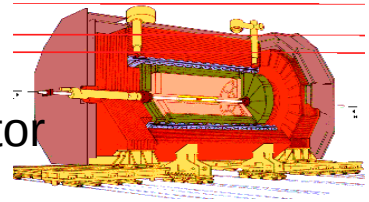
Introduction to Online System

北京大学物理学院　王大勇
dayong.wang@pku.edu.cn

The 2nd International Summer school on TeV Experimental Physics
Jinan, Aug 13, 2015

我们的数据是从哪里来的？

A modern HEP experiment



detector

Trigger and DAQ

signals

Data Acquisition System

Trigger System

decisions

raw data

Mass Storage

design feedback

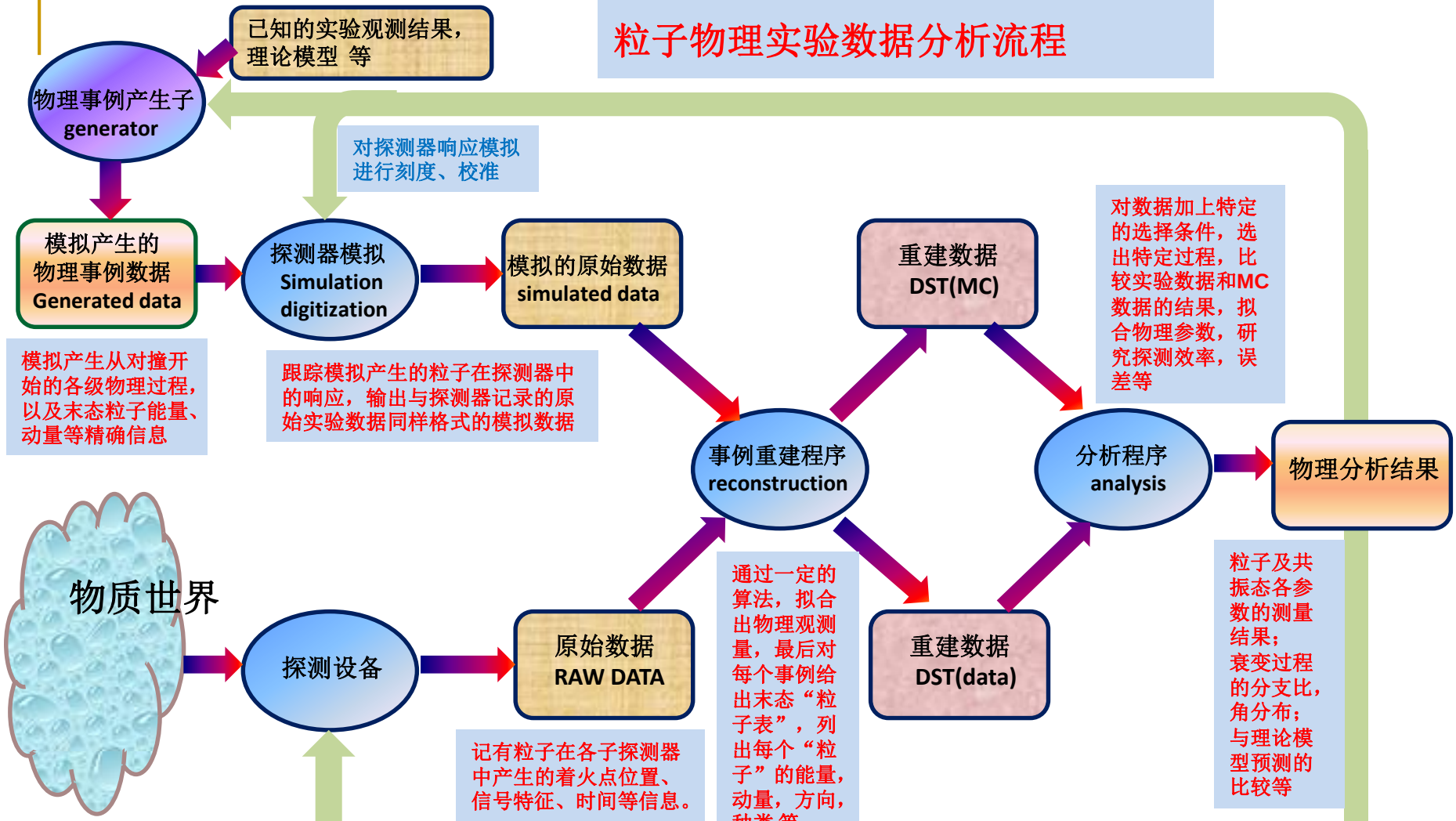
