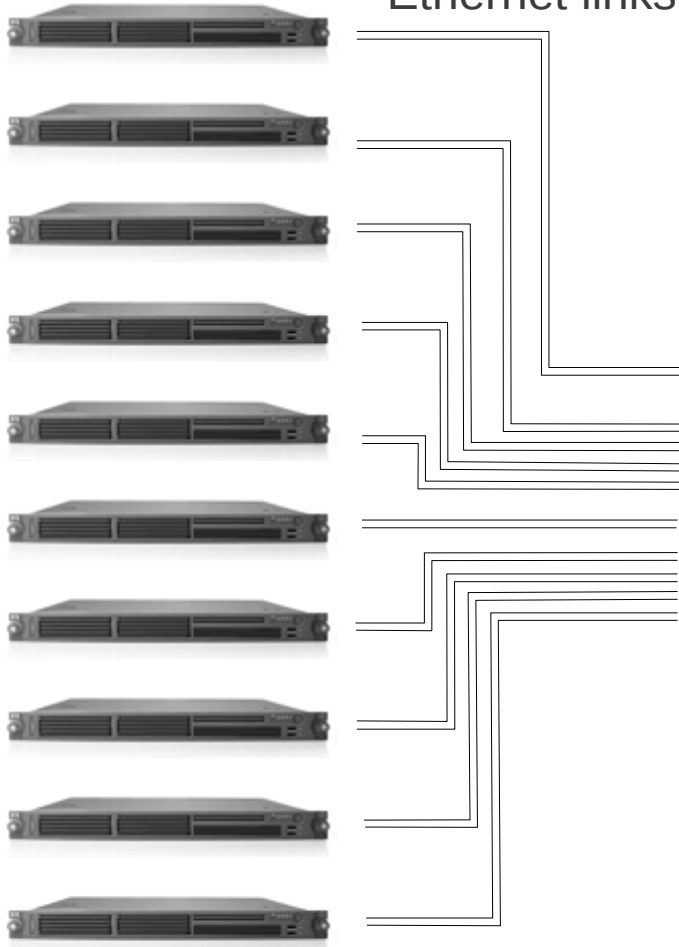
Software overview : 2nd architecture



Stimulation configuration

10 HP servers

10 * 2 * 1Gb/s
Ethernet links

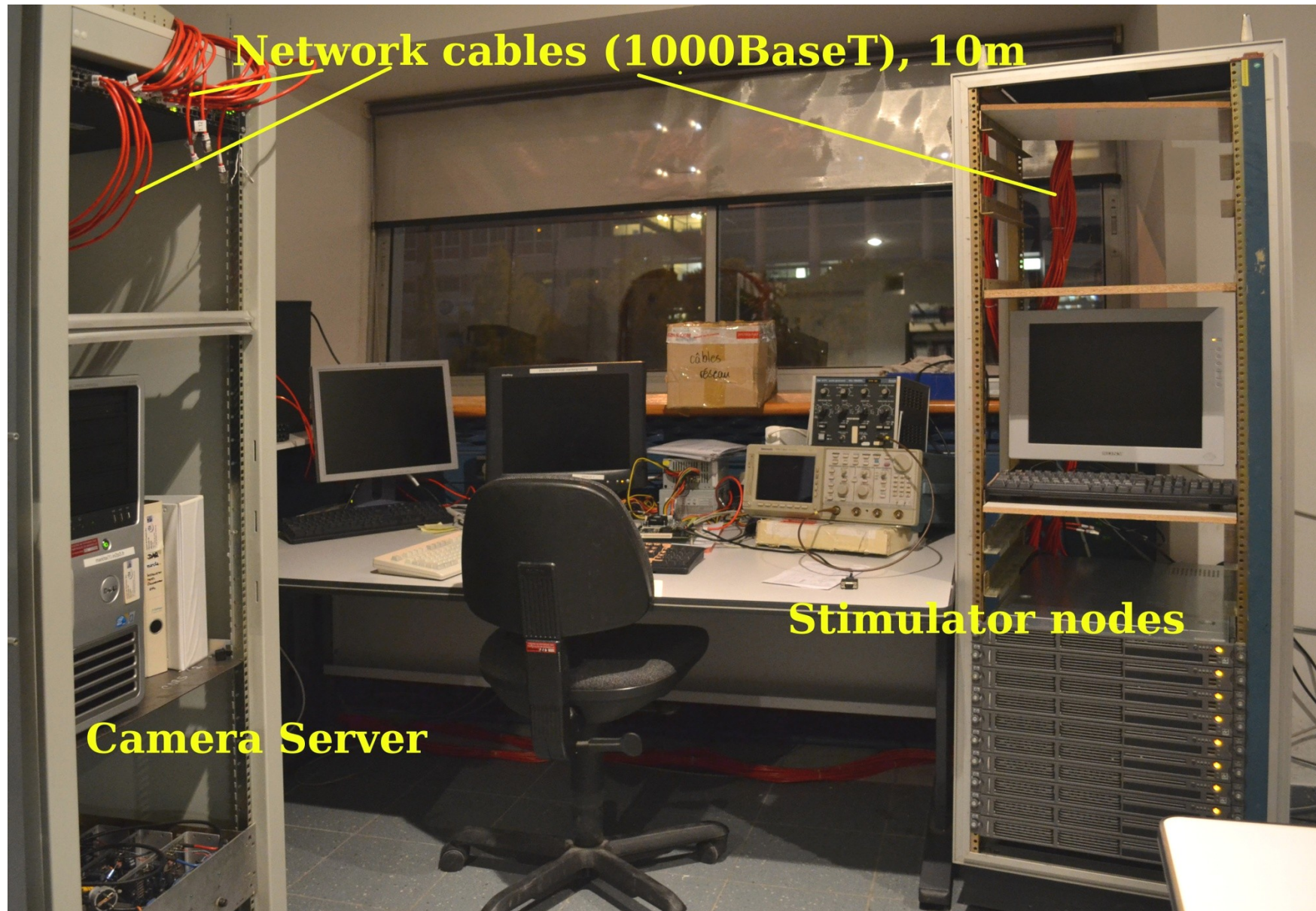


2 * 10Gb/s
Ethernet links
SFP+



One
Dell T7500 workstation

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:

