

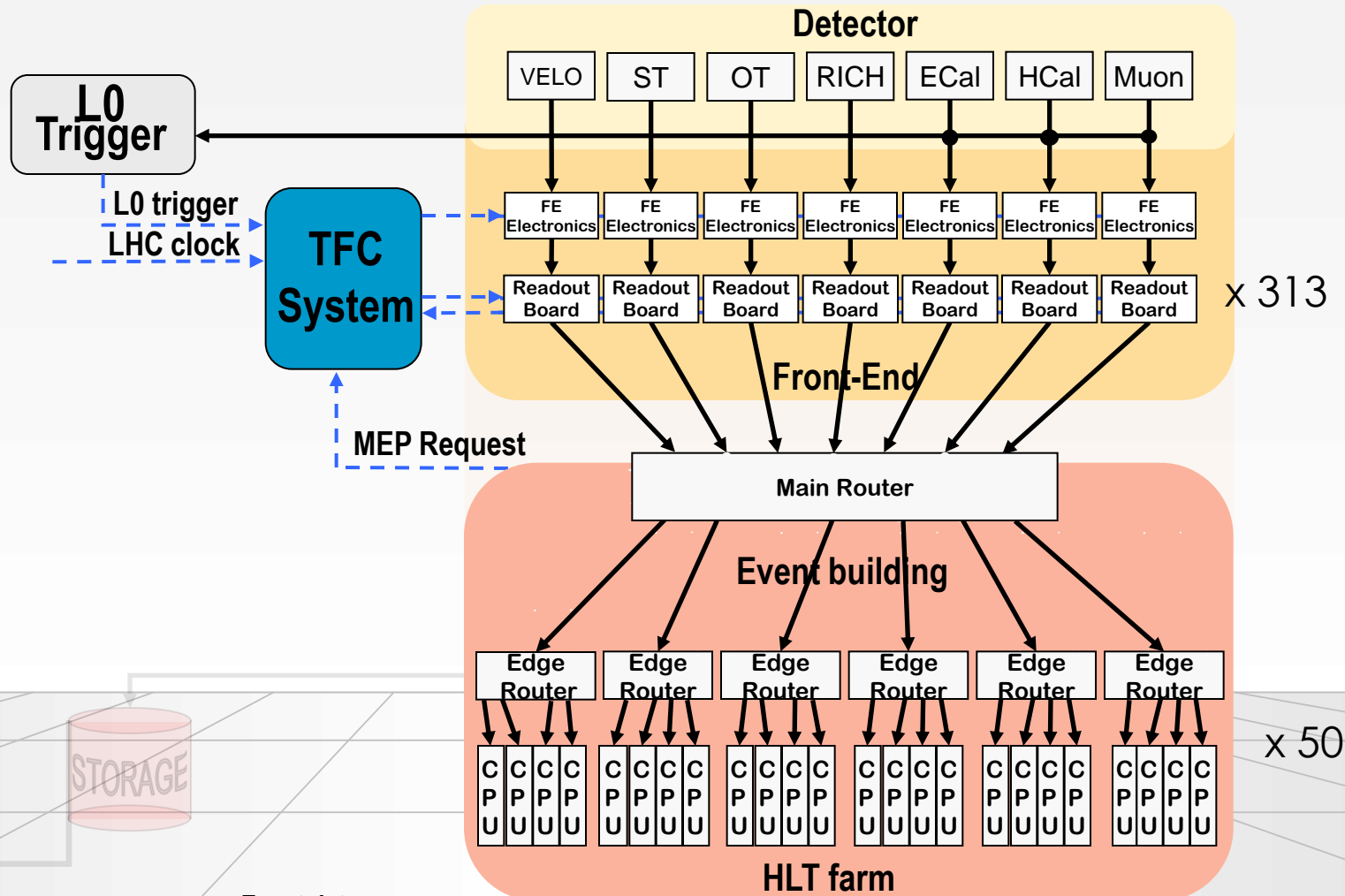
DAQ interface + implications for the electronics