Detector & Trigger Simulation

Recons-truction & Analysis

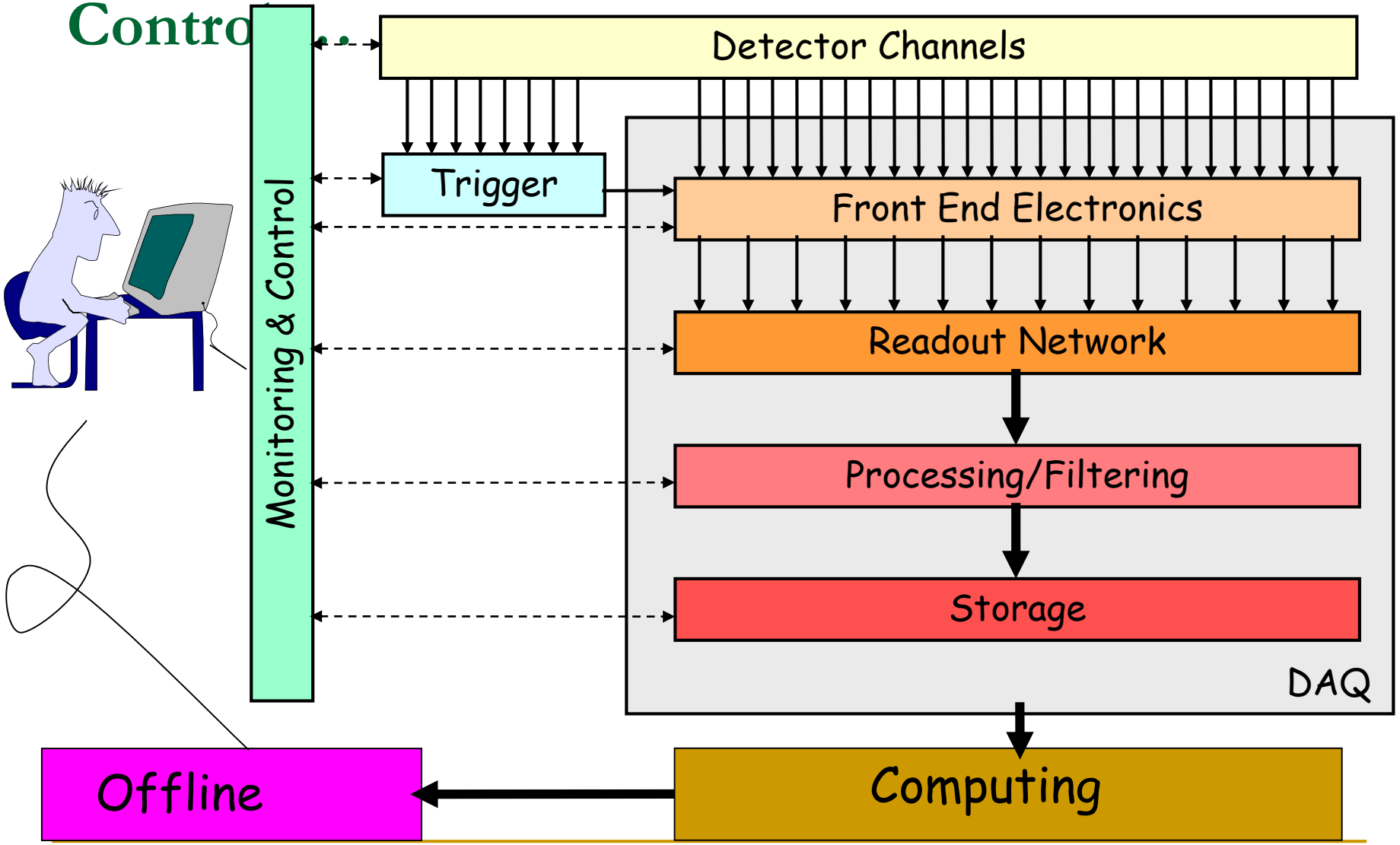
Physics Results

DAQ = Data **A**c**Q**uisition

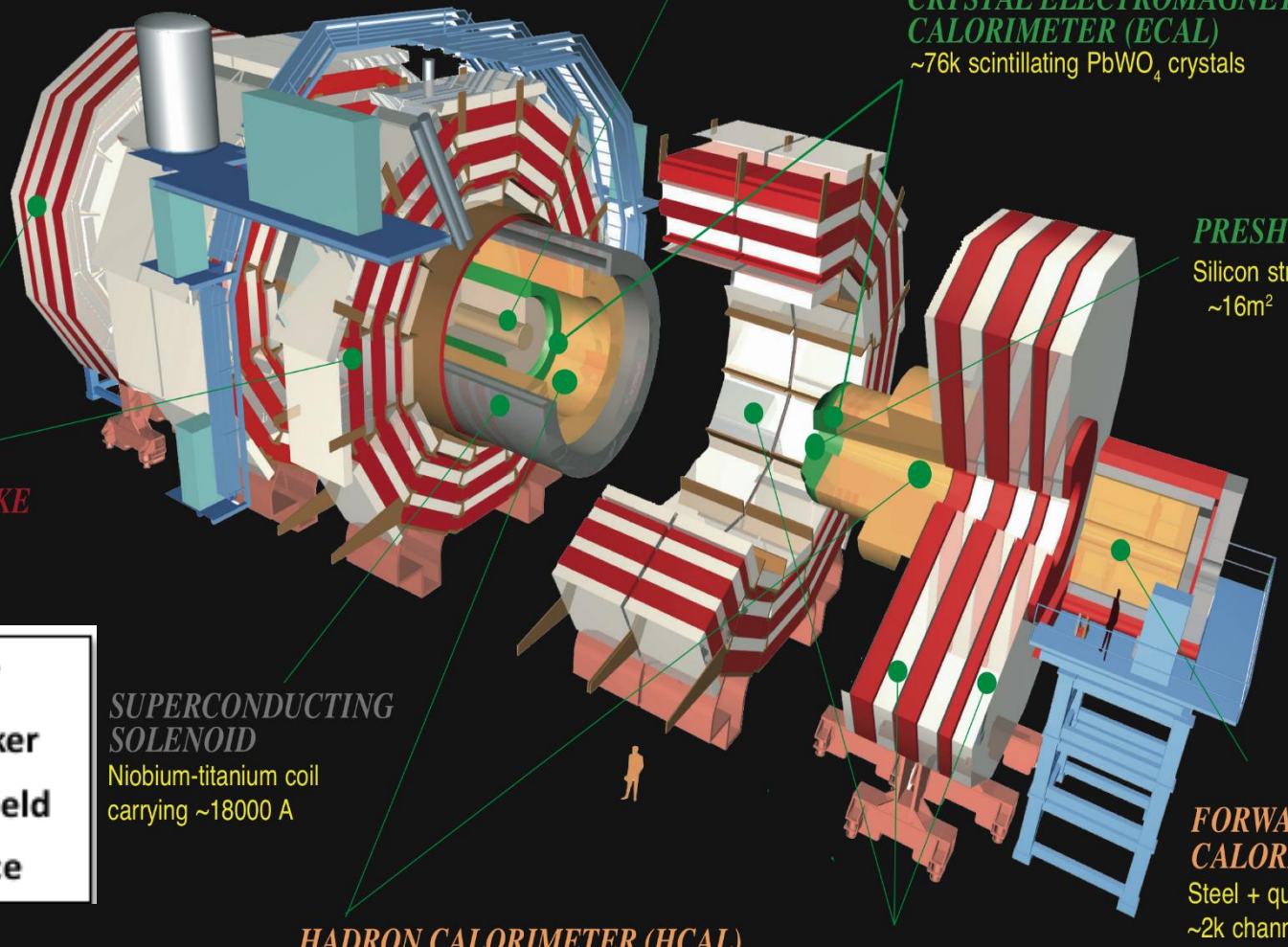
粒子物理实验数据分析流程



Online: Trigger, DAQ, Monitoring & Contro



CMS Detector



SILICON TRACKER
Pixels (100 x 150 μm²)
~1m² ~66M channels
Microstrips (80-180μm)
~200m² ~9.6M channels

CRYSTAL ELECTROMAGNETIC CALORIMETER (ECAL)
~76k scintillating $PbWO_4$ crystals

PRESHOWER
Silicon strips
~16m² ~137k channels

STEEL RETURN YOKE
~13000 tonnes

SUPERCONDUCTING SOLENOID
Niobium-titanium coil
carrying ~18000 A

FORWARD CALORIMETER
Steel + quartz fibres
~2k channels

HADRON CALORIMETER (HCAL)
Brass + plastic scintillator
~7k channels

MUON CHAMBERS
Barrel: 250 Drift Tube & 480 Resistive Plate Chambers
Endcaps: 473 Cathode Strip & 432 Resistive Plate Chambers