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

The screenshot displays the Acs Command Center v8.1 interface. The main window is titled "(project) - Acs Command Center" and features a menu bar with "Project", "Tools", "Expert", and "Help". The interface is divided into several panels:

- Common Settings:** Includes fields for "Acs Instance" (set to 0), "Cdb Root Dir" (set to /rtm/houles/TEMP/workspace/./acs/test), and radio buttons for "Localhost (single-machine project)" (selected) and "Remote (distributed project)". Under "Remote", there are options for "Use built-in ssh", "Use native ssh", and "Use Acs Daemons". Fields for "Host" and "User" are also present.
- Acs Suite:** Contains buttons for "Start", "Stop", "Kill", and "advanced".
- Deployment Info:** Shows a tree view of the deployment structure, including "Manager on 134.158.16.38, port 3000", "Containers (1)", "Client Applications (2)", and "Components (2)".
- Containers:** A table with columns "Name", "Type", and "Remote Host". It lists a single container named "Container" of type "cpp".
- Object Explorer:** A separate window showing a tree view of the object model, including "eventBuilder" and "Master".
- Operations and Attributes:** A panel for the "EventBuilder" object, showing a list of operations: "closeEventBuilder()", "getDestinationAddress()", "getDestinationPort()", and "launchEventBuilder(String, int, int, String)". It also displays the attribute "componentState" with the value "name".
- Log Console:** A panel at the bottom showing a log of messages, including "Message: Obtained reference to 'Manager'", "Message: Querying root nodes", "Message: Querying type node children of 'eventBuilder'", "Message: Connecting to 'EventBuilder'", "Message: Connected to 'EventBuilder'", "Message: Analysing attributes for 'EventBuilder'", "Message: Analysing operations for 'EventBuilder'", "Message: Invoking 'EventBuilder.launchEventBuilder()', parameters: 'destinationAddress' = '192.168.3.4', 'destinationPort' = '5565', 'bufSize' = '100', 'logFile' = 'coco.txt'", and "EventBuilder.launchEventBuilder --> Return value: null".

Stimulator

Need for a real stimulator

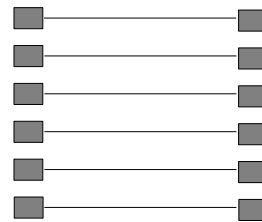
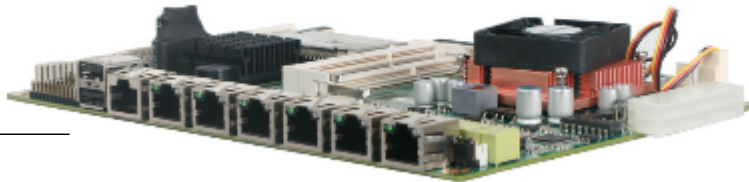
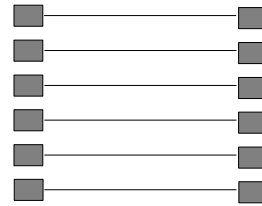
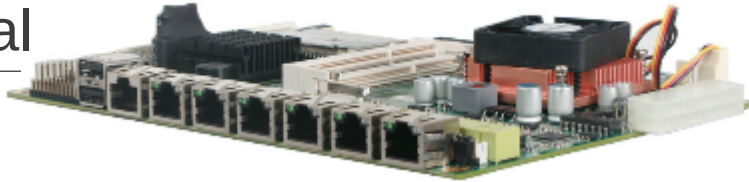
Need a stimulator to make :

- 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

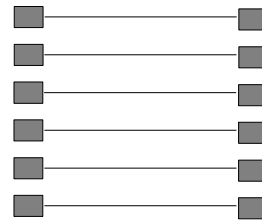
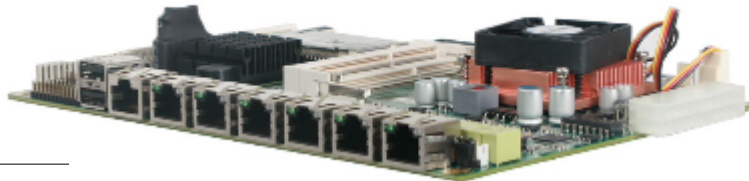
Testing configuration