Niko Neufeld

LHCb Electronics Upgrade

June 10th, 2010

The LHCb 1 MHz DAQ

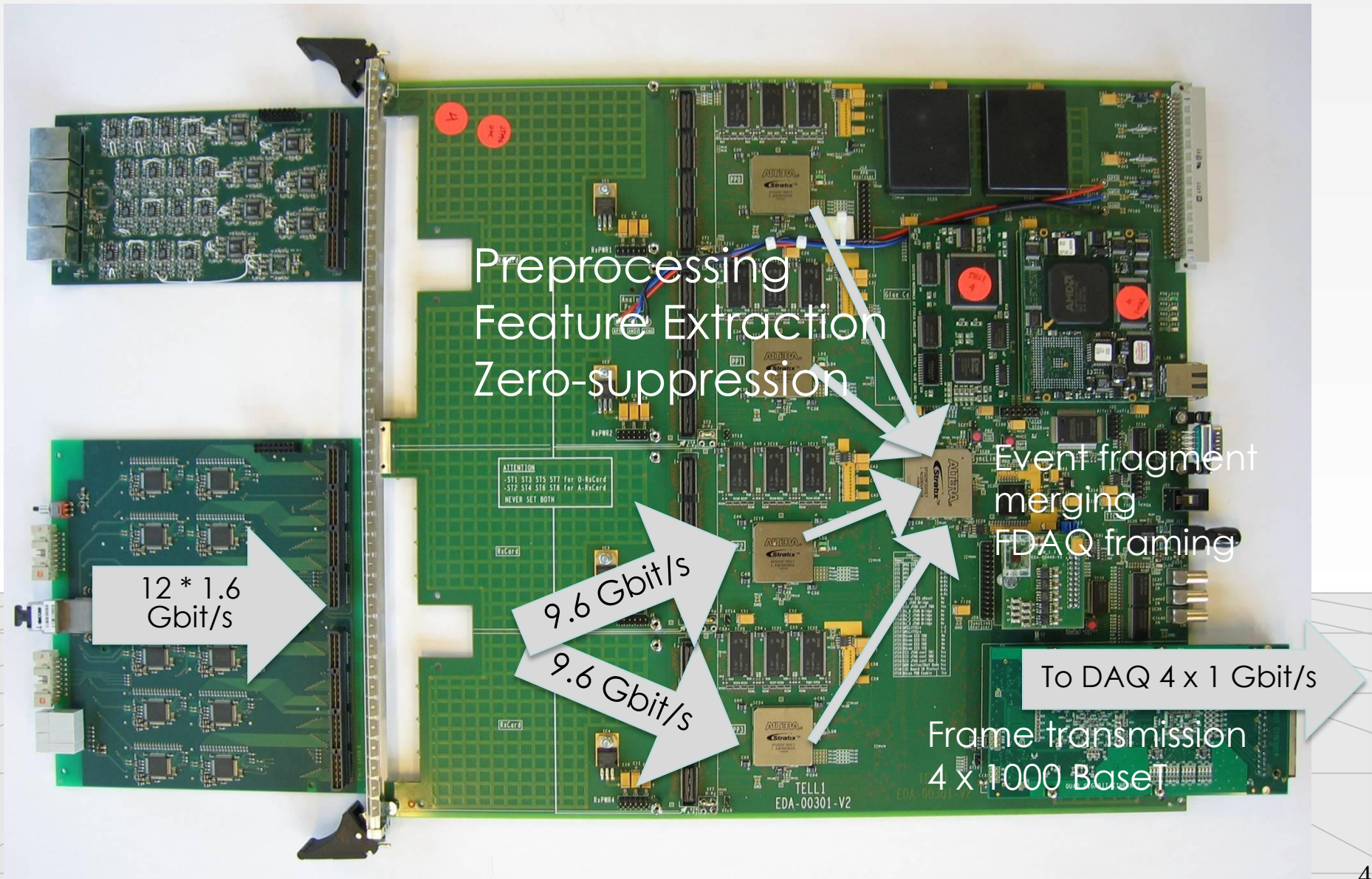


— Event data
 - - - Timing and Fast Control Signals
 — Control and Monitoring data

