Lecture 2

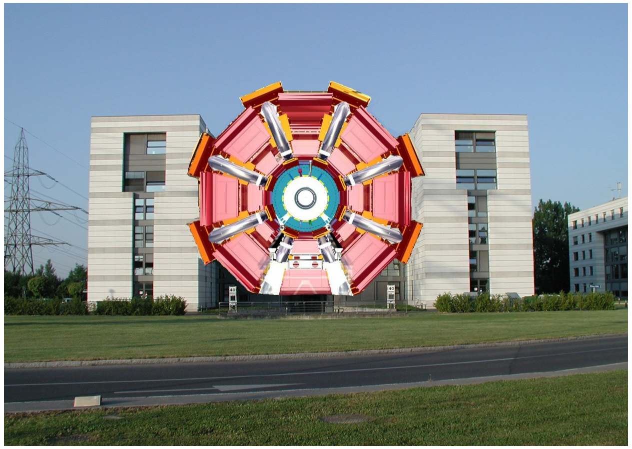
Monika Wielers

Rutherford Appleton Laboratory

- Trigger and Data Acquisition requirements for LHC
- Example: Data flow in ATLAS (transport of event information from collision to mass storage)

What are the challenges at LHC for DAQ?

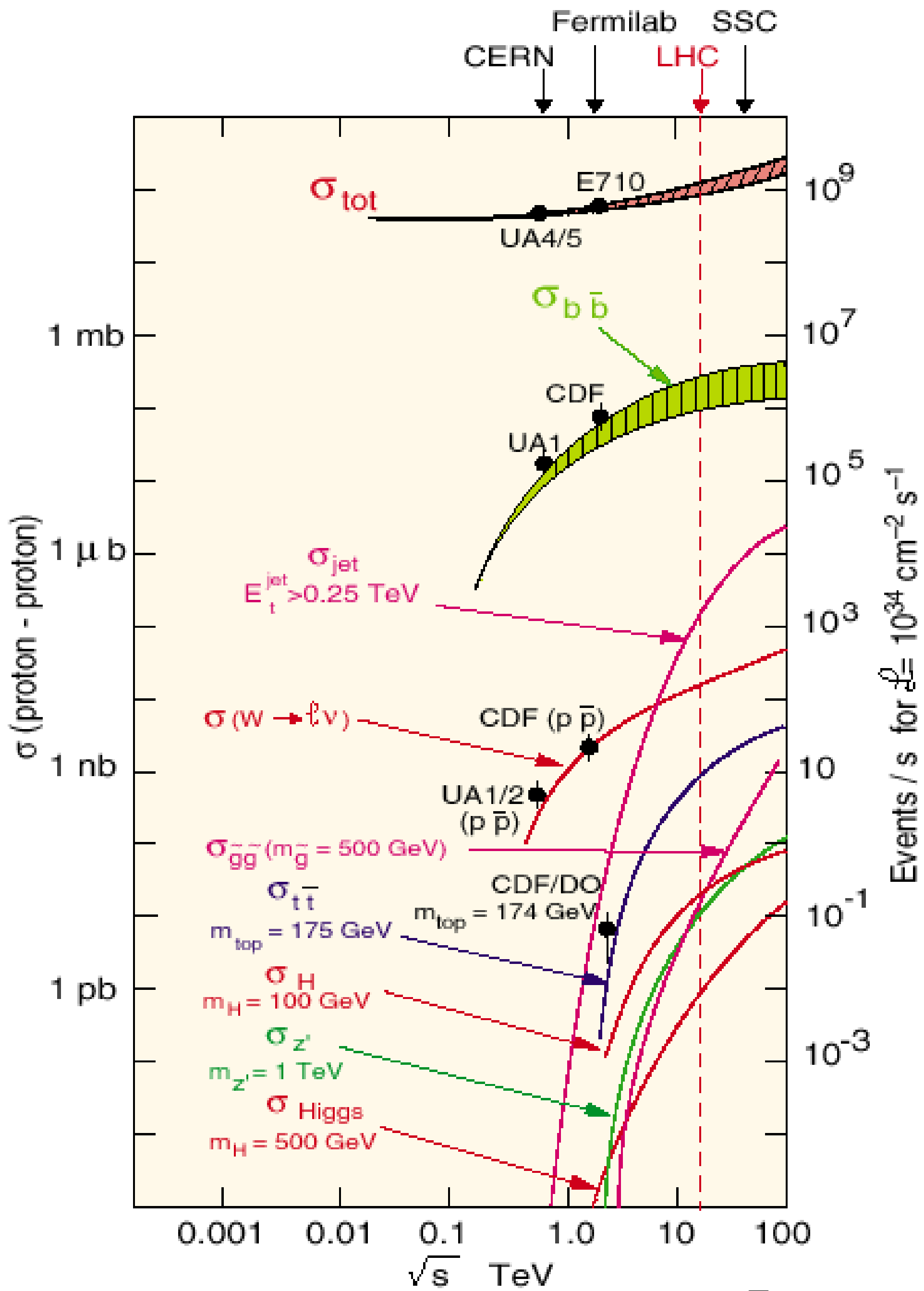
- Challenge 1
  - Physics – Rejection power
  - Requirements for TDAQ driven by rejection power required for the search of rare events
- Challenge 2
  - Accelerator – Bunch crossing frequency
  - Highest energy and luminosity needed for the production of rare events in wide mass range
- Challenge 3
  - Detector – Size and data volume
    - Unprecedented data volumes from a huge and complex detectors



Challenge 1: Physics

- Cross sections for most processes at the LHC span ~10 orders of magnitude
- LHC is a factory for almost everything: t, b, W, Z…
- But: some important signatures have small branching ratios (e.g. $H \rightarrow \gamma\gamma$, BR ~$10^{-3}$)

Process	Production Rate 10^{34} $cm^{-2}s^{-1}$
inelastic	~1 GHz
bbbar	5 MHz
$W \rightarrow l\nu$	150 Hz
$Z \rightarrow l\nu$	15 Hz
ttbar	10 Hz
Z'	0.5 Hz
H(120) SM	0.4 Hz



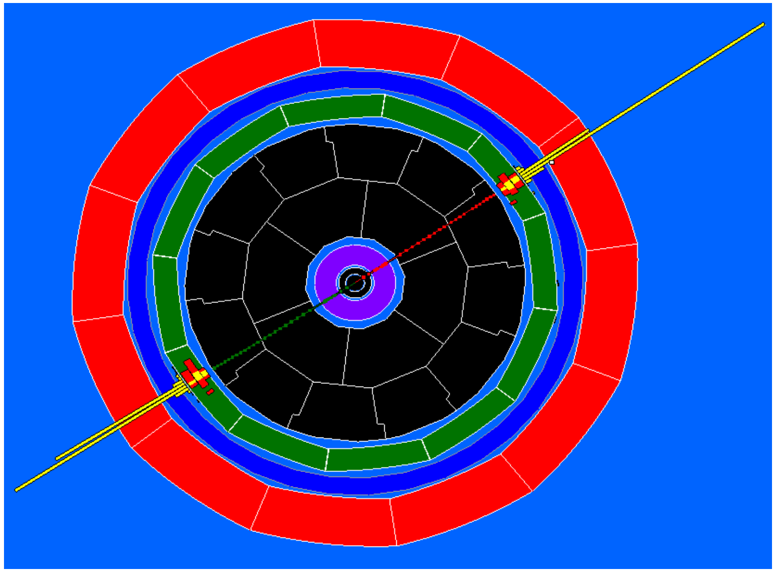
Challenge 1: Physics

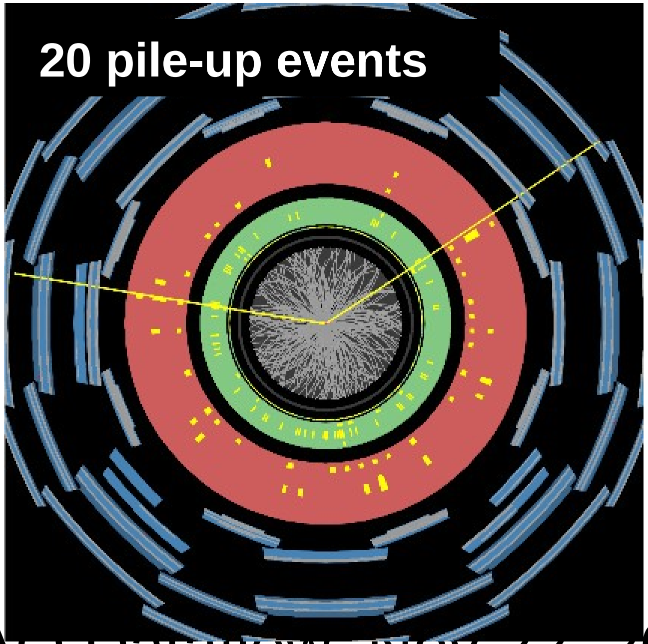
- Requirements for TDAQ driven by rejection power required for the search of rare events
- Besides the Higgs searches one of the motivations for the LHC are new particles outside the SM
  - Susy, extra-dimensions, new gauge bosons, compositeness, black holes etc.
- Be prepared for the 'new unknown', thus ensure you don't reject what is out there by your trigger
- Trigger needs to be flexible and scalable
  - Large luminosity range over lifetime of the experiments
  - "Tunable" for new physics seen

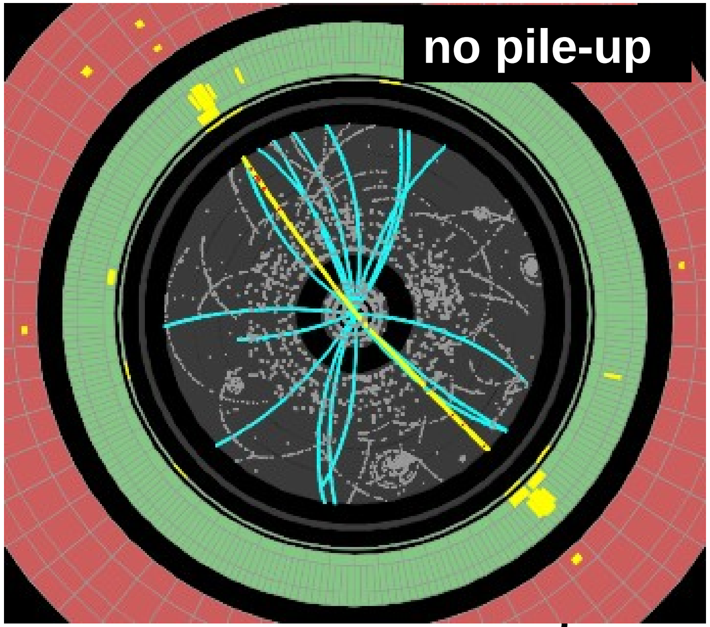
- All of this must fit in around 300 Hz of data writing to mass storage for physics analyses

Challenge 2: Accelerator

- Unlike e^+e^- colliders, proton colliders are more 'messy' due to proton remnants
- Bunch crossing frequency of 40 MHz
- LHC produces ~25 overlapping p-p interactions every 25 ns at design luminosity (in 2011 we had already up to ~20 'pile-up' events with 50ns bunch spacing)



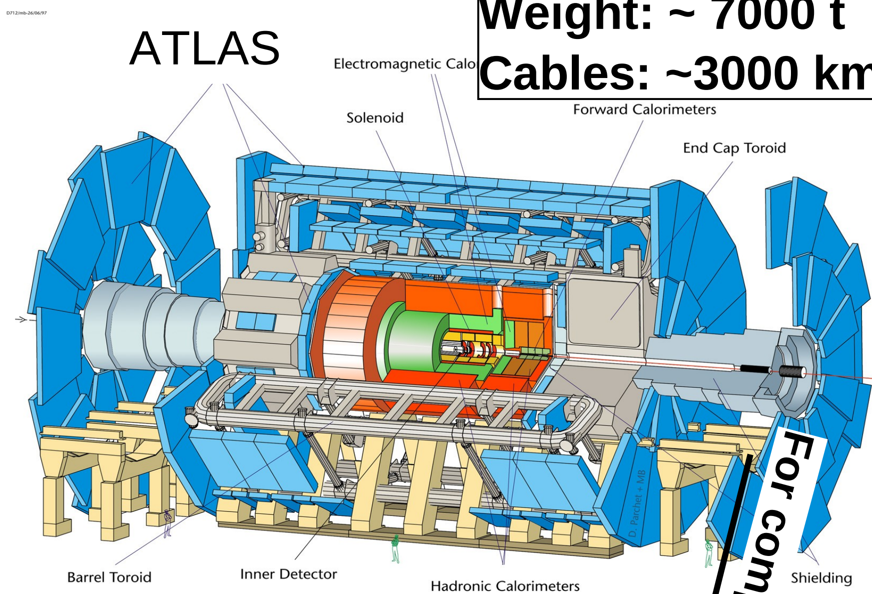


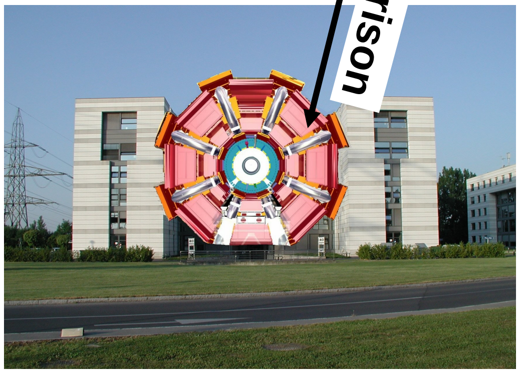


Challenge 3: Detector

- Besides being huge: number of channels are O(10^8) in ATLAS, O(10^7) in CMS, event size ~1.5 mB
  - need huge number of connections
- At 10^{34} $cm^{-2}s^{-1}$ every 25ns LHC flushes detector with ~1400 particles
- Some detectors need > 25ns to readout their channels and integrate more than one bunch crossing's worth of information (e.g. LArg readout takes ~400ns)
  - need to identify bunch crossing...
- It's On-Line (cannot go back and recover events)
  - need to monitor selection - need very good control over all conditions

Length: ~40m
Radius: ~10m
Weight: ~ 7000 t
Cables: ~3000 km



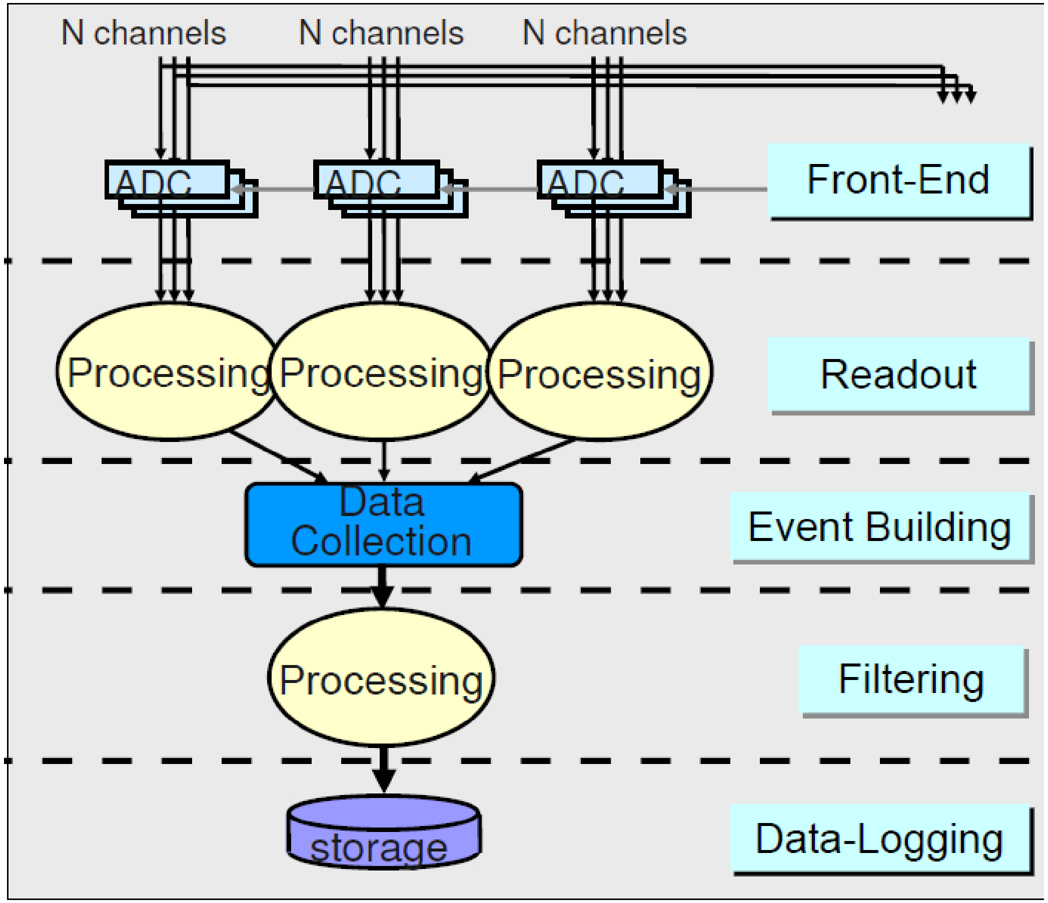


Let’s build a Trigger and DAQ for this

- What do we need?