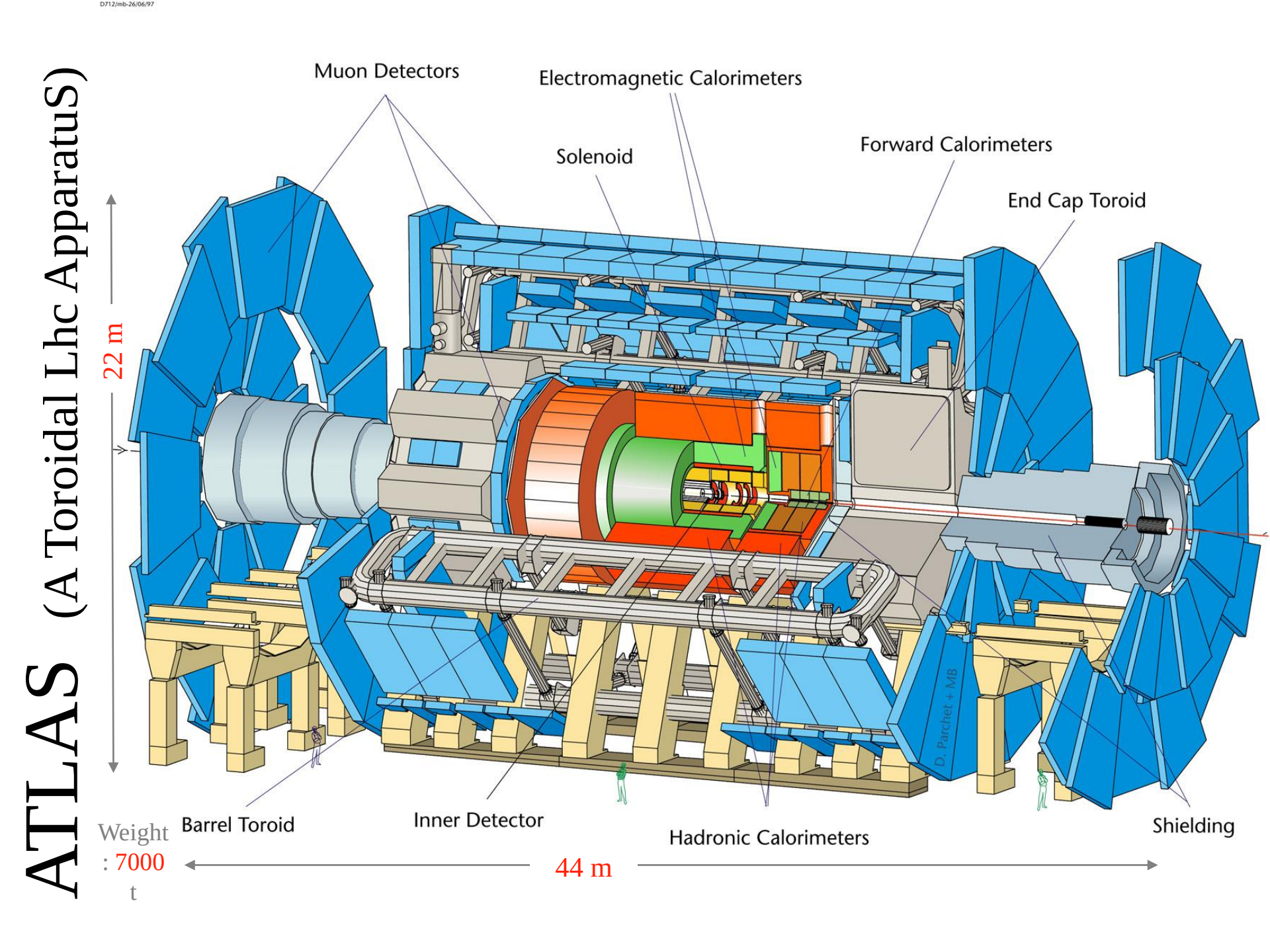
Flexible trigger

Large silicon tracker

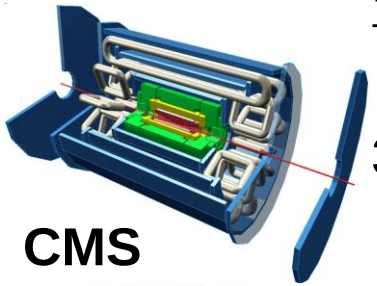
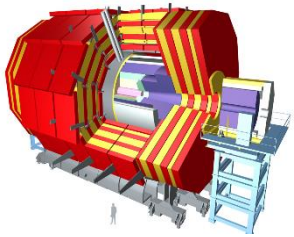
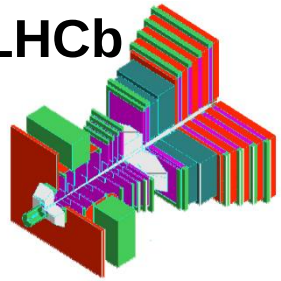
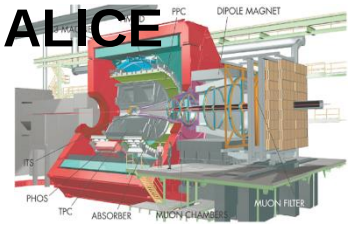
Strong magnetic field