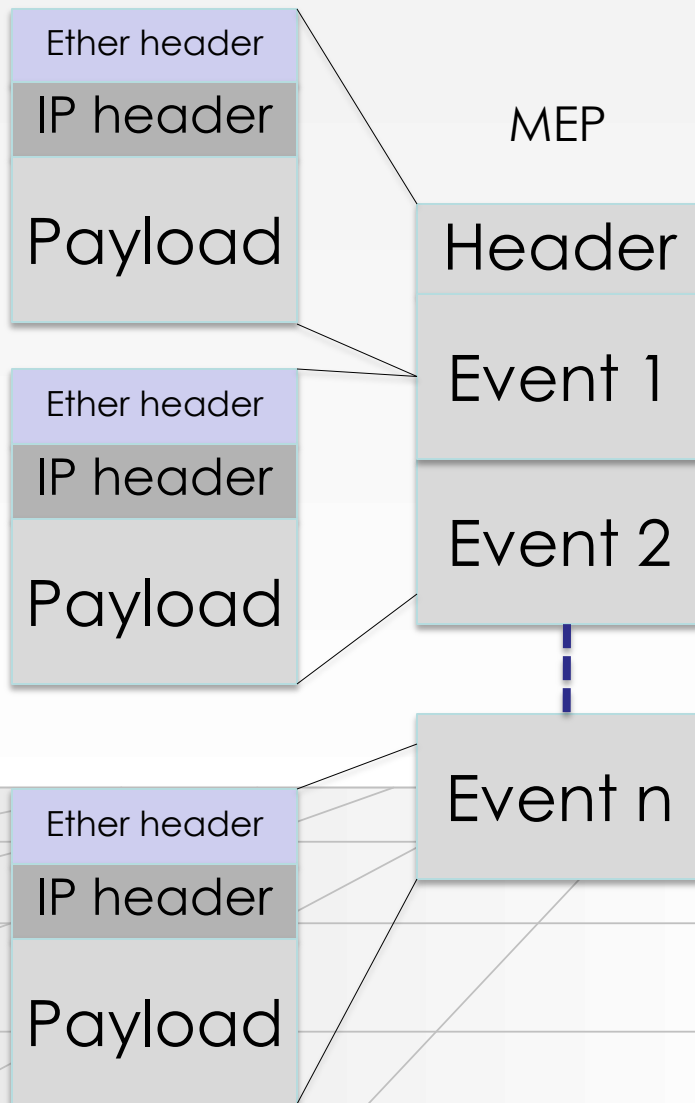
A few numbers you know...

# links (UTP Cat 6)	~ 3000
Event-size (total – zero-suppressed)	35 kB
Read-out rate	1 MHz
# read-out boards	313
output bandwidth / read-out board	up to 4 Gigabit/s (4 Ethernet links)
# farm-nodes	550 (will grow to 1500)
input bandwidth / farm-node	1 Gigabit (1 dedicated Ethernet link)
# core-routers	1 (1260 ports)
# edge routers	50 (48 ports)
Event-rate to storage	2000 Hz (nominal)