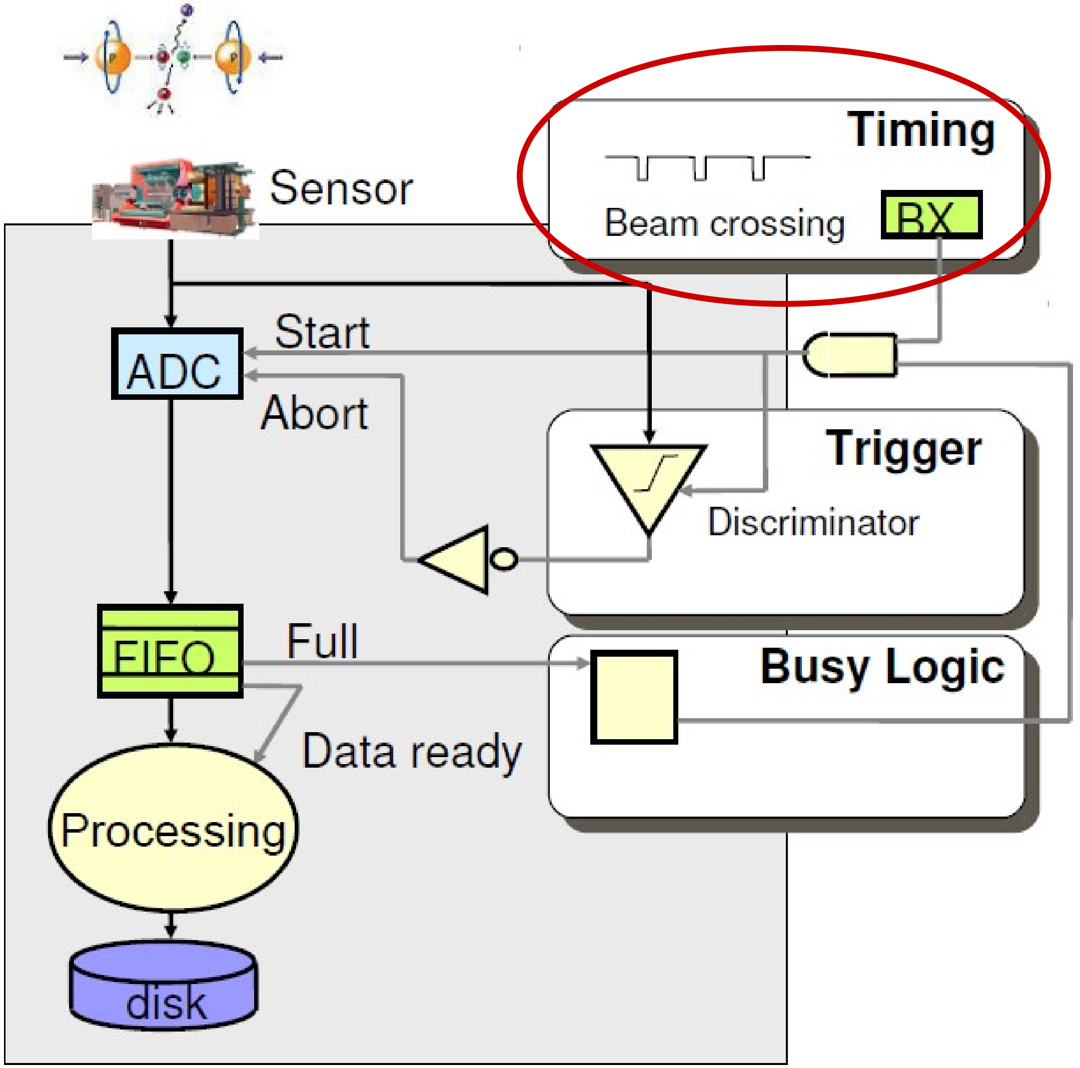
Let's build a Trigger and DAQ for this

- What do we need?
  - Electronic readout of the sensors of the detectors ("front-end electronics")
  - A system to collect the selected data ("DAQ")



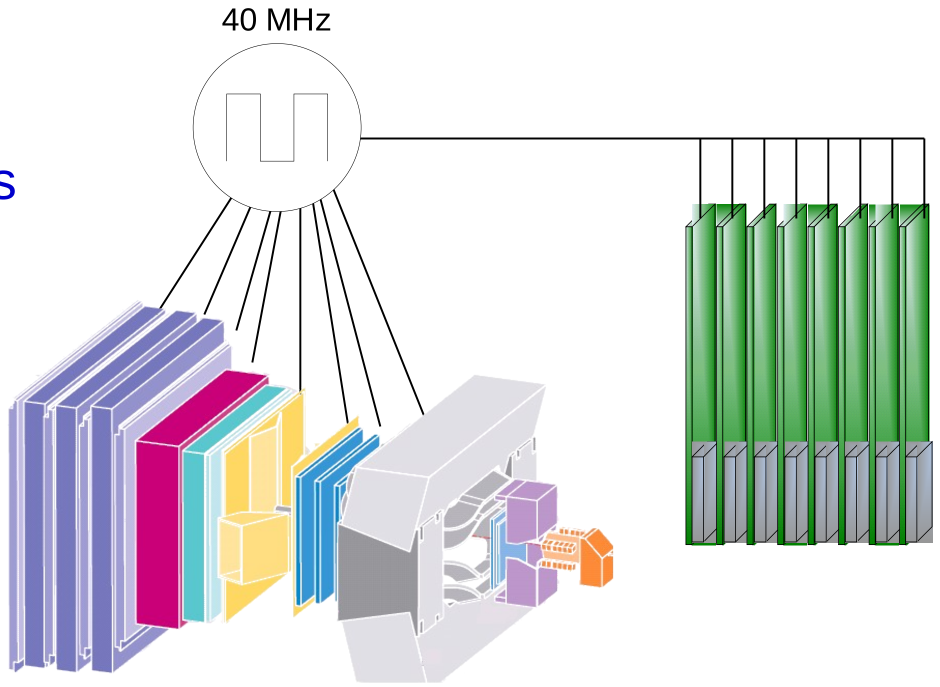
Let's build a Trigger and DAQ for this

- What do we need?
  - Electronic readout of the sensors of the detectors ("front-end electronics")
  - A system to collect the selected data ("DAQ")
  - A system to keep all those things in sync ("clock")



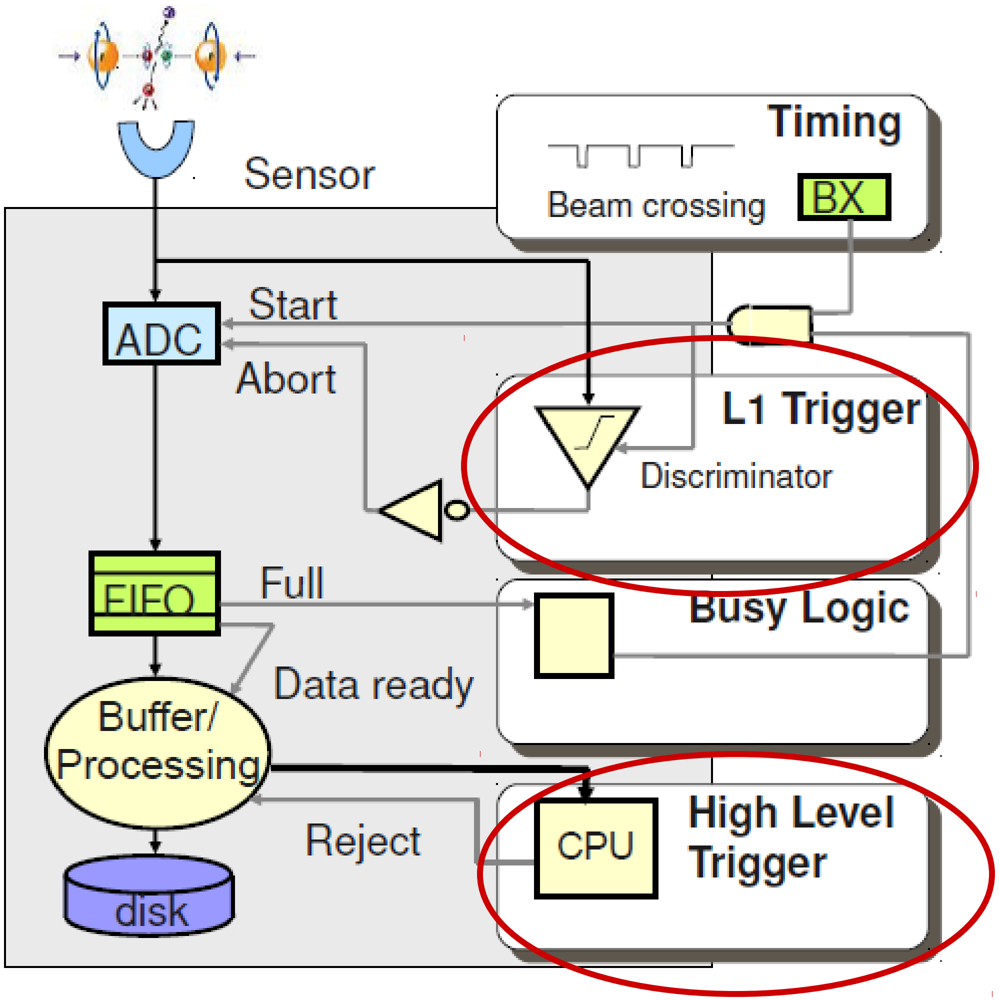
Timing

- An event is a snapshot of the values of all detector front-end readout units caused by the same collision
- A common clock signal must be provided to all detector elements
  - Since clock is a constant, detectors large and electronics fast, the detector elements must be carefully time-aligned
- Common system for all LHC experiments: TTC (Trigger, Timing and Control) based on radiation-hard opto-electronics



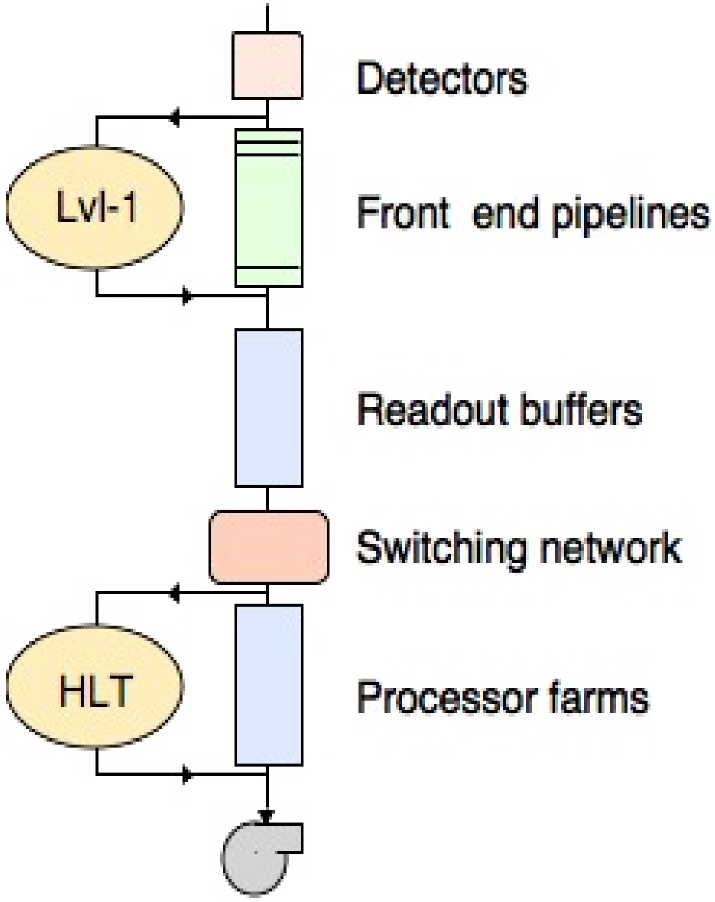
Let's build a Trigger and DAQ for this

- What do we need?
  - Electronic readout of the sensors of the detectors ("front-end electronics")
  - A system to collect the selected data ("DAQ")
  - A system to keep all those things in sync ("clock")
  - A trigger – multi-level due to complexity