Sync
Signal

EVOC NET-1820

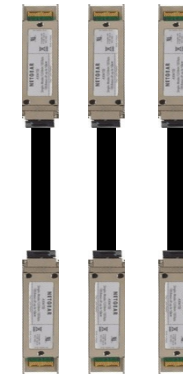
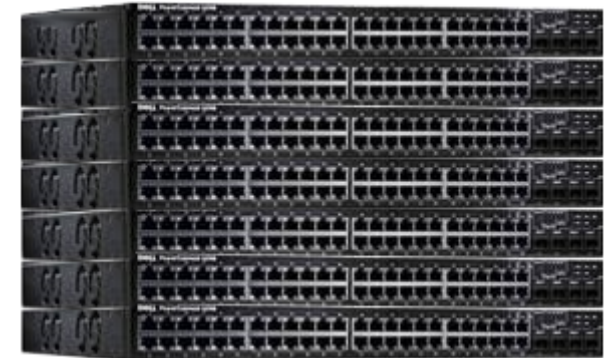


⋮



1Gb/s links

Powerconnect 6248
stack

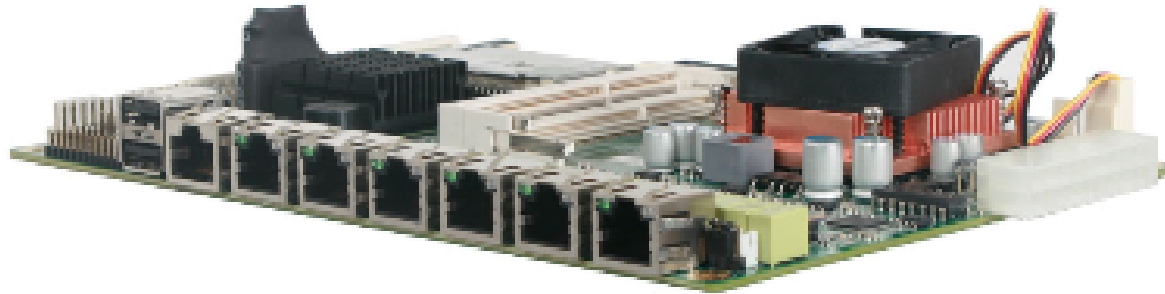


10 Gb/s
links

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 ?)

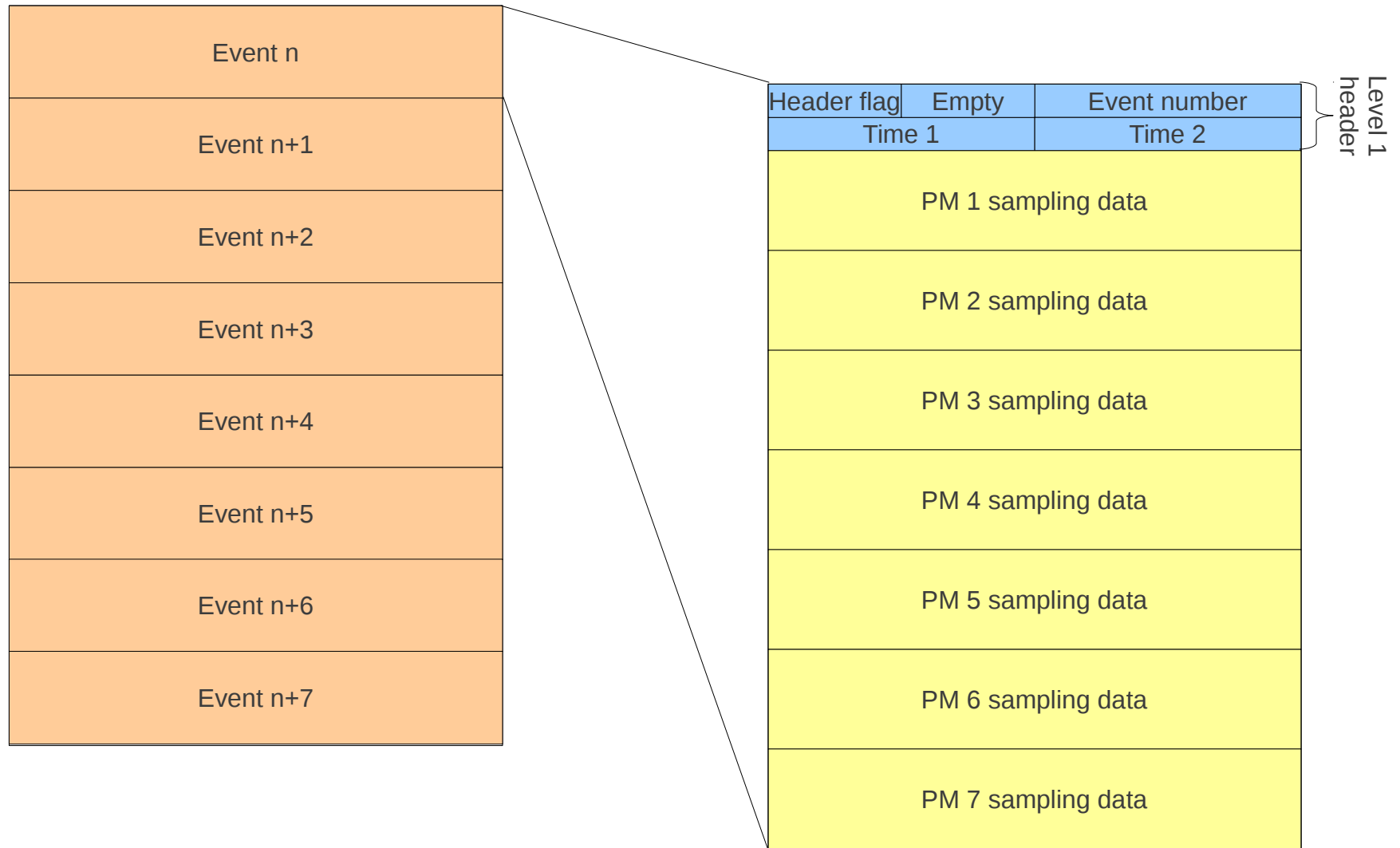
- Work with slow control and array server communications teams