The “data-source” TEL(UK)L1

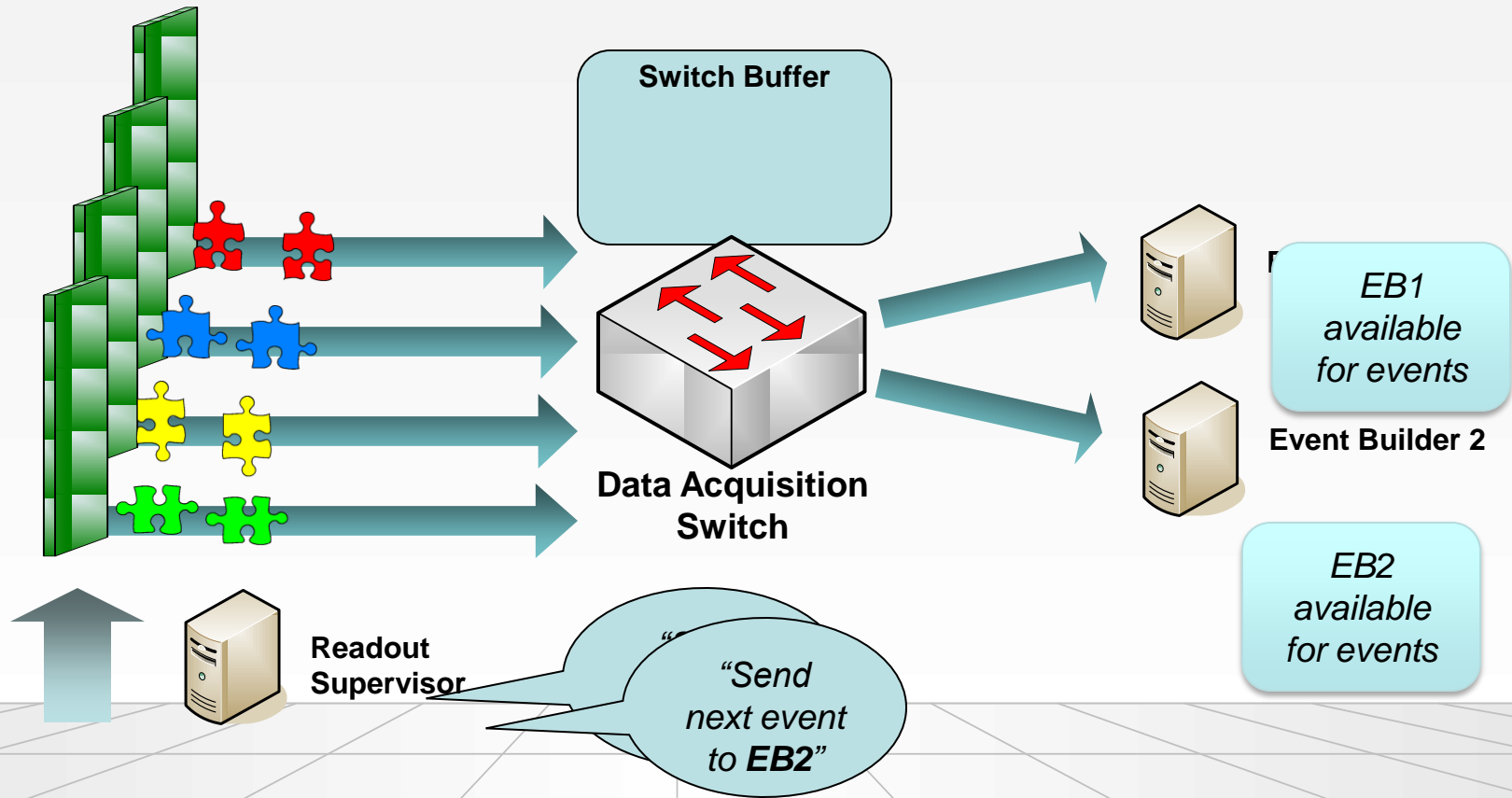


MEP



- Event fragments from up to 16 triggers packed into one Multi-Eventfragment Packet (MEP)
- A MEP must fit into one IPv4 packet (**64 kB max!**)
- IPv4 packets will be fragmented into Ethernet frames of MTU size
- MEP header is 12 bytes only (Event-number + length)
- IP address information is used in event-building
- Unreliable / no duplication / no re-transmission

Dataflow in LHCb



1 Event Builders inform Readout supervisor about availability

2 Readout Supervisor tells readout boards where events must be sent (round-robin)

3 Readout boards do not buffer, so switch must

So we can do 1 MHz

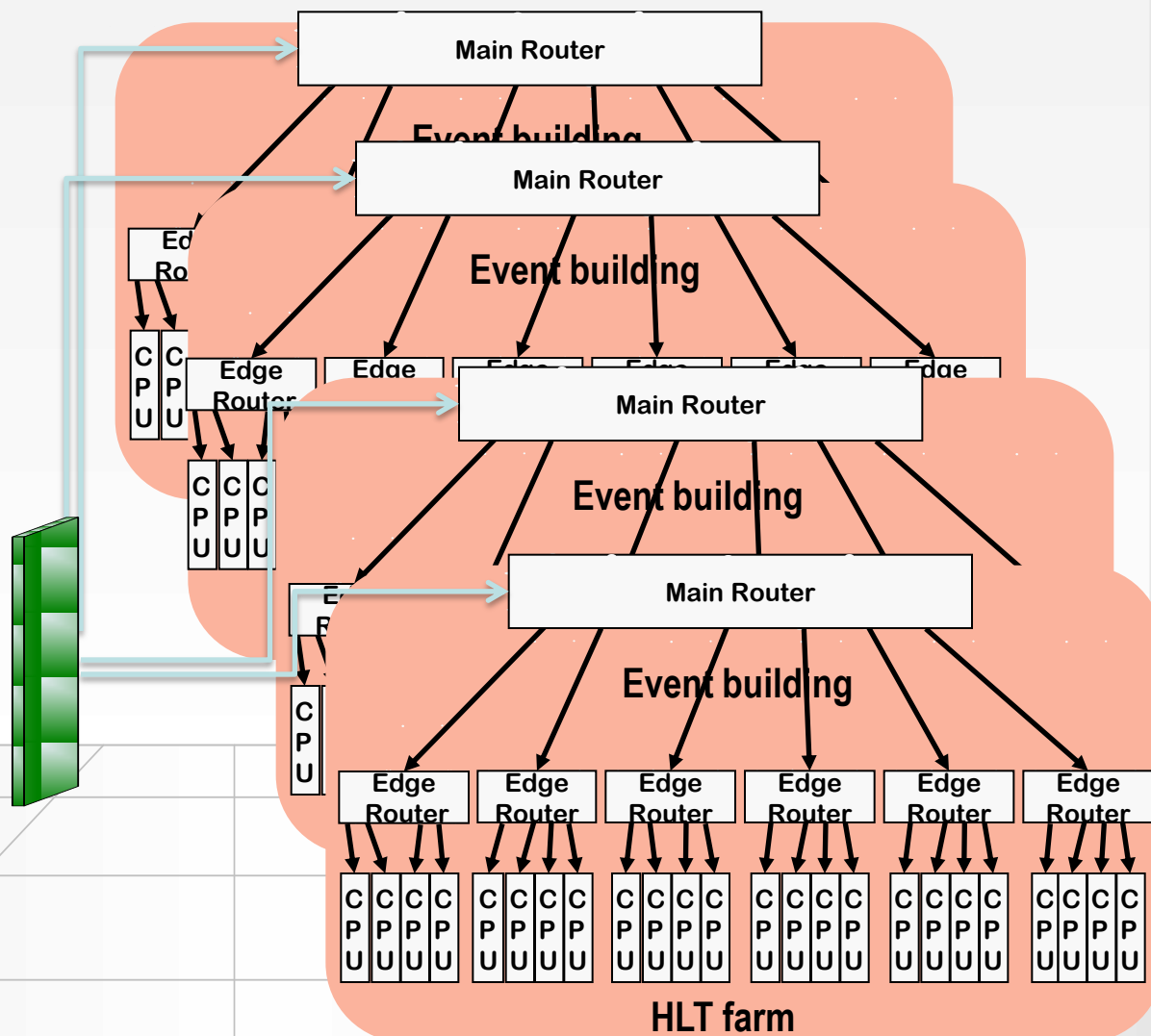
Or can we 😊?