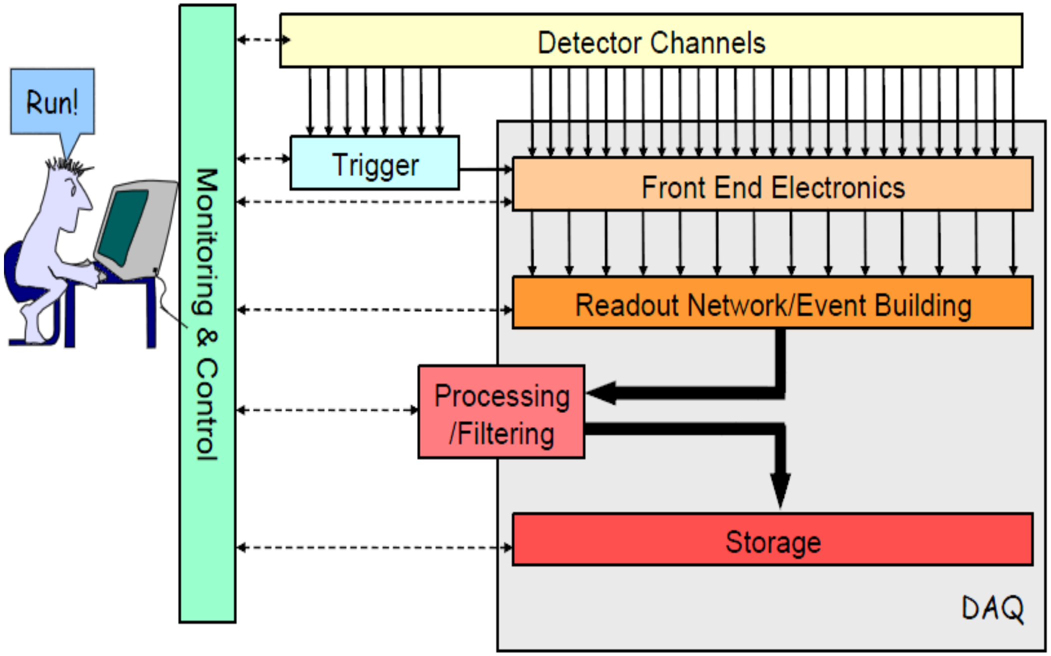
What I need to do for sure: Level-1 trigger

- No (affordable) DAQ system could read out O(10^7-10^8) channels at 40 MHz → 400 TBit/s to read out!
- What's worse: most of these millions of events per second are totally uninteresting: ~1 Higgs event every 0.01-0.001 seconds
- A filter or first level trigger (L1) must somehow* select the more interesting events and tell us which ones to deal with any further

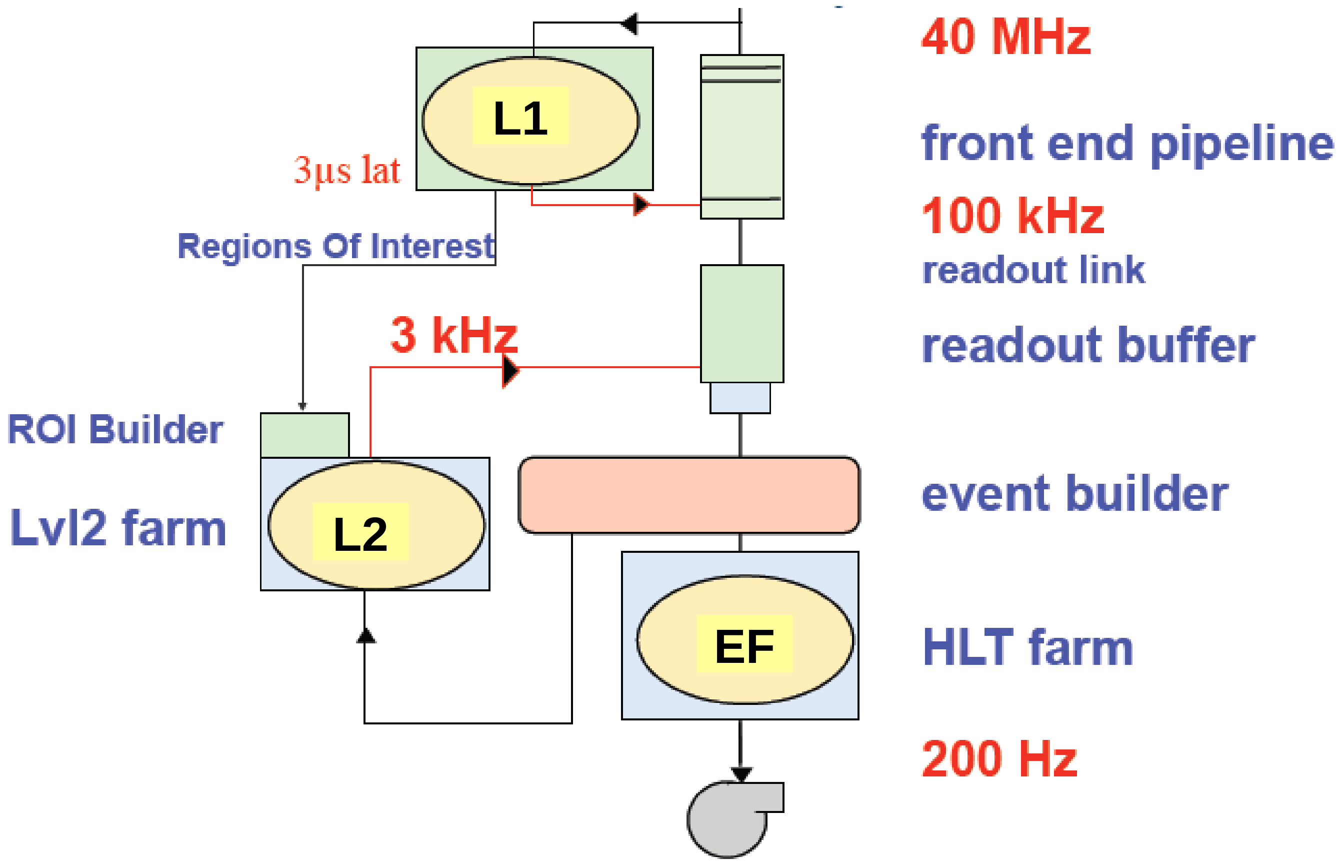


Let's build a Trigger and DAQ for this

- What do we need?
  - Electronic readout of the sensors of the detectors ("front-end electronics")
  - A system to collect the selected data ("DAQ")
  - A system to keep all those things in sync ("clock")
  - A trigger – multi-level due to complexity
  - A Control System to configure, control and monitor the entire DAQ

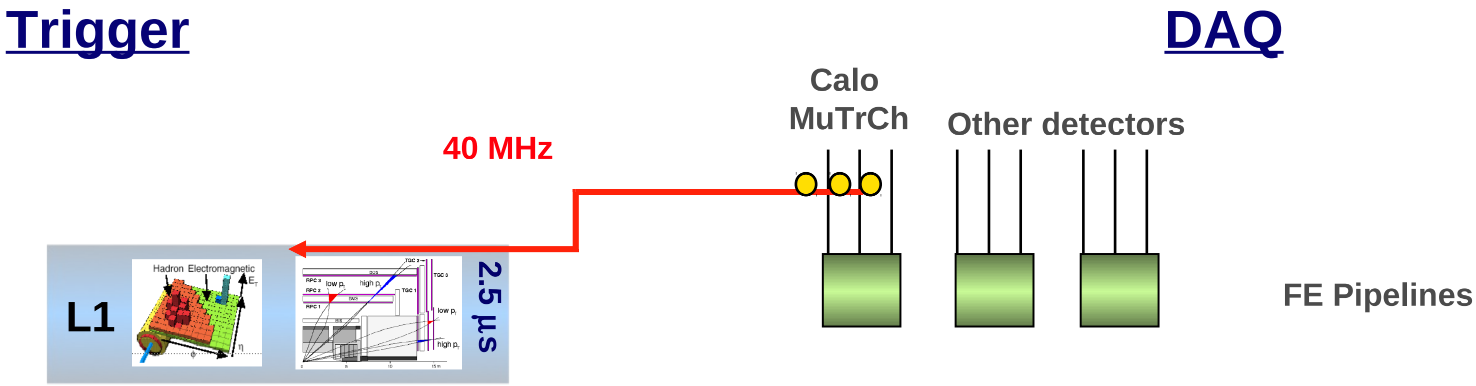


ATLAS Trigger / DAQ Data Flow



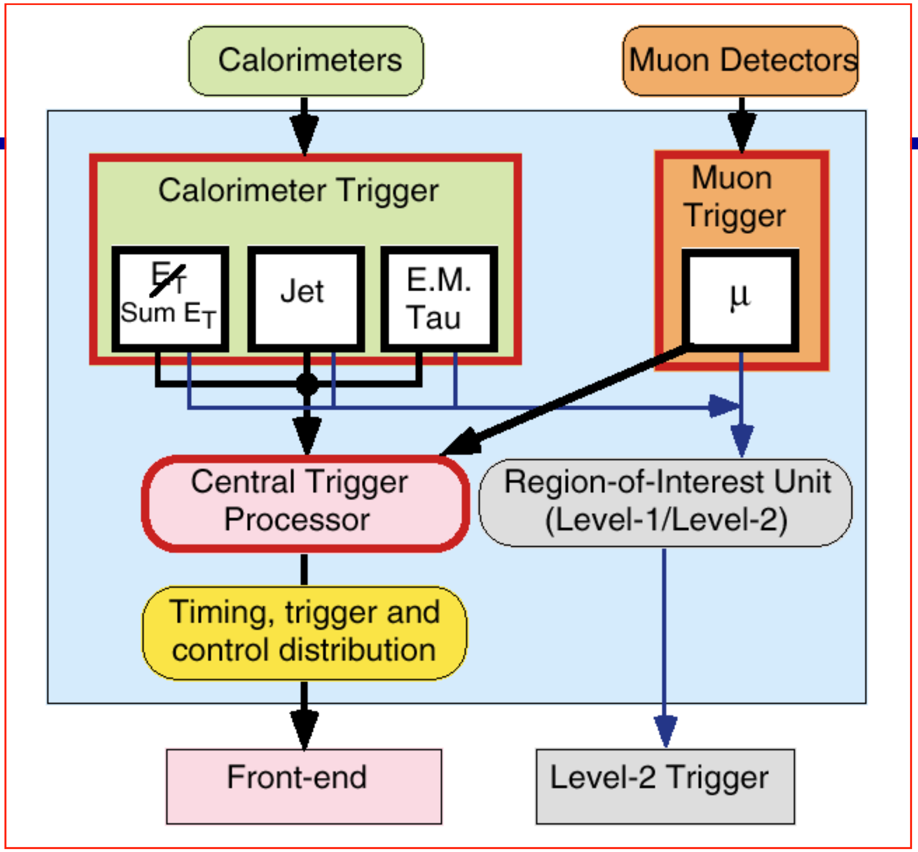
- 3-level trigger hierarchy: L1 – L2 – EF (Event Filter)
- 6 steps
  - L1 trigger
  - L1 decision
  - Readout
  - L2 Trigger
  - Event Builder
  - EF trigger
  - Final storage

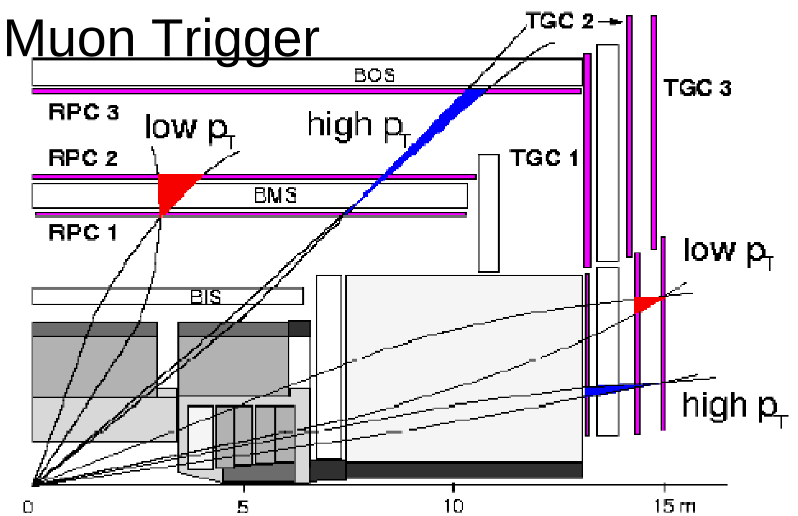
ATLAS Architecture: L1 Trigger



L1 Trigger

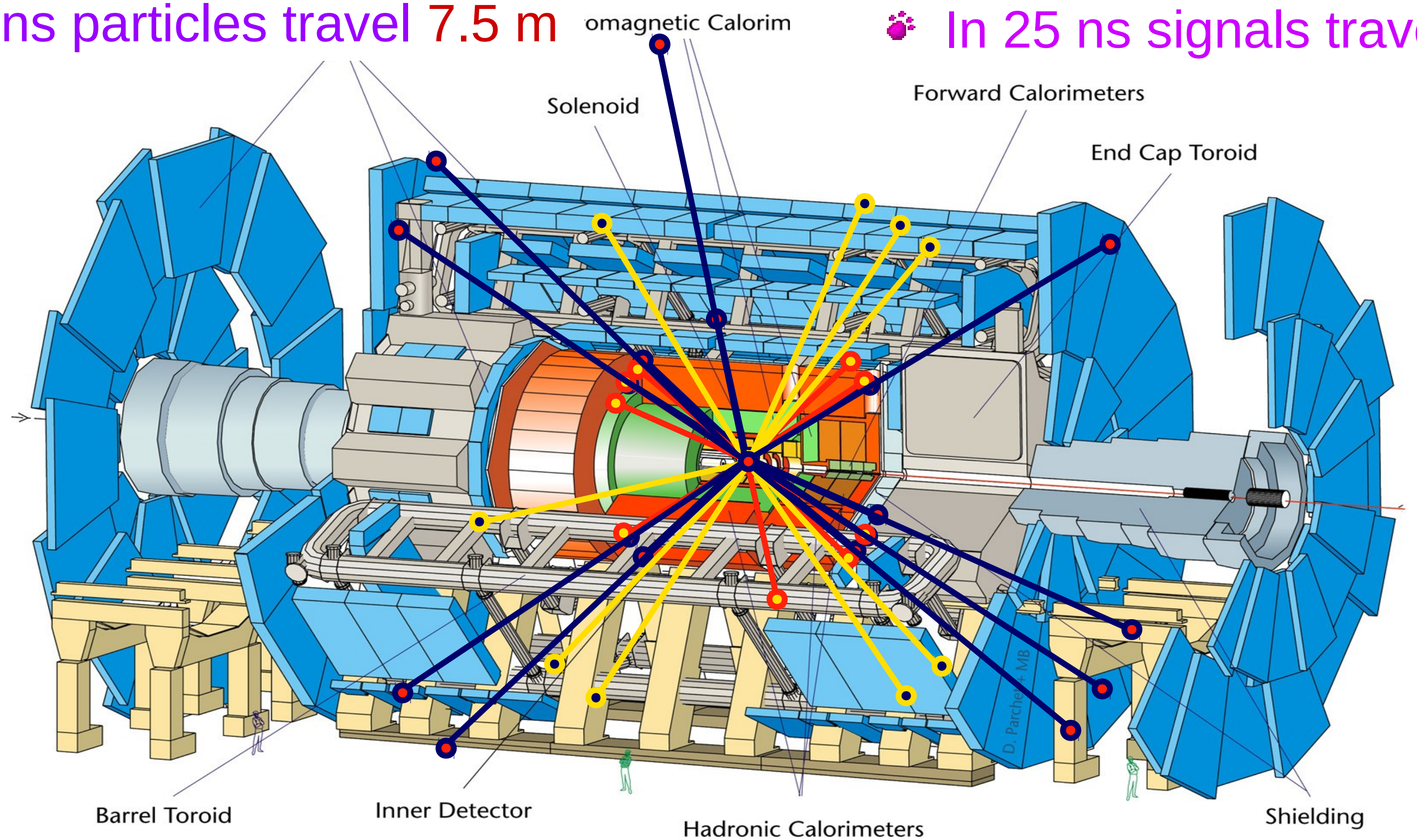
- Calorimeter and muons only
  - muons;
  - em/tau/jet calo clusters; missing E_T, sum $E_{T,}$ missing E_T significance
- Simple algorithms on reduced data granularity
- Also need bunch crossing ID
- Hardware trigger in
  - FPGA (Field-programmable gate array) and ASIC (Application Specific Integrated Circuit)
  - Programmable thresholds
  - Selection based on multiplicities and thresholds
- output L1 rate to <75 kHz (upgradable to 100 kHz)