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

Interface definition

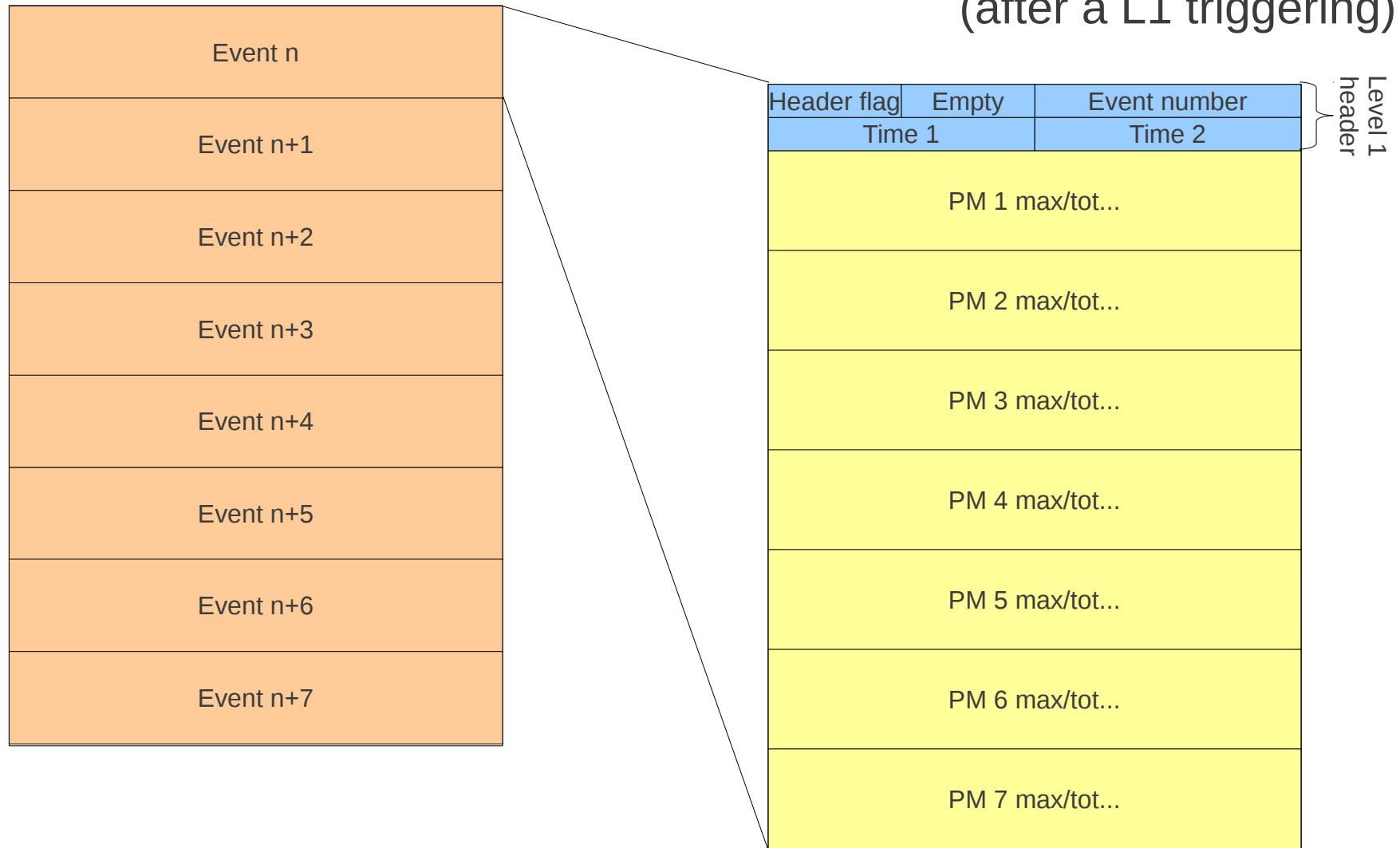
Data format : type 1.0

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



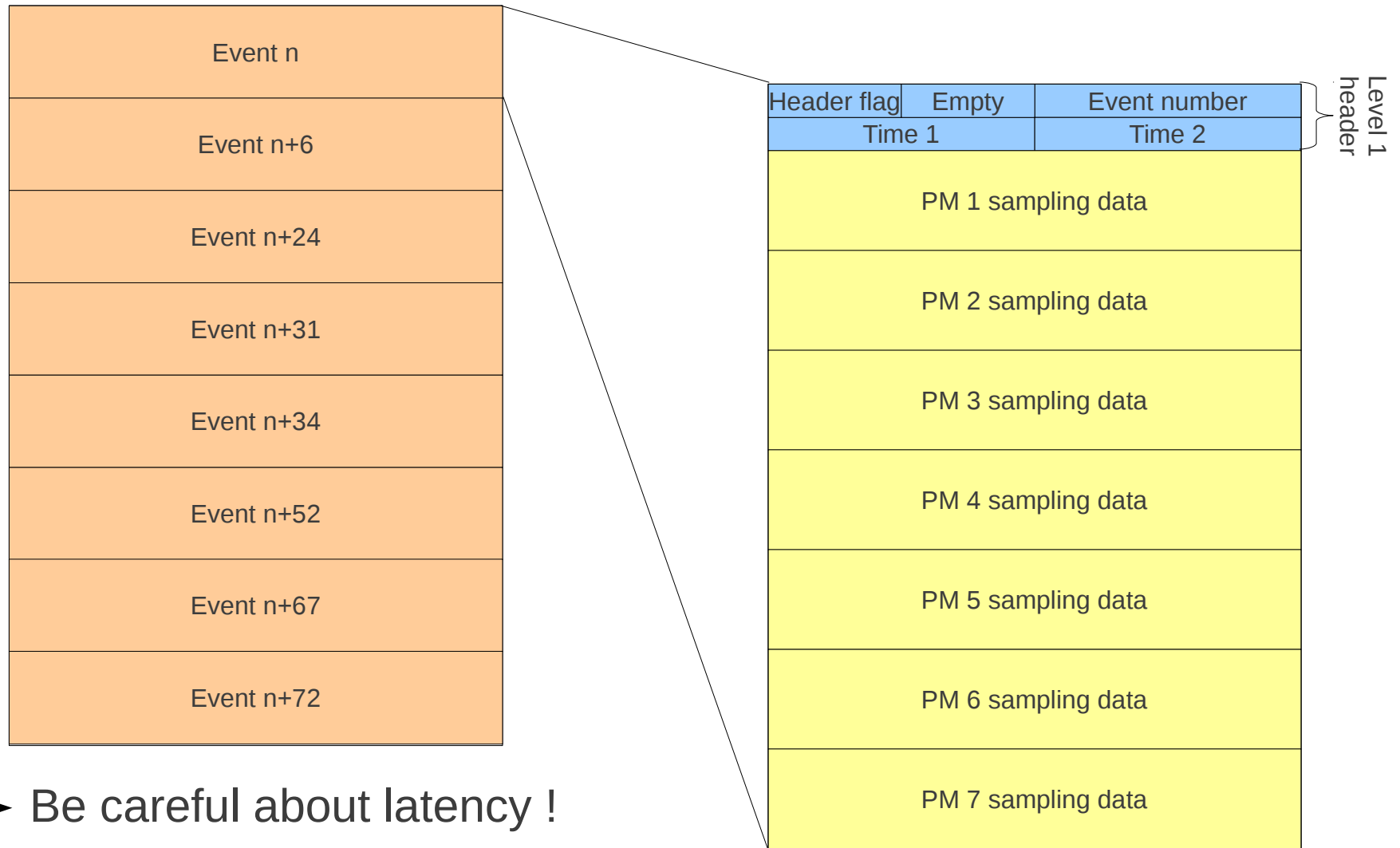