How Can we do 40 MHz?

$$40 = 10 \times 4$$

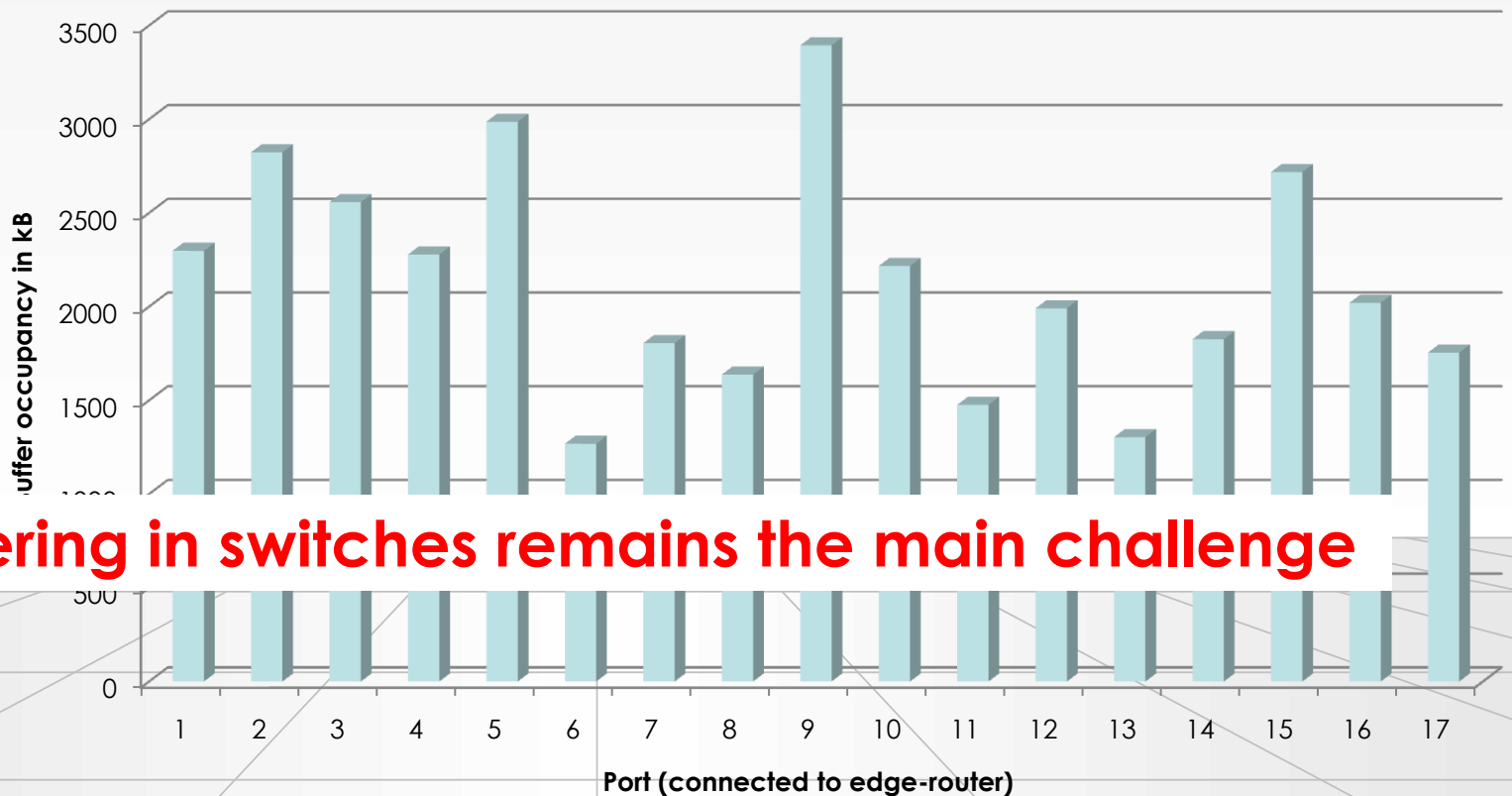
- We get a factor 10 from the link-speed: 1 Gigabit \rightarrow 10 Gigabit throughout the DAQ
- We get a factor 4 by deploying several slices of the DAQ network and farm (a concept taken from the CMS DAQ)