Level-1 trigger latency

- Interactions every 25 ns …
  - In 25 ns particles travel 7.5 m
- Cable length ~100 meters …
  - In 25 ns signals travel 5 m

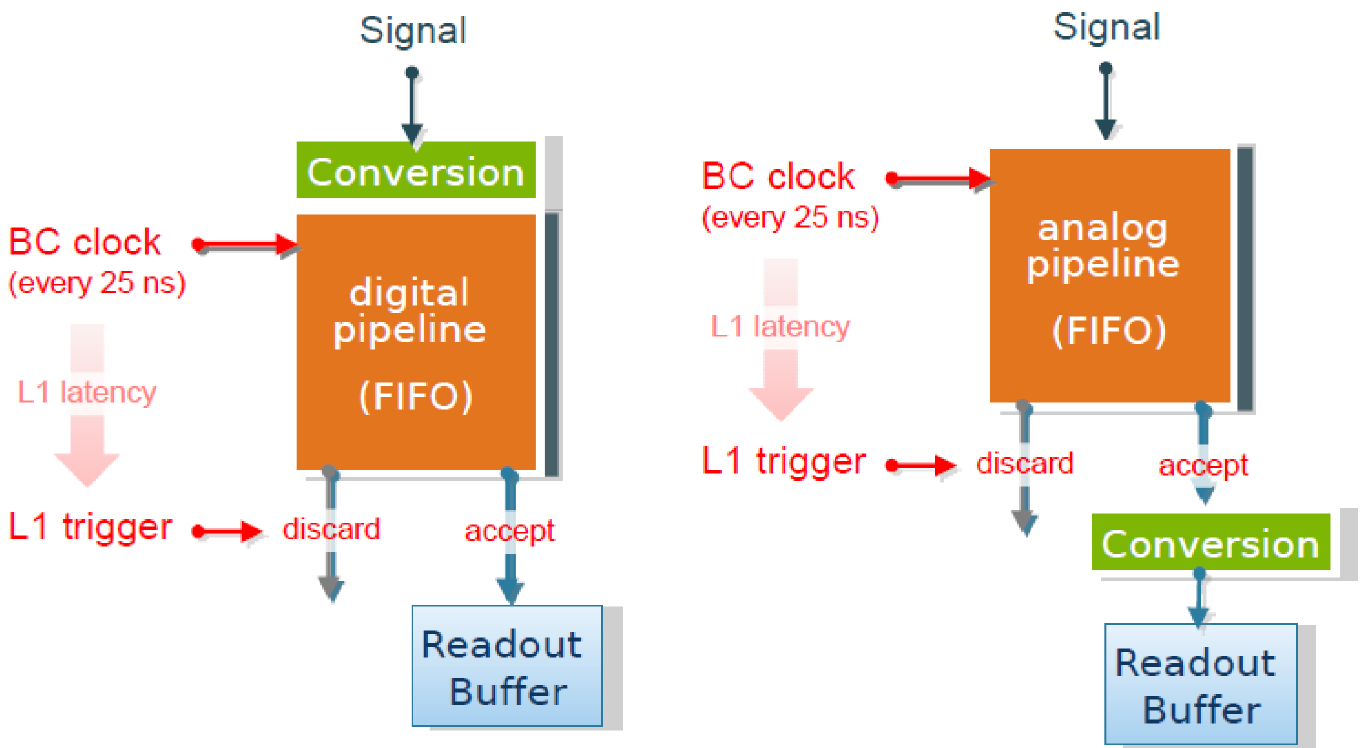


Total L1 trigger latency = *(TOF+ cables+processing+distribution)* = 2.5 μs

For 2.5 μs, all signals must be stored in electronics pipelines

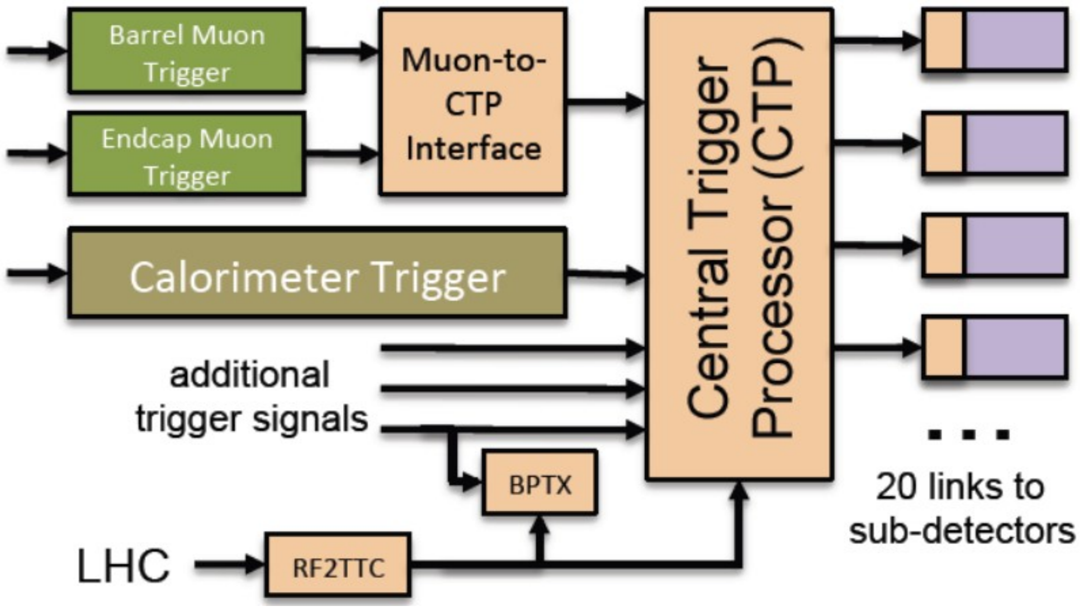
While L1 is doing its job…

- During L1 processing data for all bunch crossings buffered
- Use pipeline in data path for holding data
    - many variations (analog/digital, on/off detector)
- Use pipelined front-ends
- Length of pipeline determines maximum L1 latency



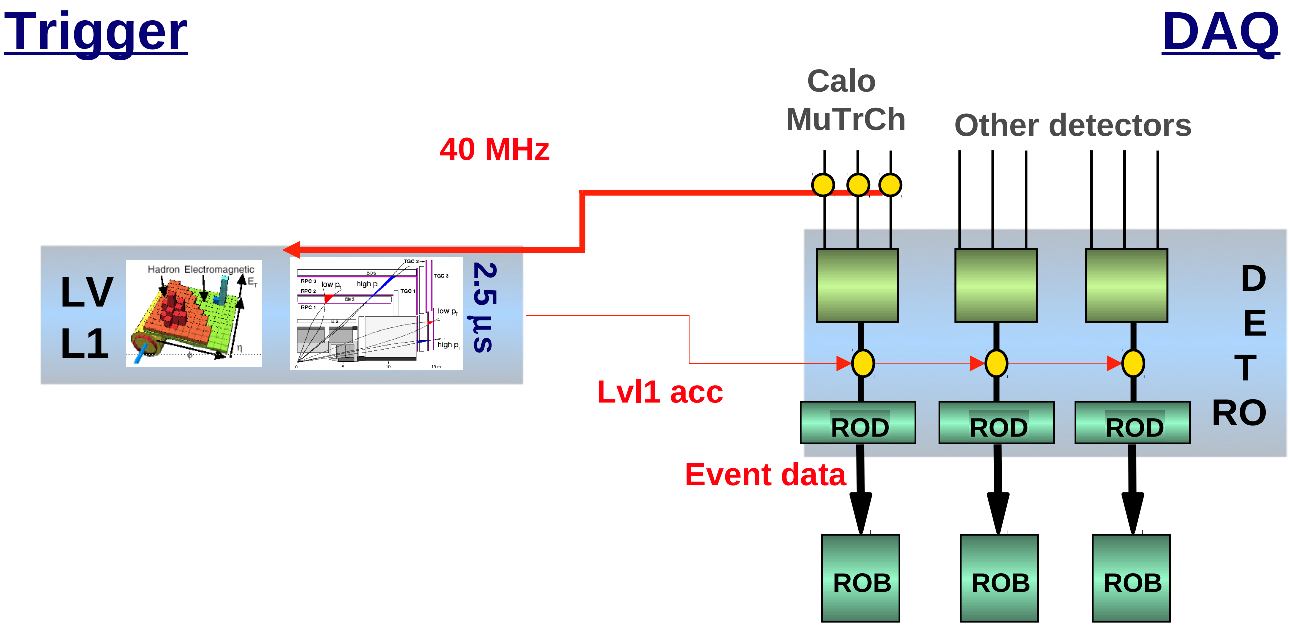
ATLAS Central Trigger Processor (CTP)

- It's here where all information from the muon and calorimeter triggers are collected (via a bus)
- Can combine info e.g. e+jet, 2μ, … to decide if you want to keep your event
  - Can also do prescaling here, e.g. accept only every 1000 event of given type



- If something interesting was found it generates the Level-1 Accept (L1A)
  - The L1A is distributed via the TTC system to the detector front-ends

ATLAS Architecture: Readout Buffer



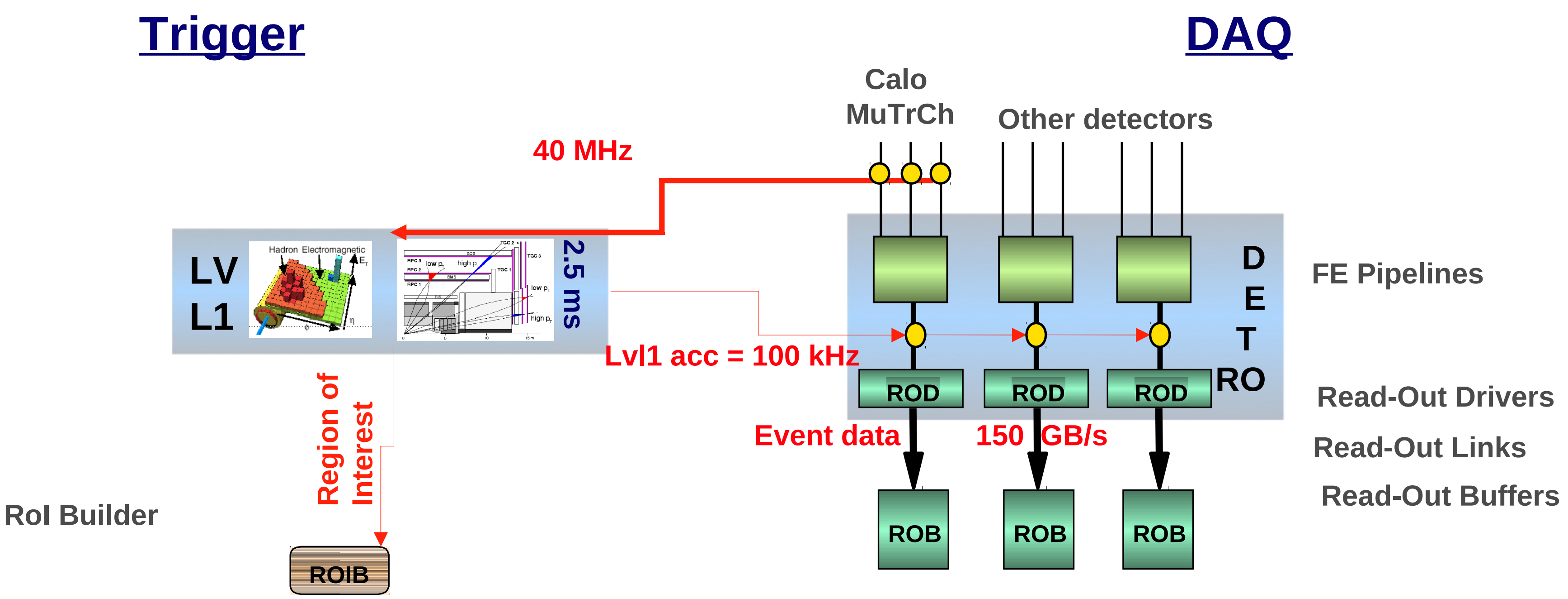
Read-Out Driver

Read-Out Links

Read-Out Buffers

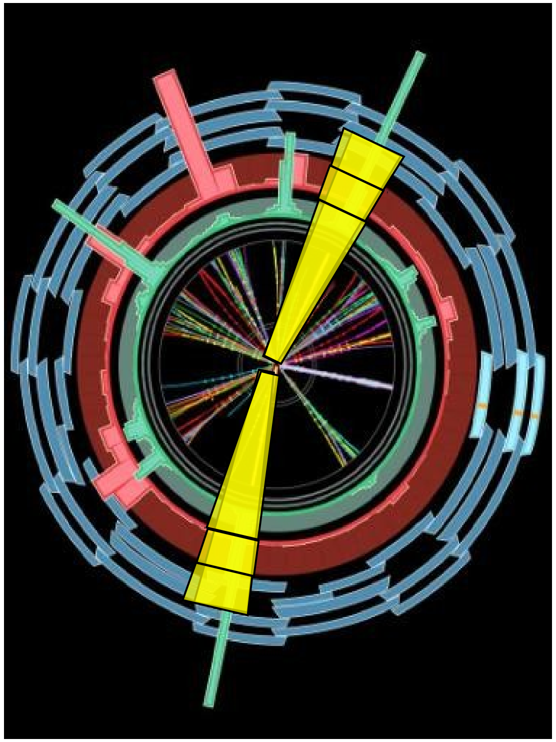
- Upon L1A signal, detector front-ends start sending data of the accepted event to the detector ROD’s (Read-Out Drivers)
- Detector ROD’s receive data, process and reformat it (as needed) and send it via fibre links (Read-out Links (ROL)) to Read-out system (ROS)
- Holds data up to L2 accept/reject
- 160 ROS PCs host ~550 ROBIN cards
  - each ROBIN card has 3 ReadOutBuffers, for a total of ~1600 ROBs
  - Each ROB has one-to-one optical ROL, connection to a ROD

ATLAS Architecture: Region of Interest Builder



Region of Interest (RoI)