Data format : type 1.1

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



Data format : type 2.0

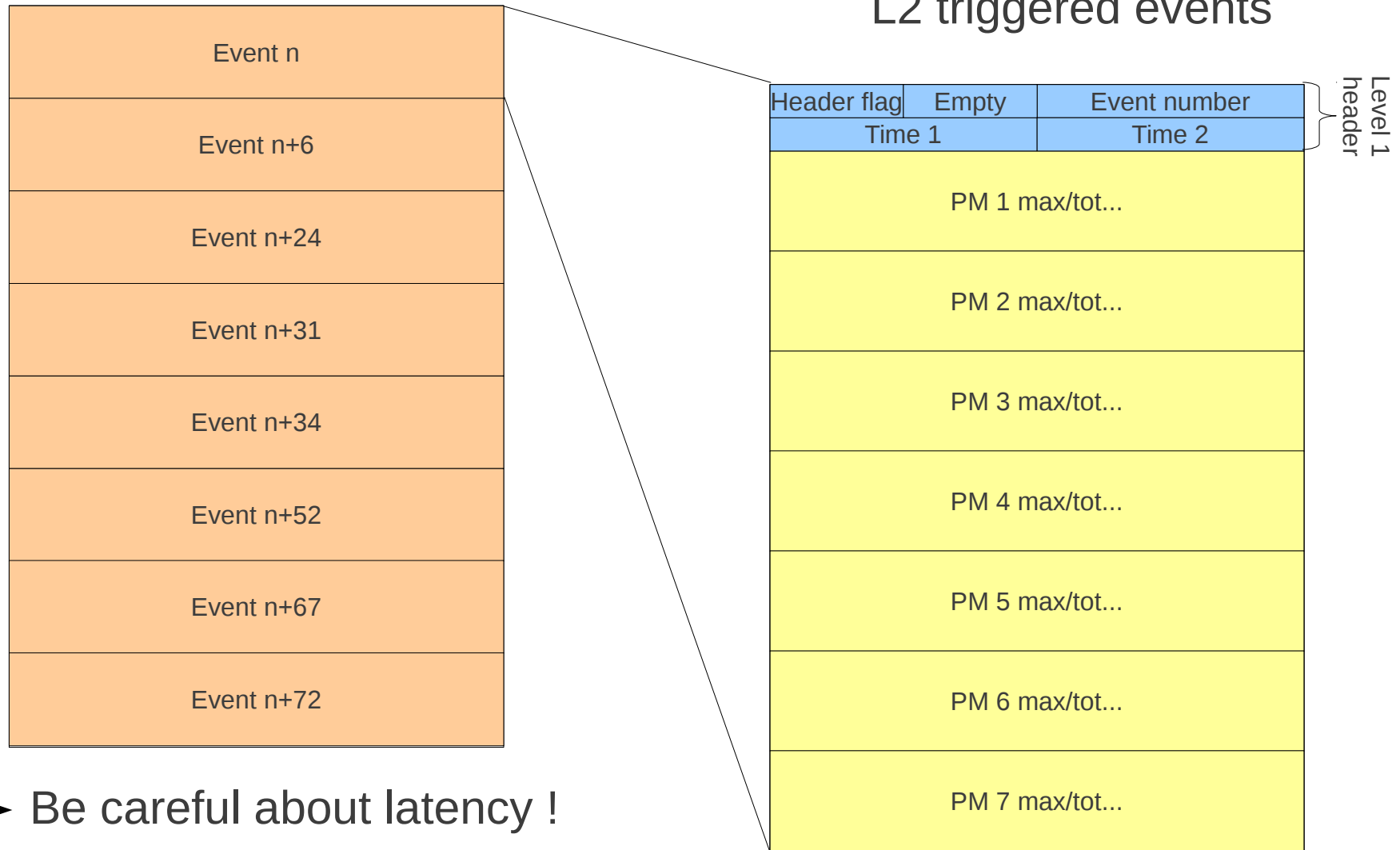
The front end electronics only transmit L2 triggered events