Buffer Usage

Buffer usage in Force10 E1200i with a test using 270 sources @ 350 kHz event-rate

- 256 MB shared between 48 ports
- 17 ports used as "output"
- Non-uniform under investigation



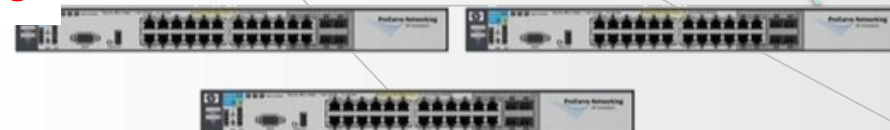
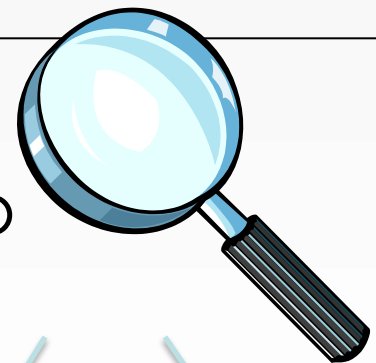
Buffering in switches remains the main challenge

Inside the network bubble

- Currently only 1 device identified which can handle the current LHCb traffic
- There are probably a few more
- These devices are very expensive and it is not clear if they will scale up by a factor 10 within 5 years →
 - try to build network from smaller & cheaper units
 - need to ease buffering requirements

Need buffering in TELL boards

Main Router



Buffering

- Need buffering for at least 4 s
 - corresponds to 4 GByte @ 20 Gbit/s per 10 Gbit/s link
- Any kind of “traffic-shaping” (random backoff, pulling the data requires buffering in the sources
- MEPs must be accessible **randomly** in this buffer (not like in a FIFO)

Reimplementing MEP

- MEP was successful but quite some work was needed to work with a “non-standard” protocol. Special tools, hardwired settings, special requirements on hard- and software
- MEP2 will be more standard and implemented in UDP →
 - IPv4 RFC 791 must be implemented, as well as ARP (RFC 826) and ICMP a.k.a. ping (RFC 1122) finally UDP (RFC 768)
 - MEP size and maximum packing factor will be increased: 2 MB or even infinite if required

Miscellaneous

- ECS
 - It must be possible to take a snapshot of a number of MEPs by latching them into a memory which can be accessed from the ECS
 - ECS must be accessible via a Gigabit Ethernet link which is strictly independent from the data links
 - It must be possible to reset the ECS interface out-of-band
 - It must be possible to do a consistent latch of all counters pertaining to the data-flow in a coherent way so that they can be read by the ECS
 - Write-only registers are not permitted.
- Debugging
 - A data-generator must be provided which works at wire-speed for all output links and is completely independent of the input (front-end).