Broad acceptance

Total weight	: 14000 tonnes
Overall diameter	: 15.0 m
Overall length	: 28.7 m
Magnetic field	: 3.8 T


ATLAS (A Toroidal Lhc ApparatuS)
Muon Detectors
Electromagnetic Calorimeters
Solenoid
Forward Calorimeters
End Cap Toroid
22 m
Weight : 7000 t
Barrel Toroid
Inner Detector
Hadronic Calorimeters
Shielding
44 m
D. Parchet + MB

LHC Trigger and DAQ summary

	No.Levels Trigger	First Level Rate (Hz)	Event Size (Byte)	Readout Bandw.(GB/s)	Filter Out MB/s (Event/s)
ATLAS 	**3**	**10^5** LV-2 **10^3**	**10^6**	**10**	**100** (10^2)
CMS 	**2**	**10^5**	**10^6**	**100**	**100** (10^2)
LHCb 	**3**	LV-0 **10^6** LV-1 **4 10^4**	**2x10^5**	**4**	**40** (2x10^2)
ALICE 	**4**	Pp-Pp **500** p-p **10^3**	**5x10^7** **2x10^6**	**5**	**1250** (10^2) **200** (10^2)

- 实验的取数环境
  - 加速器的时间结构，事例率和数据率，数据量，死时间
- 触发判选系统
  - 对触发判选系统的要求，多级触发，触发判选的物理原则、方案的实现和系统性能的监督，触发效率的测量，高亮度下的触发判选
- 数据获取系统
  - 数据获取系统的任务，单CPU的数据获取系统，新一代高能物理实验的数据获取系统，实时操作，系统数据获取系统的仿真
  - 在线监测与控制
  - 计算系统：实验数据的记录、(储存、传输与处理)

高能实验的取数环境

- 高能核物理和粒子物理实验中，除了感兴趣的好事例以外，还存在着上万倍甚至更多的本底，需要一个触发判选系统实时地排除大量本底，选出有意义的好事例，再通知数据获取系统收集处理。
- 各探测器在实验装置中彼此相对独立，而触发判选和数据获取系统则有关实验装置的整体，它的结构取决于探测器及其读出电子学的具体结构、加速器的时间结构、触发率、数据量以及实验目的。

加速器的时间结构

- 在对撞机实验中，一些粒子束团(长度为0.1~5cm)和另一些粒子束团周期性地对撞。对撞周期为μs级的实验中，触发判选系统有可能在下一次对撞来到之前做出初步判选。对撞周期小于0.5μs级时，考虑到探测器信号的延迟(如漂移时间)、电缆的延迟等，触发判选系统无法在一个对撞周期内做出任何判选，必须用完全不同的流水线结构。
- 在同步加速器上的固定靶实验中束流的时间结构不一样。束流的同步加速时间10s左右，在加速期间没有束流轰击靶和探测器，然后是束流引出时间。对中微子实验用快引出，约几毫秒，而强子实验一般用慢引出，约1~2s。

事例率和数据率

实验的灵敏度可用亮度来度量，它表示当反应截面为1cm^{2}时每秒钟的事例数。在固定靶实验中，亮度L是单位面积内的靶原子数与束流流强的乘积。在对撞机的情况下，亮度

$$L = n_1 n_2 f K_b / 4\pi\sigma_x\sigma_y$$

式中f为束流绕环转动频率，K_b为每束内的束团数，n_1、n_2为每个束团内的粒子数，σ_x、σ_y表征高斯分布的束团横向截面。

一些对撞机及相应探测器的参数

对撞机	最大束流能量/GeV	亮度/(10^{30} cm^{-2} s^{-1})	周长/km	束团数	对撞周期/ns	探测器	电子学道数
DAΦNE	0.75	5	0.0977	30~120	2.7 ~ 10.8	KLOE	23k
BEPC	2.2	5	0.2404	1	802	BES	20k
CESR	6	830	0.768	9X4	14-220	CLEO	400k
LEP	101	10000	26.66	5120	22000	ALEPH etc	100~ 300k
KEKB	8+3.5	3000	3.016	1658	2	BELLE	133k
HERA	e30+ p920	14	6.336	189+180	96	H1 etc	250k
Tevatron	1000	210	6.28	36	396	CDF etc	75~100k
LHC	7000	10000	26.66	2835	25	CMS etc	100M