➔ Be careful about latency !

Data format : type 2.1

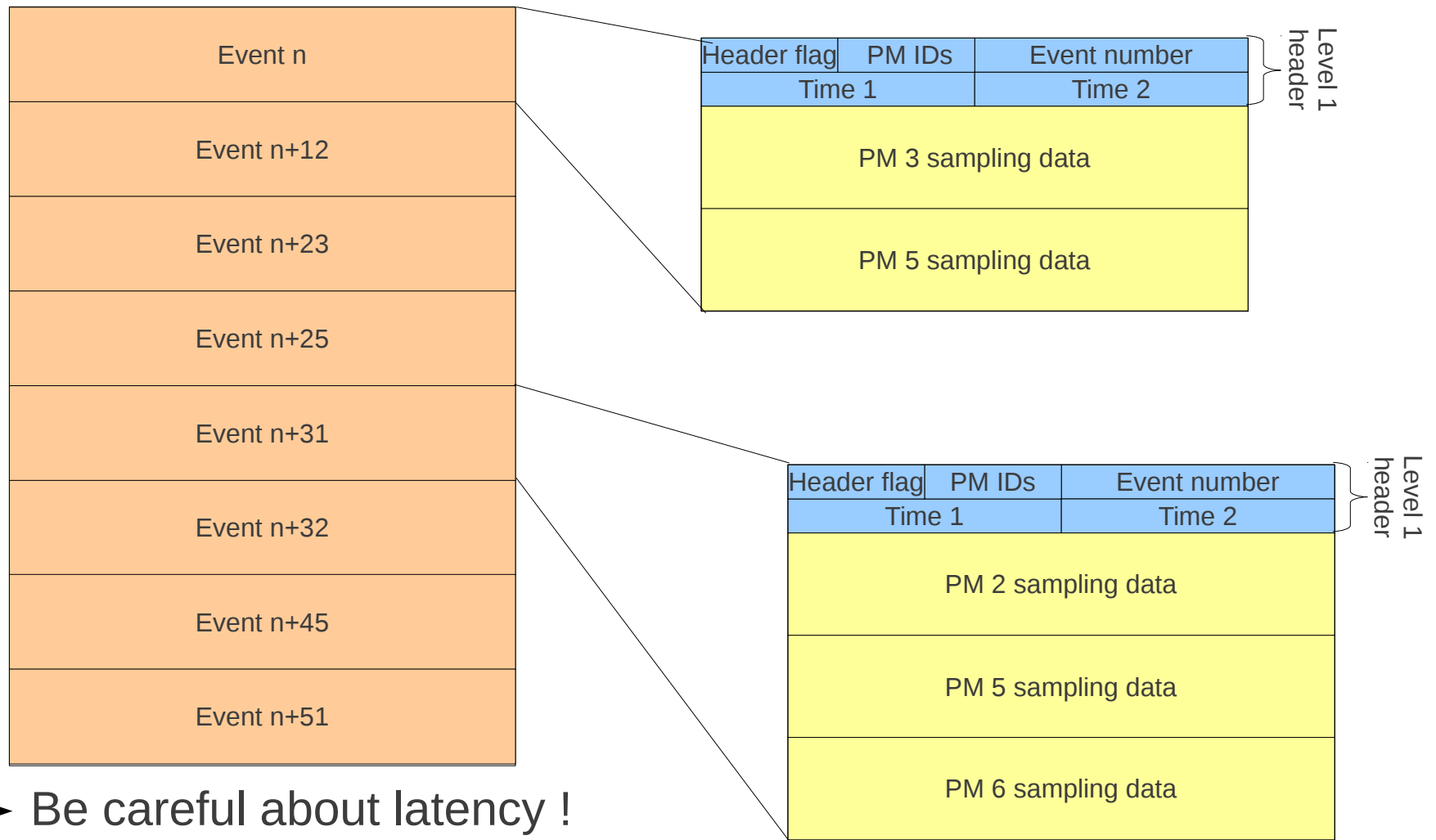
The front end electronics transmit a single value for each PM in L2 triggered events



→ Be careful about latency !