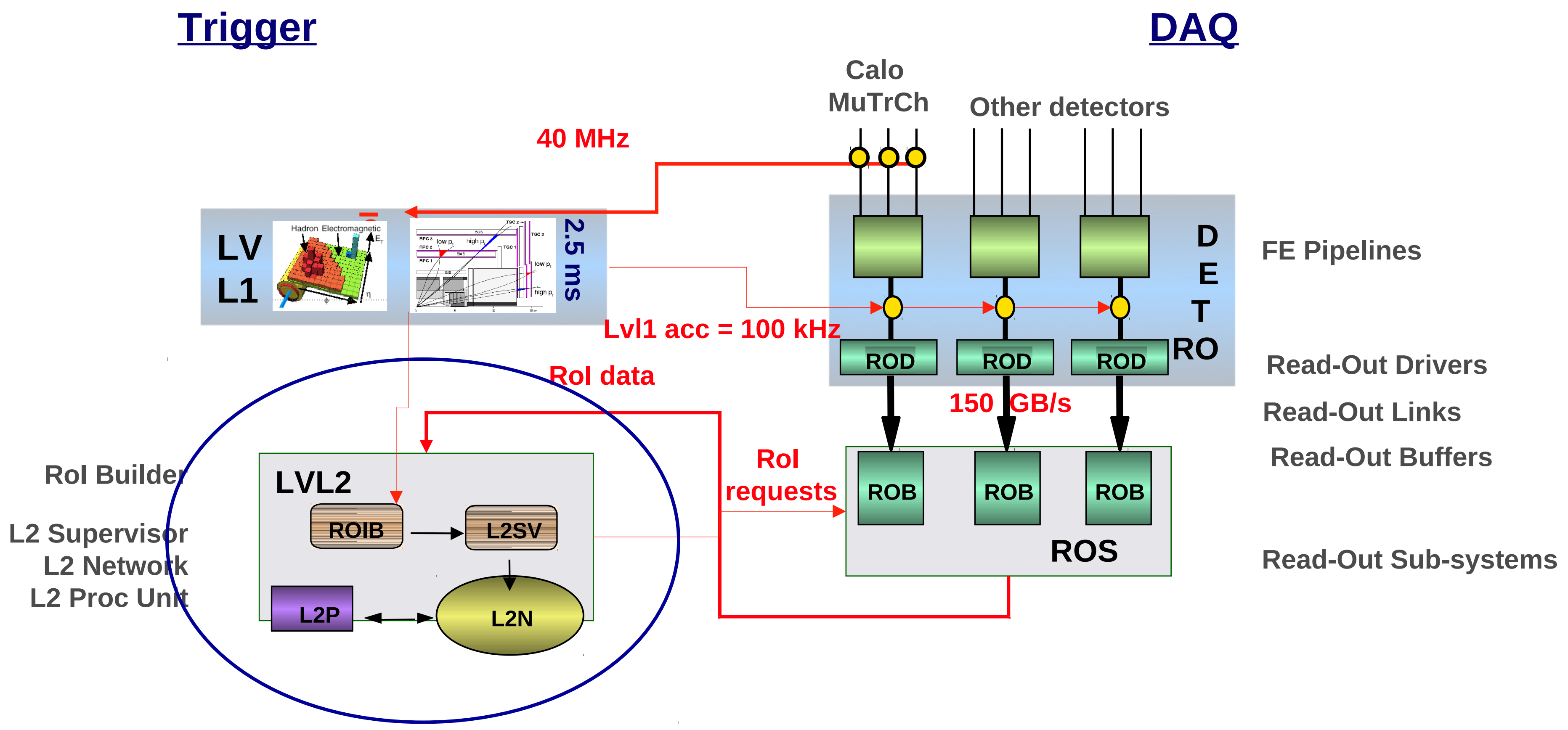
- L1 result contains the (η,φ) coordinates of regions containing high-p_T L1 trigger objects
- There is a simple correspondence $\eta-\phi$ region ↔ ROB number(s) (data fragments containing a certain number of readout units)
  - Identify for each RoI the list of ROBs with the corresponding data from each detector (quick procedure)
- RoIB are VME boards with FPGAs L2 (on average) has to process only 1-4% of the data volume; save on
  - Processing time
  - Bandwidth
- Note: RoI approach only used by ATLAS at LHC



Examples for RoI-based triggers:

- Muons
- Electrons/Photons
- Jets
- Taus

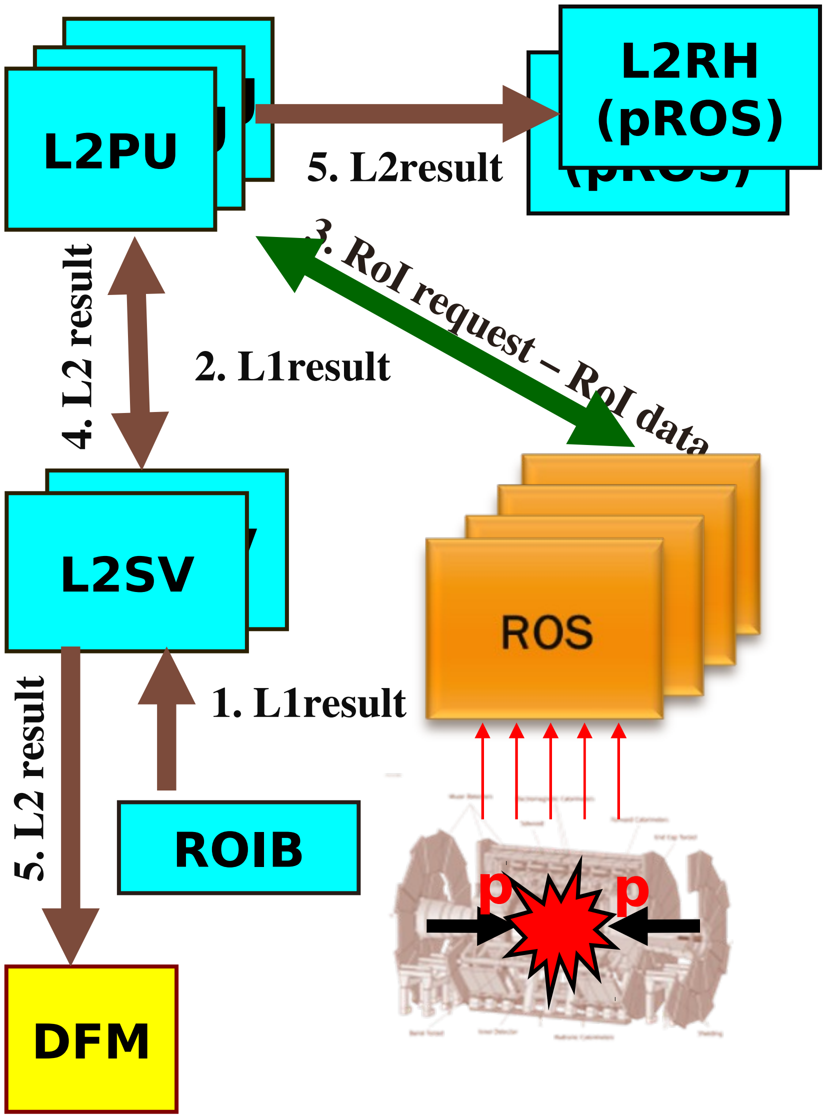
ATLAS Architecture: Level-2 Trigger



L2 Trigger

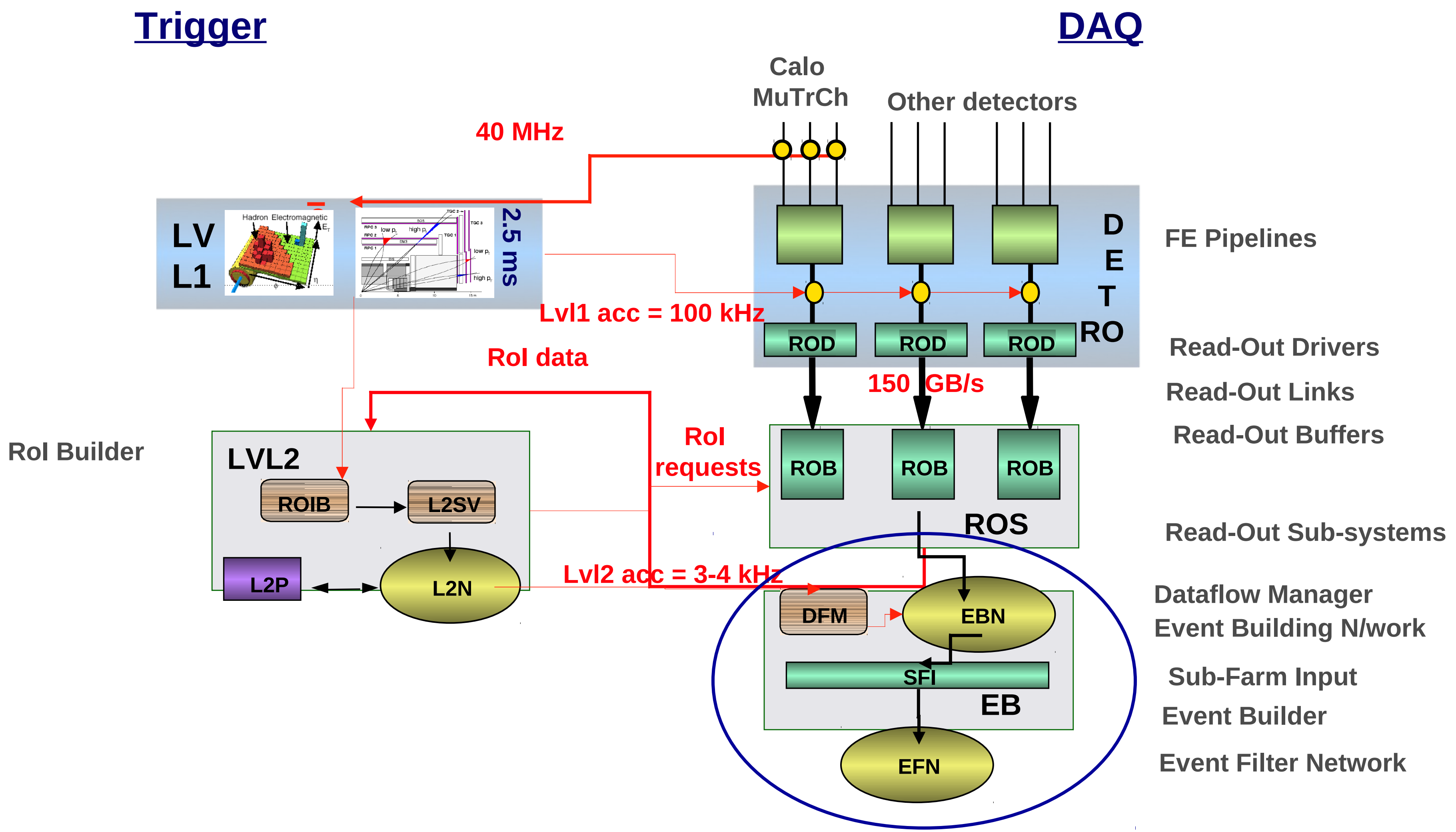
- Software trigger running on a farm of PCs
- Aim
  - overall time budget in L2 : 40 ms average
  - rejection factor : x 30
- Processing scheme
  - Fast selection algorithms depending on input object
  - Identify objects using “simple” criteria
  - combine objects to test event topology

L2 Trigger

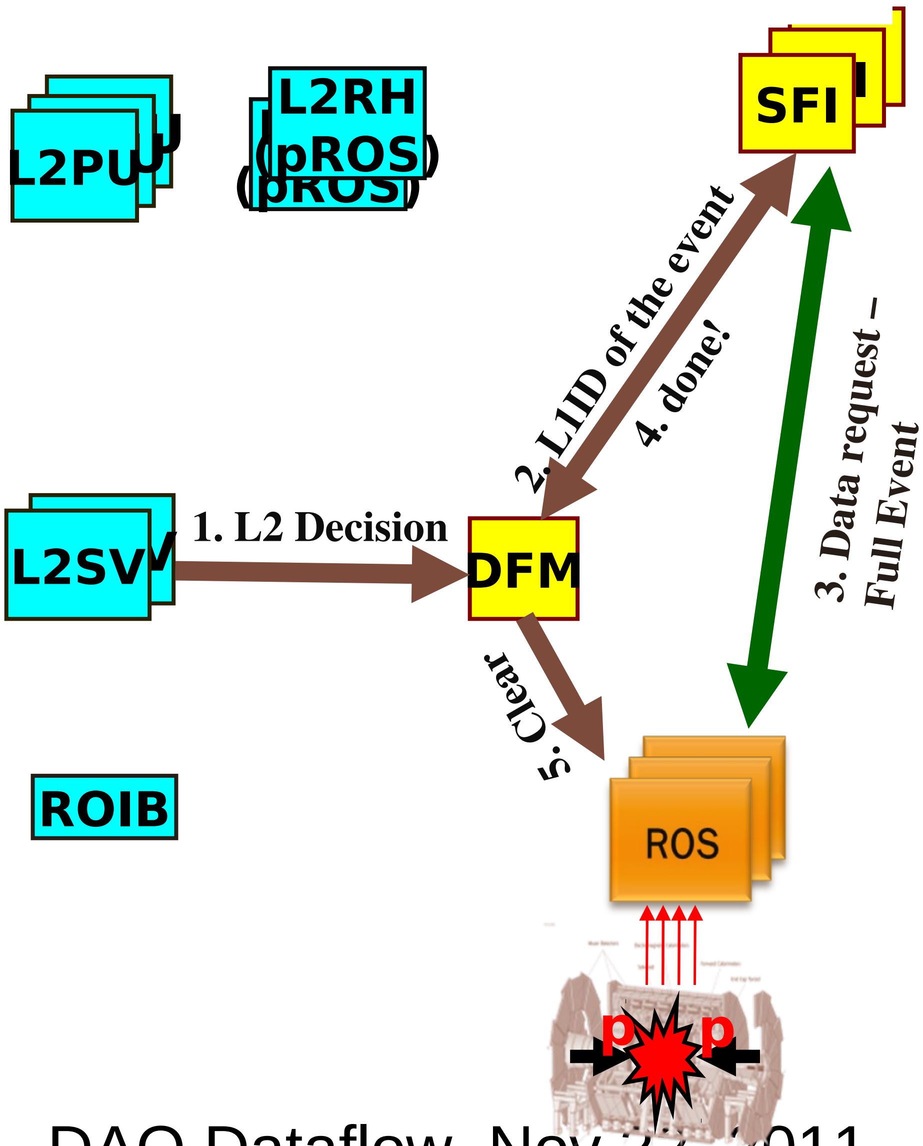


1) Region of Interest Builder (RoIB) passes formatted information to one of the L2 supervisors (L2SV).
2) L2 supervisor selects one of the processors in the L2 farm to process RoI
3) L2 processor (L2PU) requests data from the ROSs
4) L2PU runs selections, produces an accept or reject and informs the L2 supervisor.
5) For an accept result is stored in L2 ResultHandler (L2RH)
6) L2 supervisor passes decision to the DataFlow Manager (controls Event Building).