Luminosity Basics

Mean number of inelastic interactions per beam-cross **ε*μ = Mean number of interactions per beam-cross seen by detector**

$$\mathcal{L} = \frac{\mu n_b f_r}{\sigma_{inel}} = \frac{\mu_{vis} n_b f_r}{\sigma_{vis}}$$

Inelastic cross section (unknown)

Cross section seen by detector

- σ_{vis} is determined in dedicated fills based on beam parameters

Calibrating σ_{vis} in VdM Scans

- The Luminosity in terms of beam densities ρ_1 and ρ_2:

$$\mathcal{L} = n_b\, f_r\, n_1 n_2 \int \rho_1(x,y)\rho_2(x,y)dxdy$$

- Only if the integral factorises into independent x & y components:

$$\mathcal{L} = \frac{n_b\, f_r\, n_1 n_2}{2\pi\, \Sigma_x \Sigma_y}$$

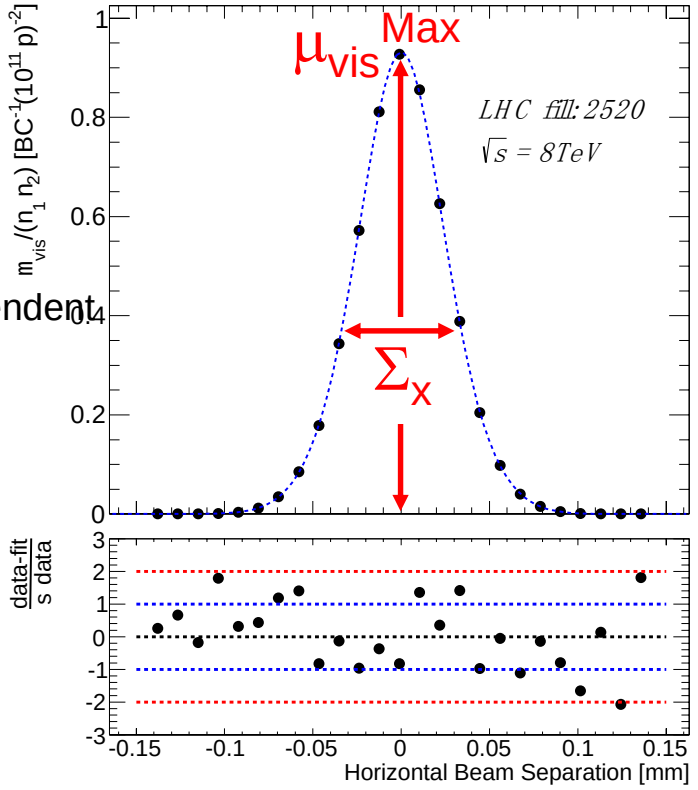
- This and previous slide imply

Measured in vdM scan Detectorindependent

$$\sigma_{vis} = \mu_{vis}^{Max} \frac{2\pi \Sigma_x \Sigma_y}{n_1 n_2}$$