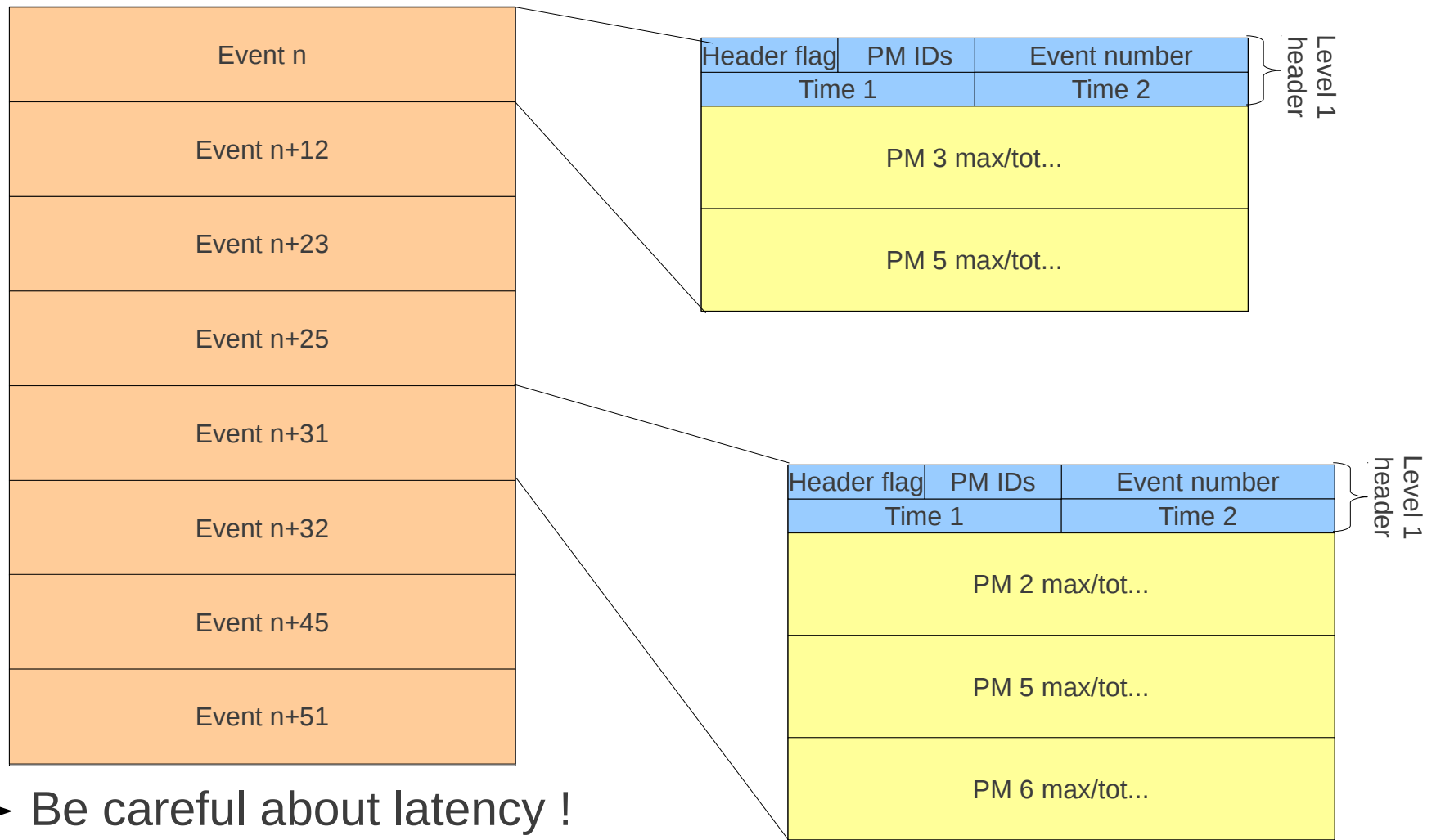
Data format : type 3.0

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



Data format : type 3.1

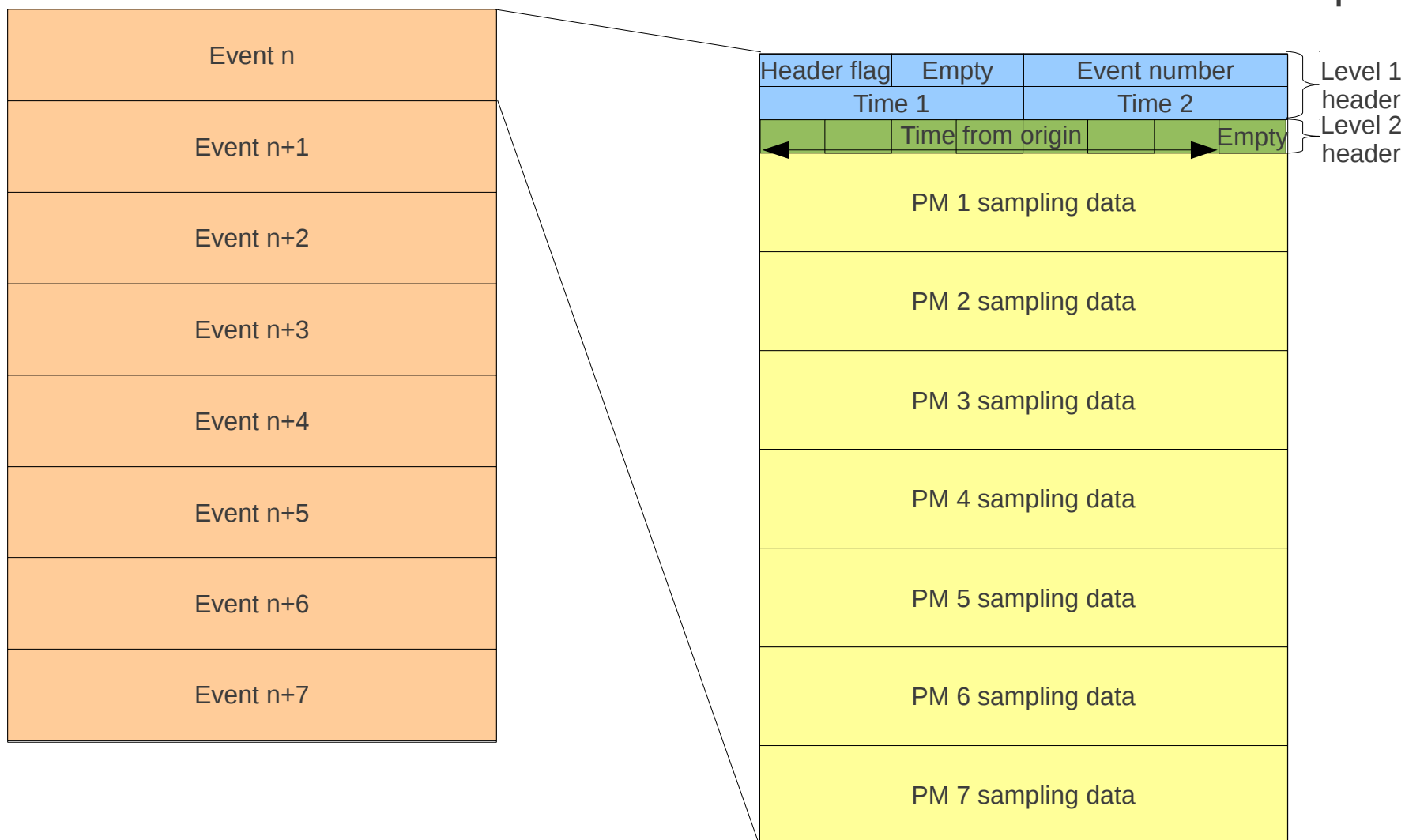
The front end electronics transmit a single value for triggered PM in L2 triggered events