ATLAS Architecture: Event Builder



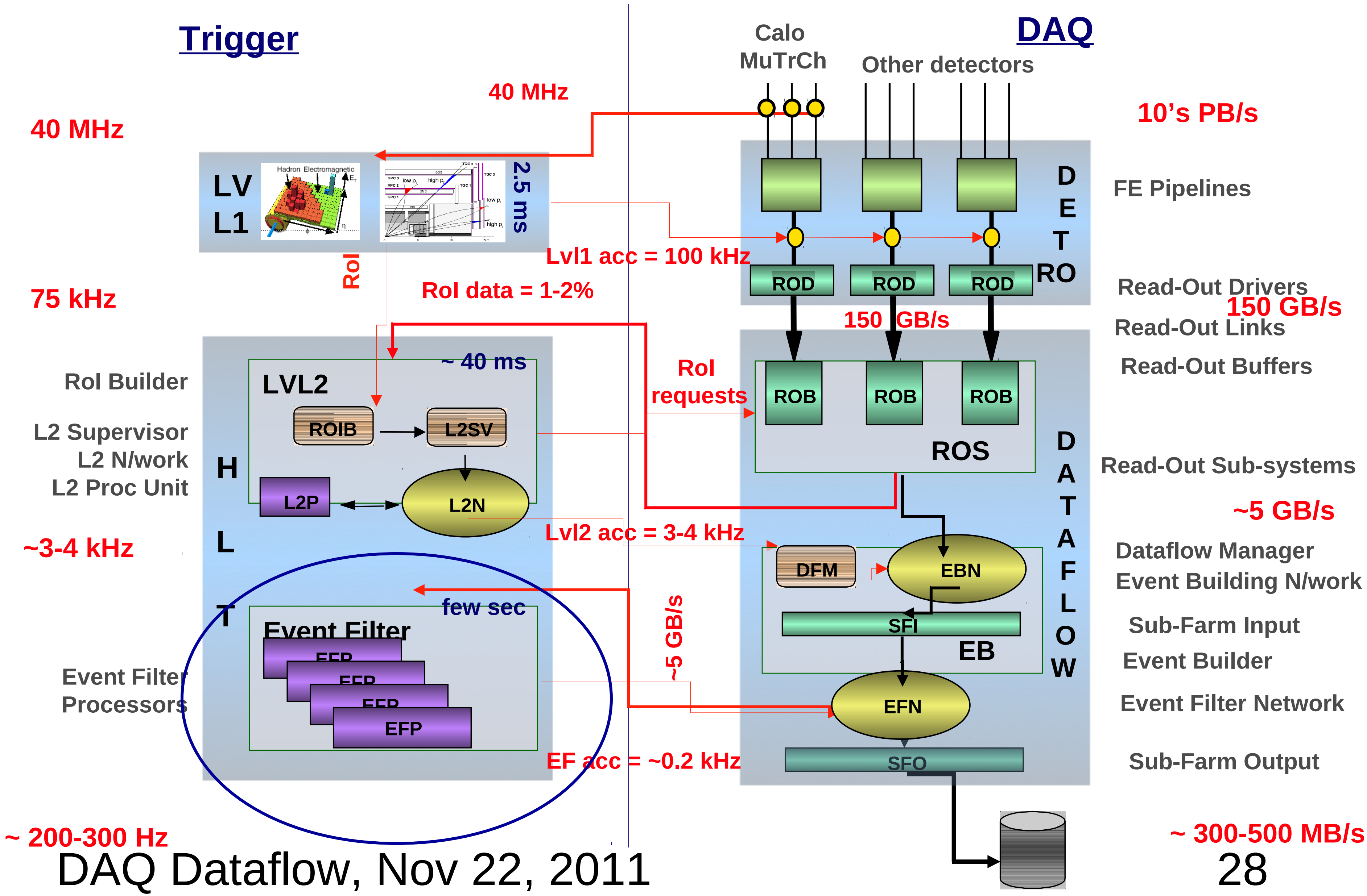
Event Builder



1) L2 Supervisor informs DataFlow Manager (DFM) of event accepted by L2
2) DFM selects a Sub-Farm Input (SFI) and sends to SFI the request to build the complete Event
3) SFI requests ROS's to send event data (L2 pulls event)
4) When done SFI informs DFM.
5) For rejected events and for events for which event Building has completed DFM sends "clears" to the ROSs (for 100 - 300 events together).

- Network traffic for Event Building is ~5 GB/s

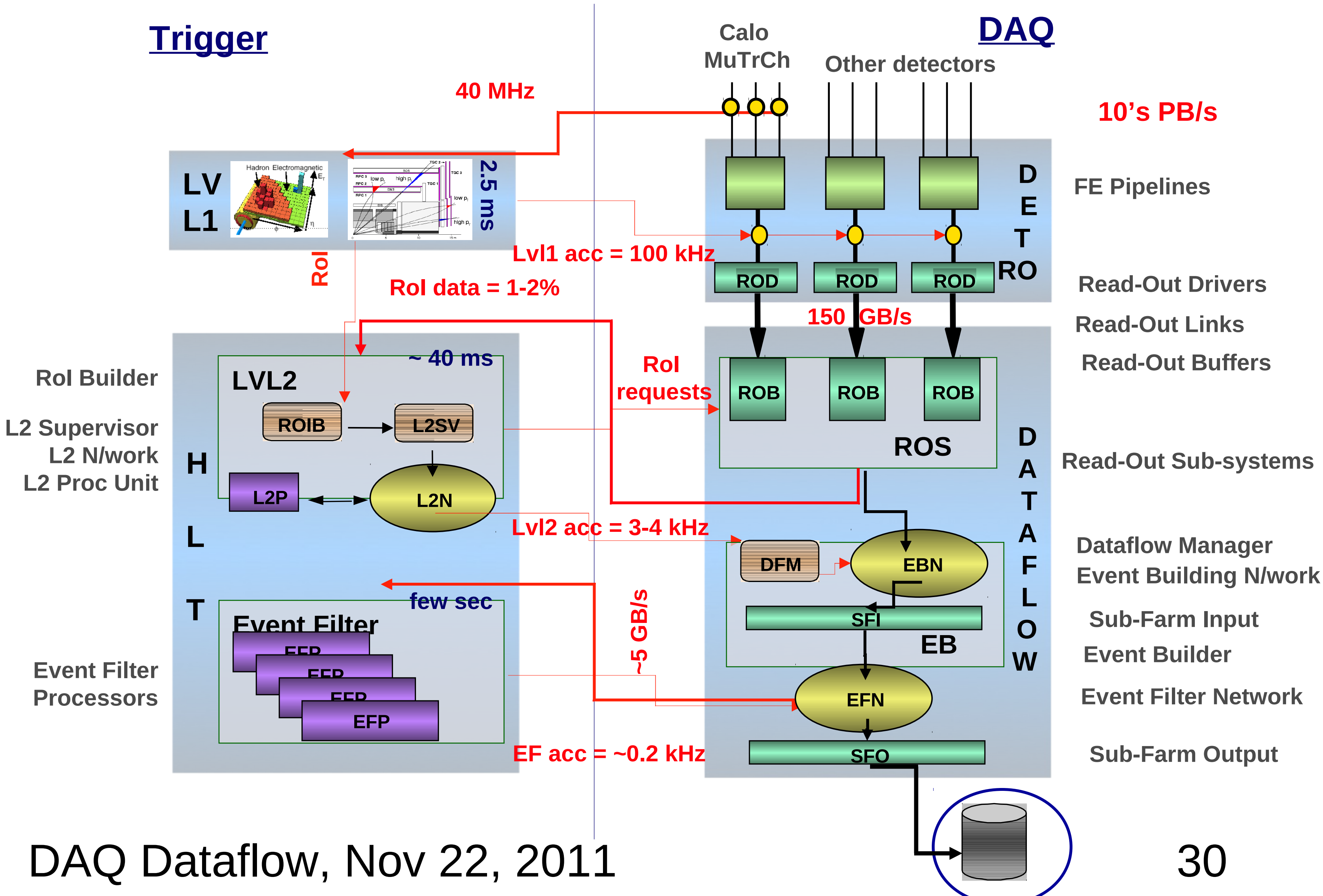
ATLAS Architecture: Event Filter



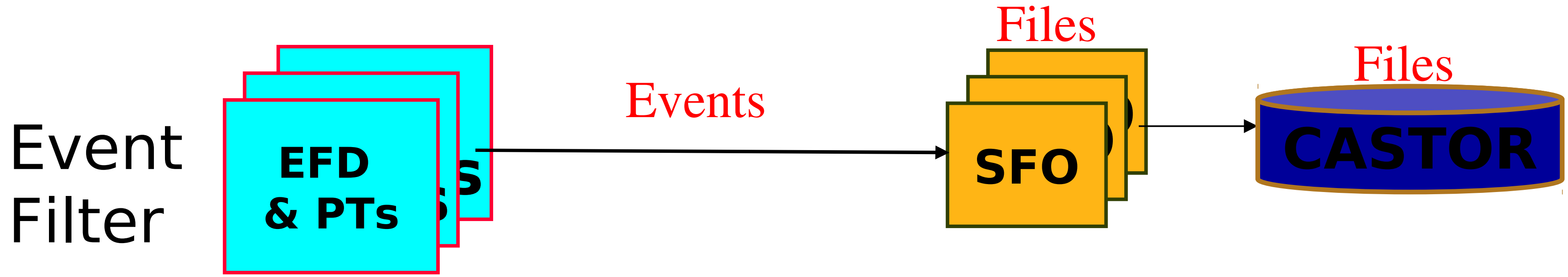
Event Filter

- Final selection in software triggers using large commercial PC farms
  - Latency ~ 4s
  - access to full granularity and offline reconstruction -like algorithms
- Note, there is a flexible boundary between L2 and EF farm

ATLAS Architecture: Storage



Data Logger



- Sub-farm output (SFO)
  - Receive events and write them into files on local disks
  - Dedicated nodes with high performance RAID disks
- Events are sorted out to different files according to their “Physics Stream” content (e.g. “Muons”, “Jets”, “EGamma”, etc.):
  - Events belonging to multiple streams will end up in multiple files
- Files are closed when they reach 2 GB or at end of a luminosity block
  - To assist with overall normalisation each run is sub-divided into periods of a ~ minutes called a luminosity block. During each block the beam luminosity should be constant and can also exclude any blocks where there is a known problem)
- Closed files are finally transmitted via GbE to the CERN Tier-0 for off-line analysis, subsequently erase from local SFO disk

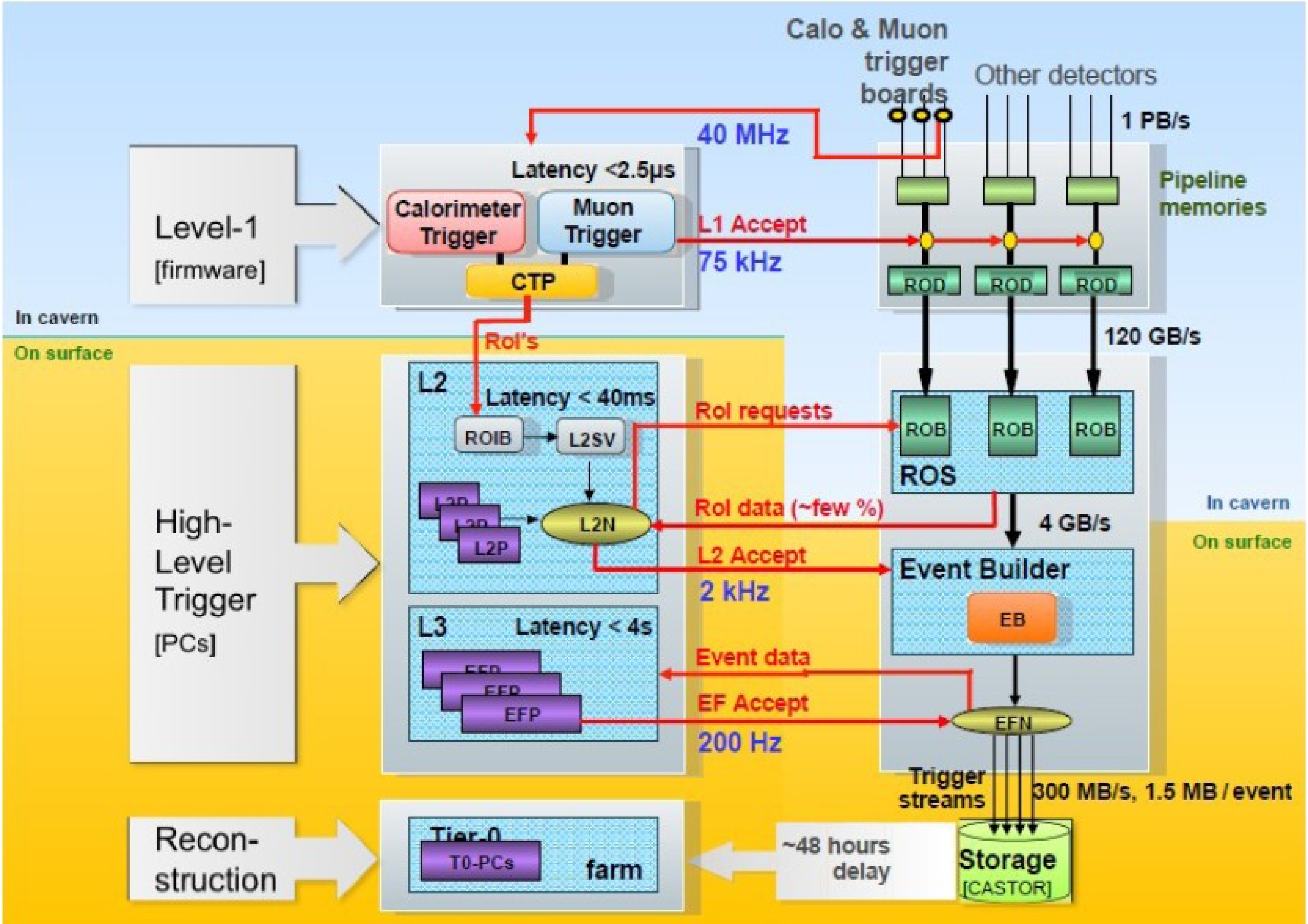
Summary

- Challenge to design efficient trigger/DAQ for LHC
  - Very large collision rates (up to 40 MHz)
  - Very large data volumes (tens of MBytes per collision)
  - Very large rejection factors needed (>10^5)
- Pipelined readouts and fast, parallel custom electronics enable triggers to work at 25 ns collision spacing
- Large networking switches allow high-rate/volume event building
- Large parallel commercial PC farm used to process events with advanced algorithms and high rejections
- Used ATLAS event DataFlow as an example of a large TDAQ system
  - L1 trigger – ReadOut System – L2 trigger – Event Builder – EF – data logger