Detectordependent

Measured by beam instrumentation



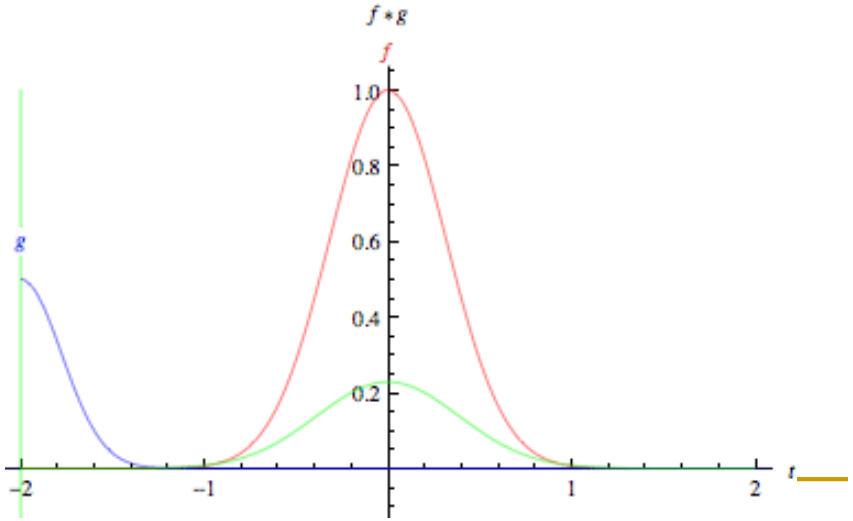
Van der Meer Scan

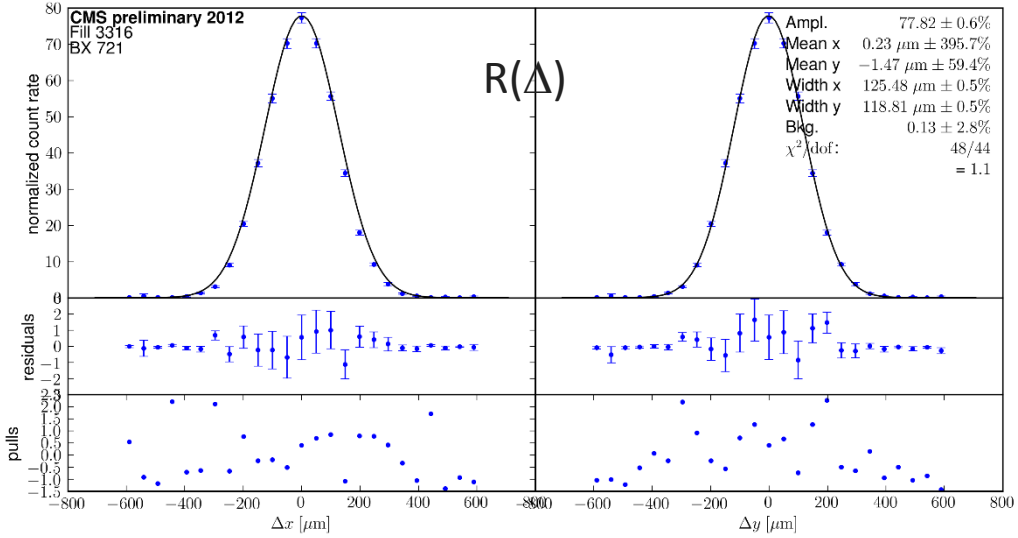
Assuming factorizable gaussian for the beam density function (not too bad as approximation)

$$\rho(x,y) = \rho(x)\rho(y) \propto \exp(-\frac{x^2}{2\sigma_x^2})\exp(-\frac{x^2}{2\sigma_y^2}) \Rightarrow F(\Delta_x,\Delta_y) \propto \exp(-\frac{\Delta_x^2}{2\Sigma_x^2})\exp(-\frac{\Delta_y^2}{2\Sigma_y^2})$$

The resulting "effective area" is then just:

$$\Sigma_x\Sigma_y = \sqrt{\sigma_{x,1}^2+\sigma_{x,2}^2}\sqrt{\sigma_{y,1}^2+\sigma_{y,2}^2}$$





好事例率：

$$n_{ph} = \sigma L$$

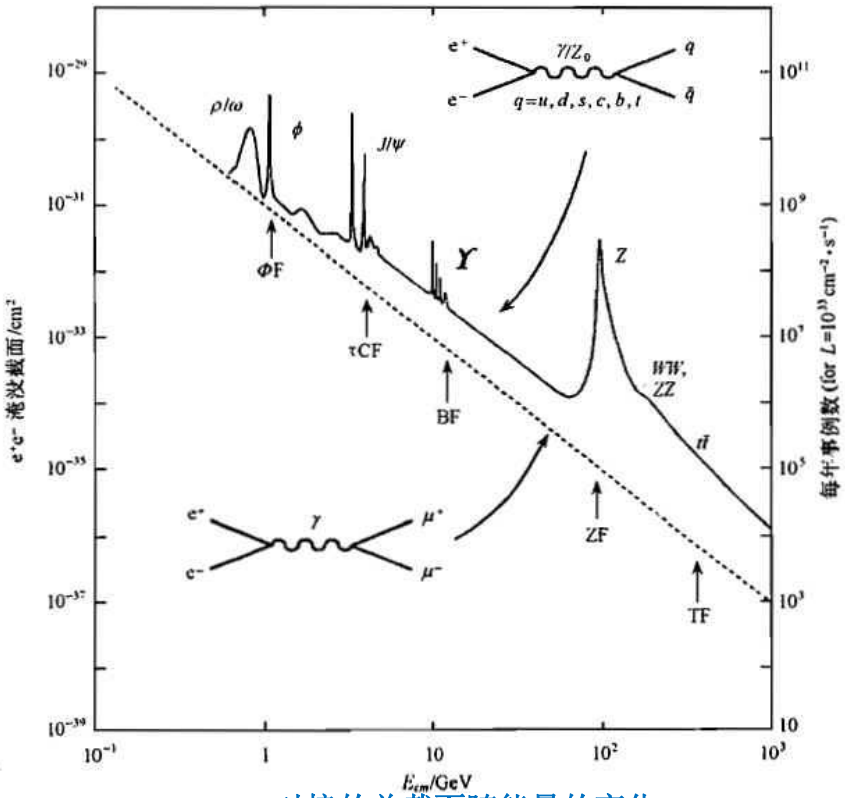
σ为所研究的物理过程的作用截面。但是有些共振峰如J/Ψ的宽度仅87keV，小于对撞机的能散度△E(~ 800 keV)，则只有峰宽范围内的亮度有效。实际运行中，BEPC在J/Ψ共振区n_{ph} <10Hz；在非共振区n_{ph} <1Hz。