Data format : type X.Y.1

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

Example 1.0

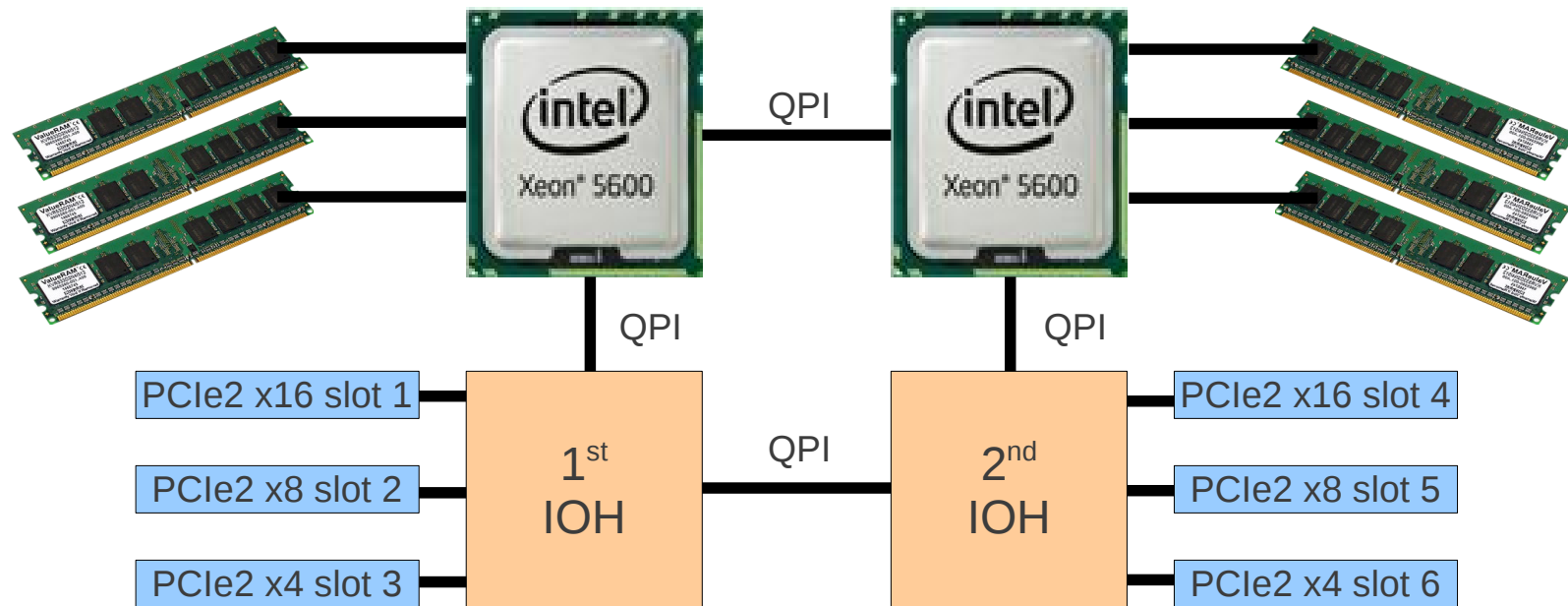


Data format : discussion

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