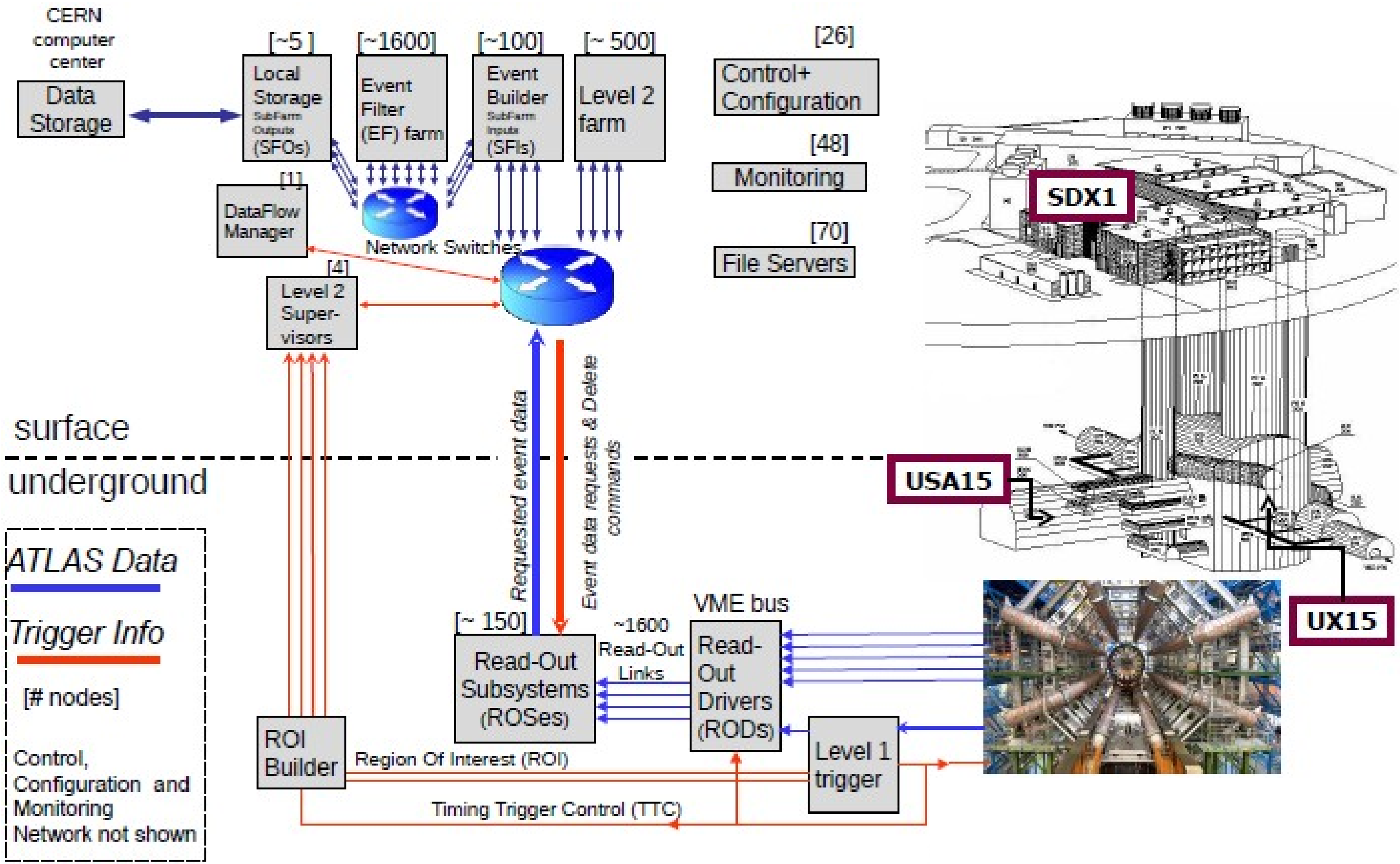
- We'll look in detail at the trigger aspects in the next lecture

Backup

The ATLAS Trigger/DAQ System



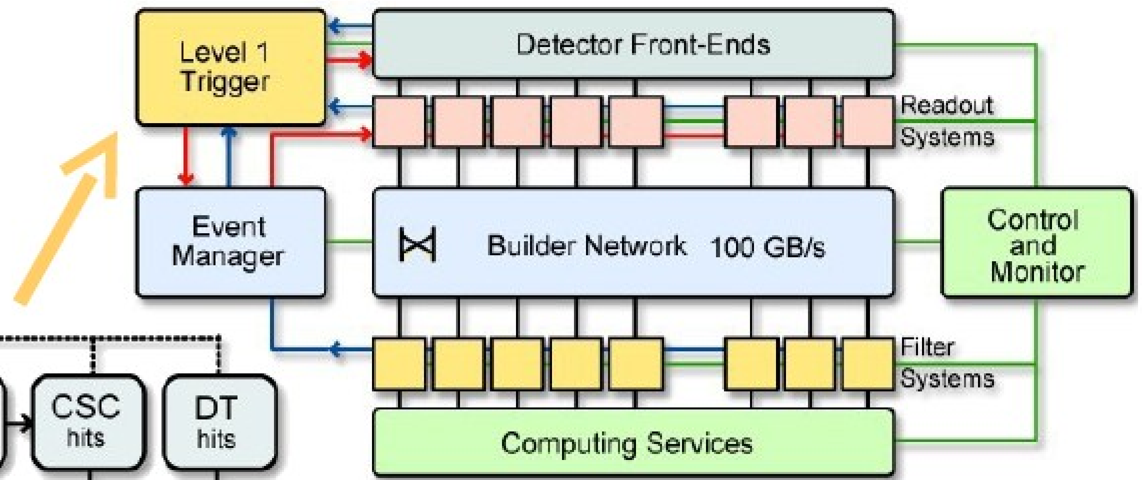
ATLAS Trigger / DAQ Data Flow

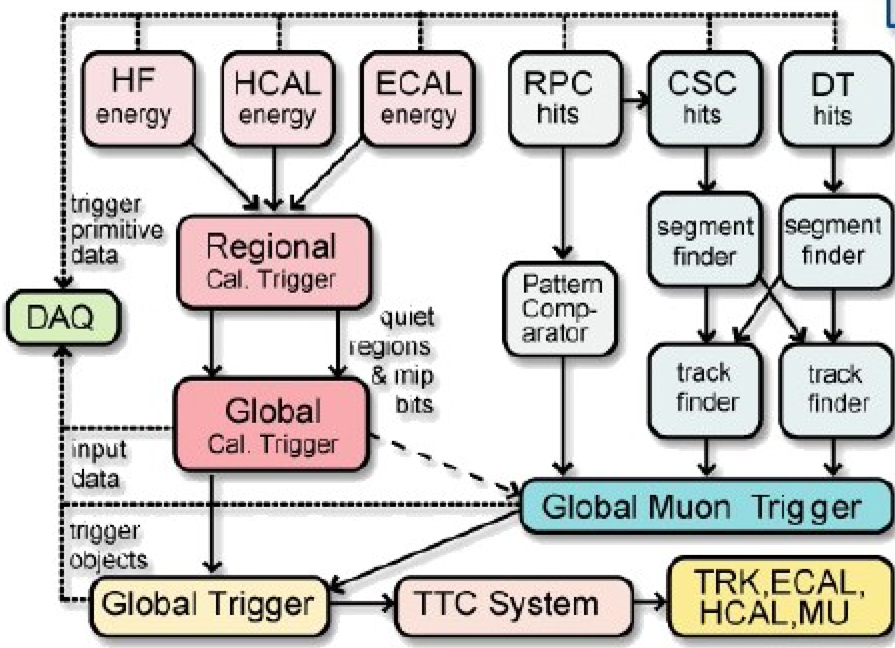


The CMS Trigger/DAQ System

Overall Trigger & DAQ Architecture: 2 Levels

Level-1 Trigger:
- 3.2 μs latency
- 100 kHz output

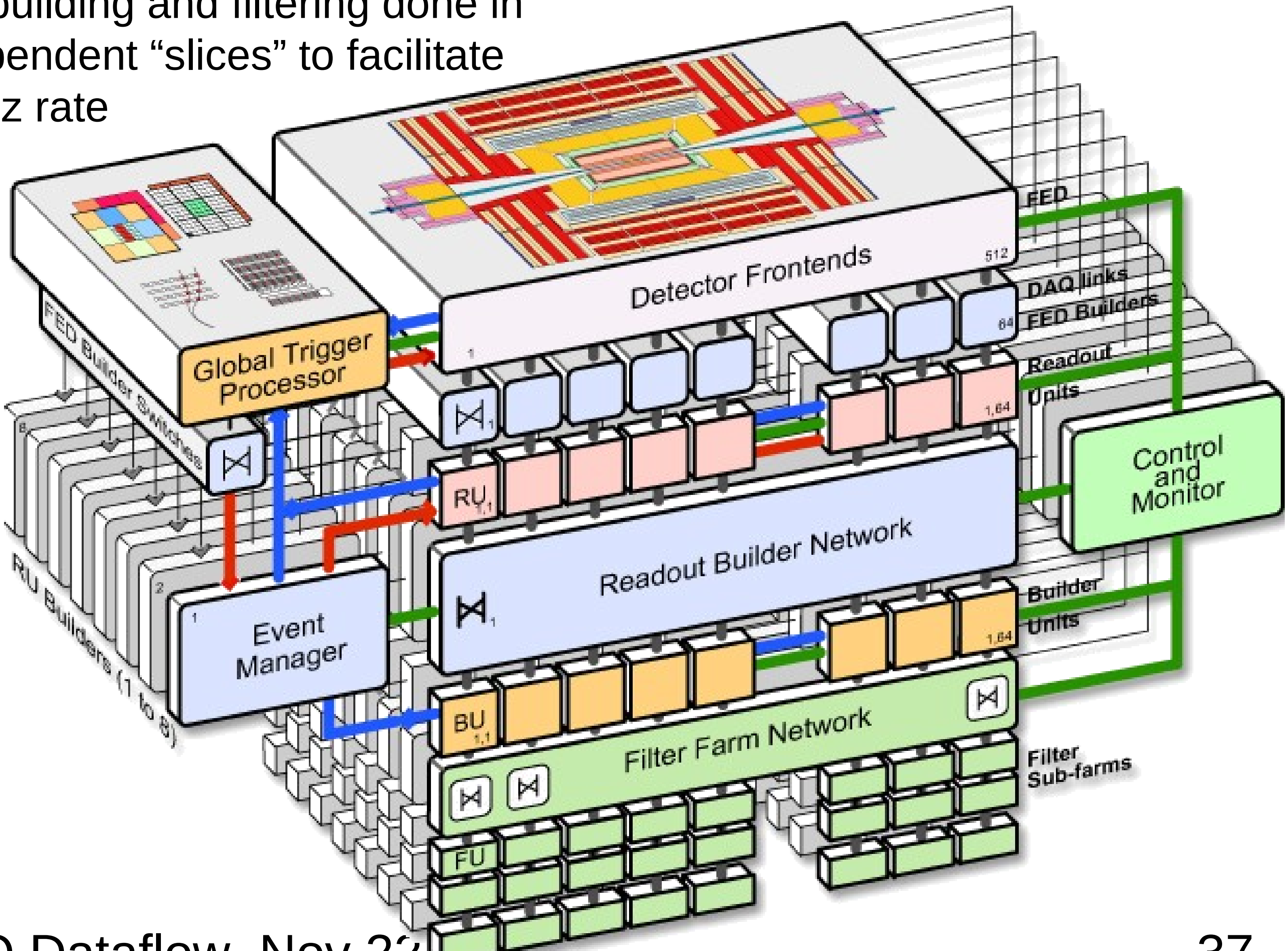




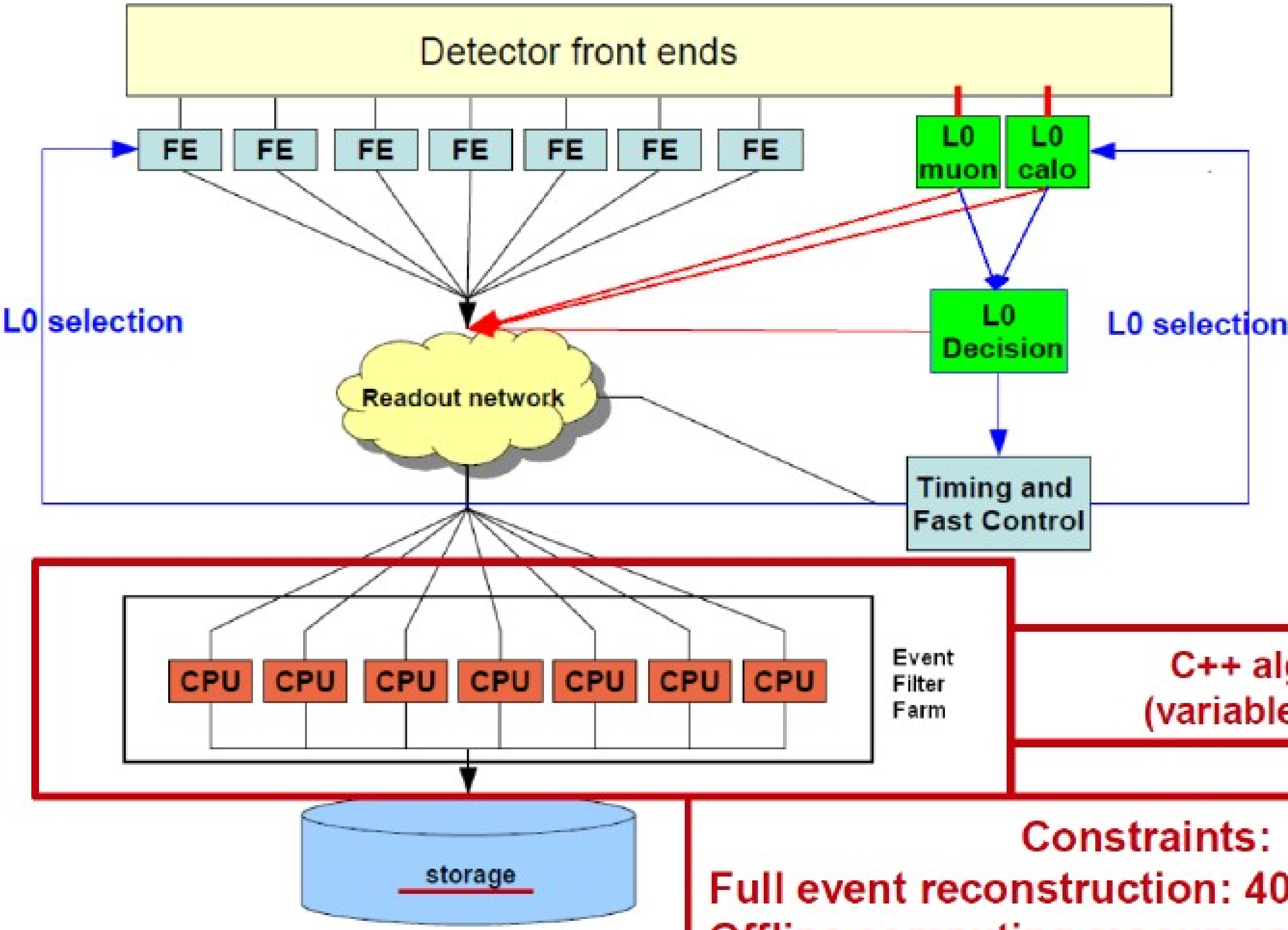
DAQ/HLT:
- Event building at full L1 rate
- Average event size: 1 MB
- Average HLT time: 50 ms
- Output rate: 100-300 Hz

CMS “3D” Event Builder

Event building and filtering done in 8 independent “slices” to facilitate 100 kHz rate



LHCb DAQ System



L0 Latency: 4μs
L0 rate: 1 MHz

C++ algorithm
(variable latency)

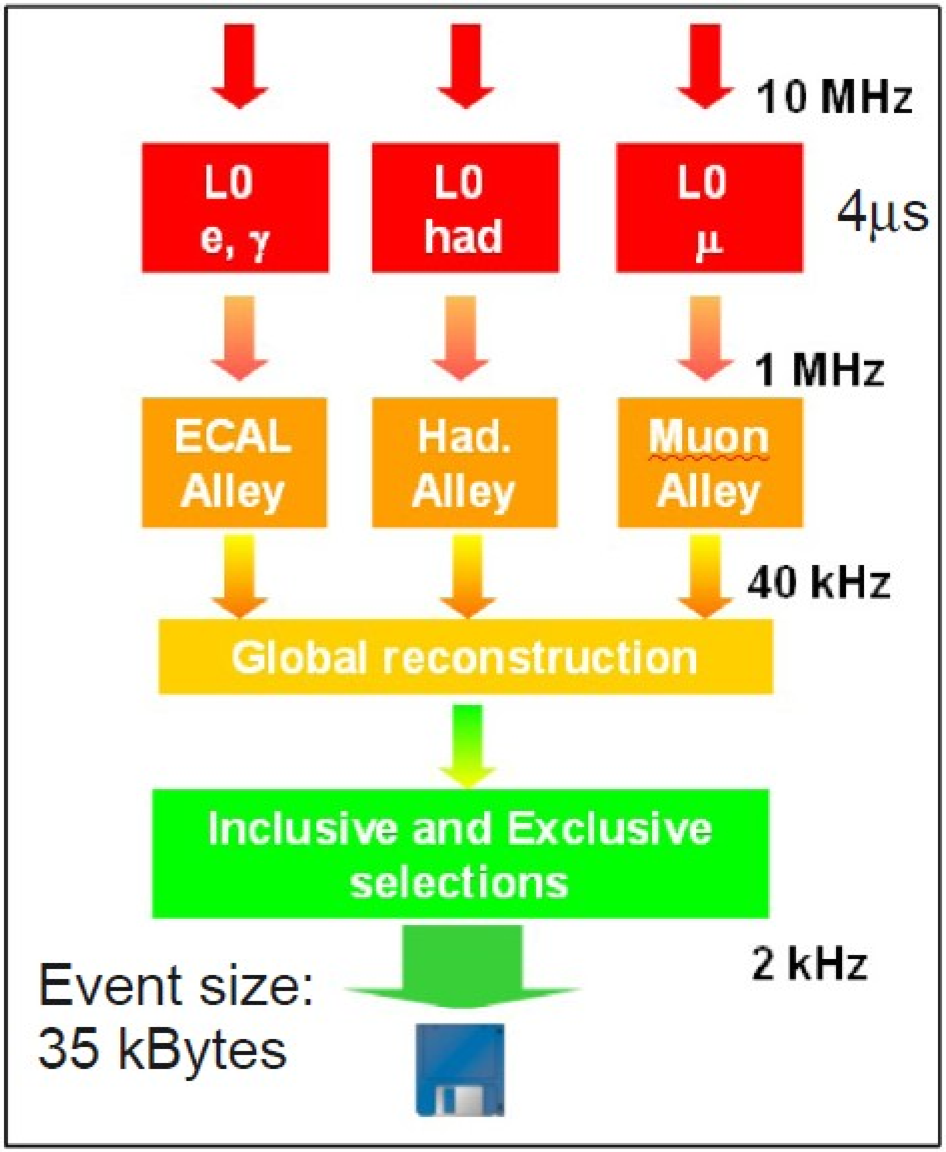
Constraints:
Full event reconstruction: 40 kHz (Hlt1)
Offline computing resources: 2 kHz (Hlt2)

LHC-b Trigger System

Level-0 (hardware): High E_T / p_T candidates

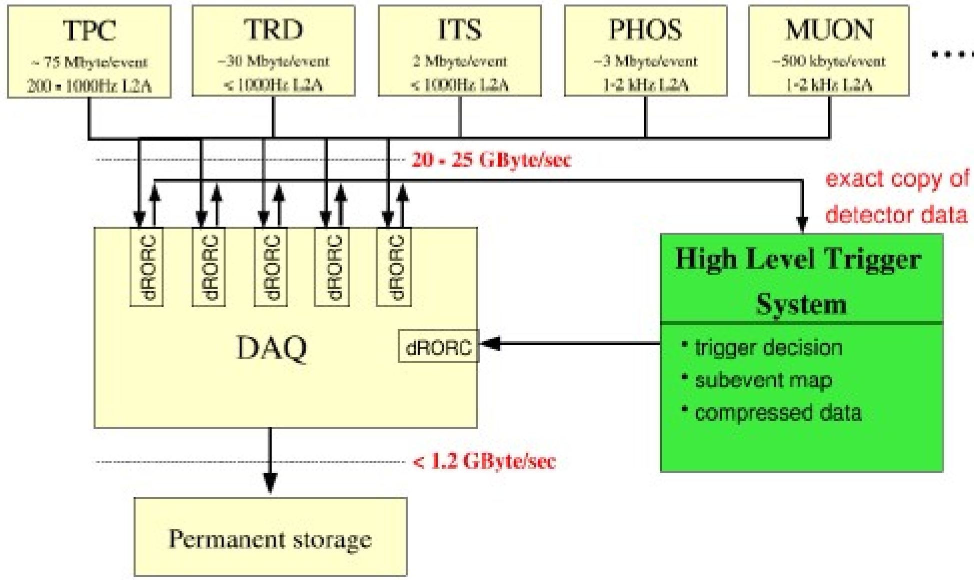
HLT1 (software):
- Partial Reconstruction on ROIs to confirm L0 candidates
- Use VELO for IP filter
- Add extra tracks

HLT2 (software):
- Full Reconstruction of event
- RICH available for PID
- Few tracks (inclusive)
- All tracks (exclusive)



ALICE Trigger/DAQ System

ALICE has different constrains:
- Low rate (max 8 kHz of Pb+Pb)
- Very large events (>40 Mbytes)
- Slow detector (TPC ~ 100 µs)



3 levels of hardware triggers:

Collision

L0: Trigger detectors detect collision (V0/T0, PHOS, SPD, TOF, dimuon trigger

L1: select events according to
- centrality (ZDC, ...)
- high-pt di-muons
- high-pt di-electrons (TRD)
- high-pt photons/pi0 (PHOS)
- jets (EMCAL, TRD)

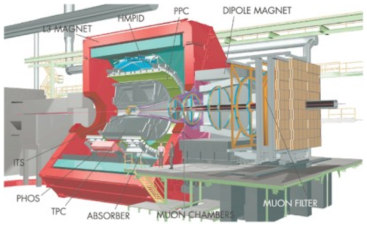
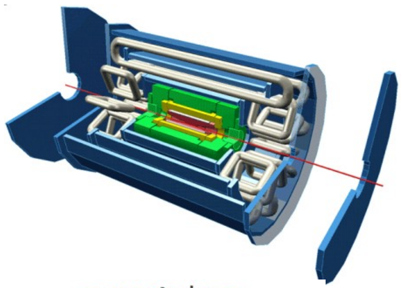
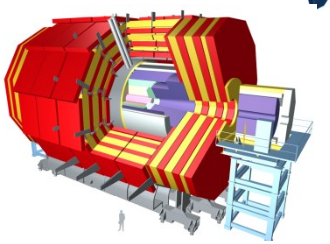
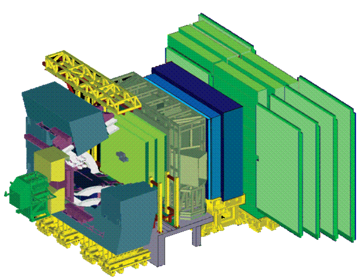
L2: reject events due to past/future protection

HLT process data in parallel with event building
Also does event compression

HLT rejects events containing
- **no J/psi, Y**
- **no D0**
- **no high-pt photon**
- **no high-pt pi0**
- **no jet, di-jet, γ-jet**

0 1.2 6.5 88 t [µsec]

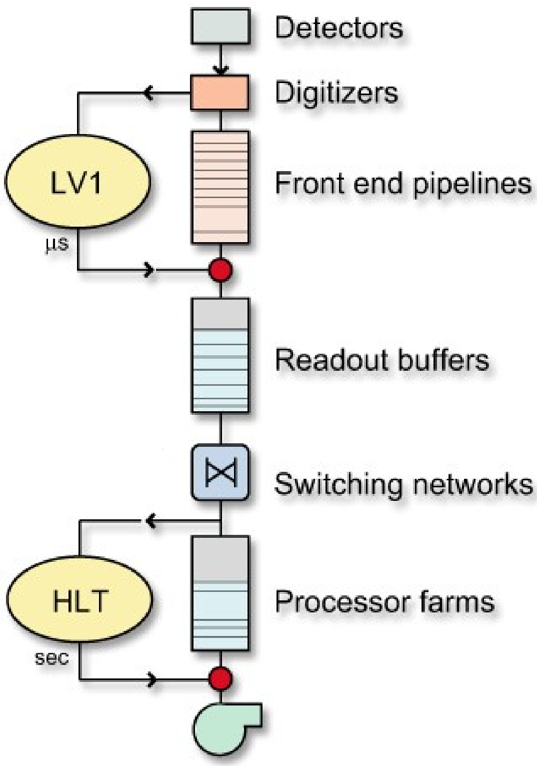
Trigger/DAQ parameters

	No.Levels Trigger	Level-0,1,2 Rate (Hz)	Event Size (Byte)	Readout Bandw.(GB/s)	HLT Out MB/s (Event/s)
	4	Pb-Pb 500 p-p 10^3	$5x10^7$ $2x10^6$	25	1250 (10^2) 200 (10^2)
	3	LV-1 10^5 LV-2 $3x10^3$	$1.5x10^6$	4.5	300 ($2x10^2$)
	2	LV-1 10^5	10^6	100	~1000 (10^2)
	2	LV-0 10^6	$3.5x10^4$	35	70 ($2x10^3$)

CMS Event Building

40 MHz
Clock driven
Custom processors

100 kHz
Event driven
PC network



Level-1 Trigger
Custom design

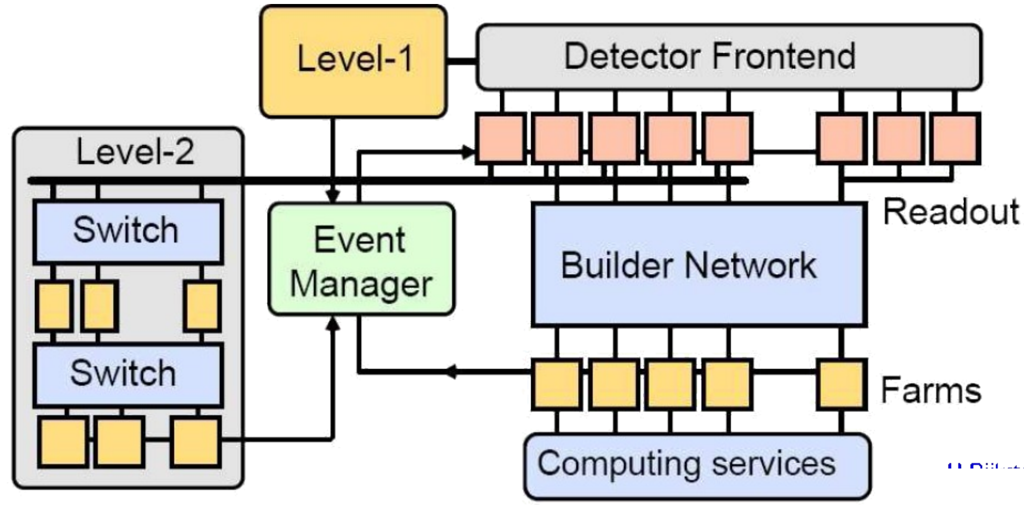
High-Level Trigger

Industry products

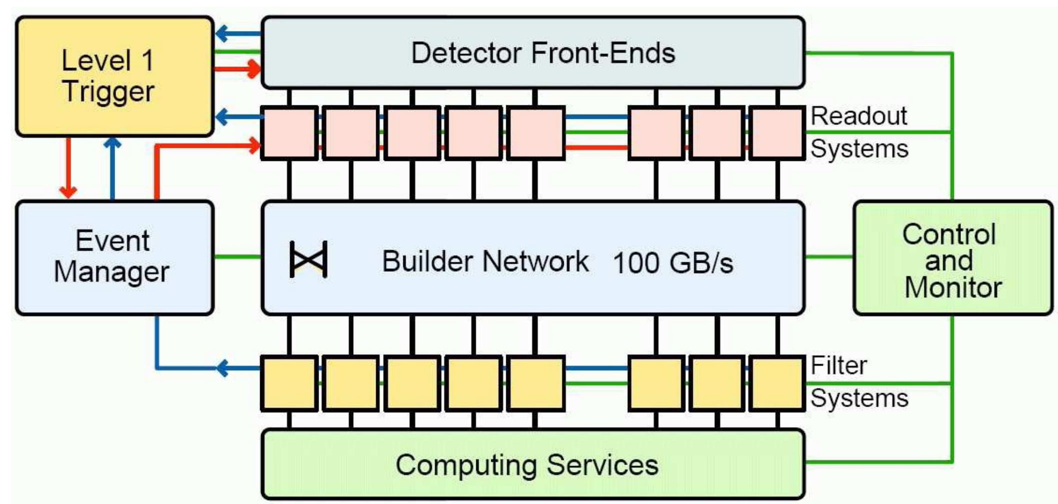
Level-1 output / HLT input	**100 kHz**
Network bandwidth	**1 Terabit/s**
HLT output	**10^2 Hz**
Invest in data transportation and CPU	

Event Builder

- Send a part first (RoI)
- Run L2 algorithms and decide if you want to keep the event
- If yes, send complete event data



- Alternative (used by CMS, Alice ad LHCb
  - Send everything, ask questions later
  - Much higher demand on networking



Lot’s of Abbreviations…

- Read-Out Drivers (ROD):
  - subdetector-specific,
  - collect and process data (no event selection)
- Read-Out Link (ROL)
  - 160 MByte/s optical fibre
- Read-Out Buffer input stage (ROBIN) card
  - Part of Readout system
  - 64-bit 66 MHz PCI card - 3 ROL inputs
- Read-Out Subsystem (ROS)
  - Set of PCs
  - Each PC contains 4 ROBINs => 12 ROLs per ROS PC