本底事例率：

宇宙线本底：宇宙线通量约180 m^{-2} s^{-1}，探测器面积约10-100 m^2，所以宇宙线本底事例率为10^3-10^4 s^{-1}。

丢失束流粒子本底：由于束流粒子和真空中残余气体的作用以及束团内、束团间粒子的相互作用，使粒子动量改变而脱轨，这些脱轨的粒子如果击中探测器，就造成束流本底事例。例如BEPC估算有约3X10^5 s^{-1} 。在对撞区，本底随Φ方向的变化和对撞机结构密切相关。

其他本底：如同步辐射造成的散粒状击中本底和电子学干扰造成的假事例等。

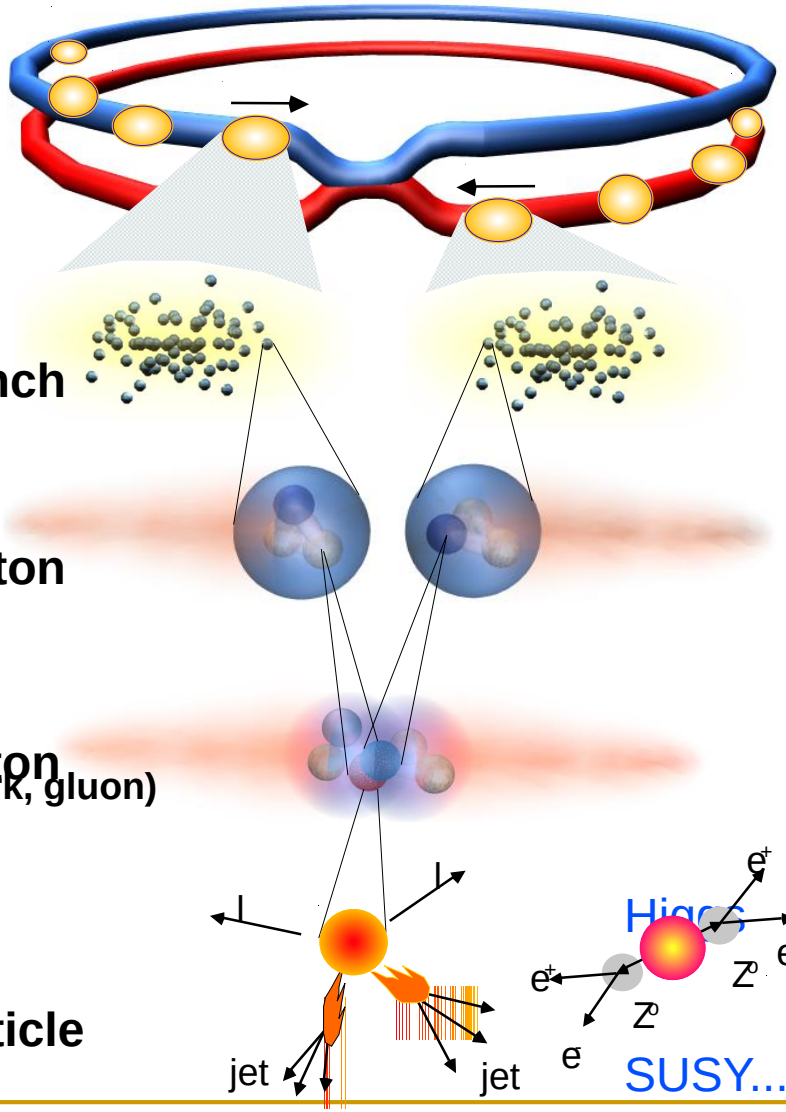


e+e-对撞的总截面随能量的变化.

不感兴趣的物理事例：

和e+e-对撞机不同，在pp对撞机(如Teva t ron , LHC)及强子束固定靶实验中强相互作用占优势，总作用截面高达100mb。当亮度L= 10^{34} m^{-2} s^{-1}时，事例率n_{ph} >$10^9$$s^{-1}$，即每次对撞就可能发生~30个事例，但其中绝大多数是物理上不感兴趣的小角度散射(minimum bias events)，真正感兴趣的如H^0事例差不多每分钟才产生一个。这样就要在10^{10}个本底事例中选出一个H^0事例。

Collisions at the LHC:



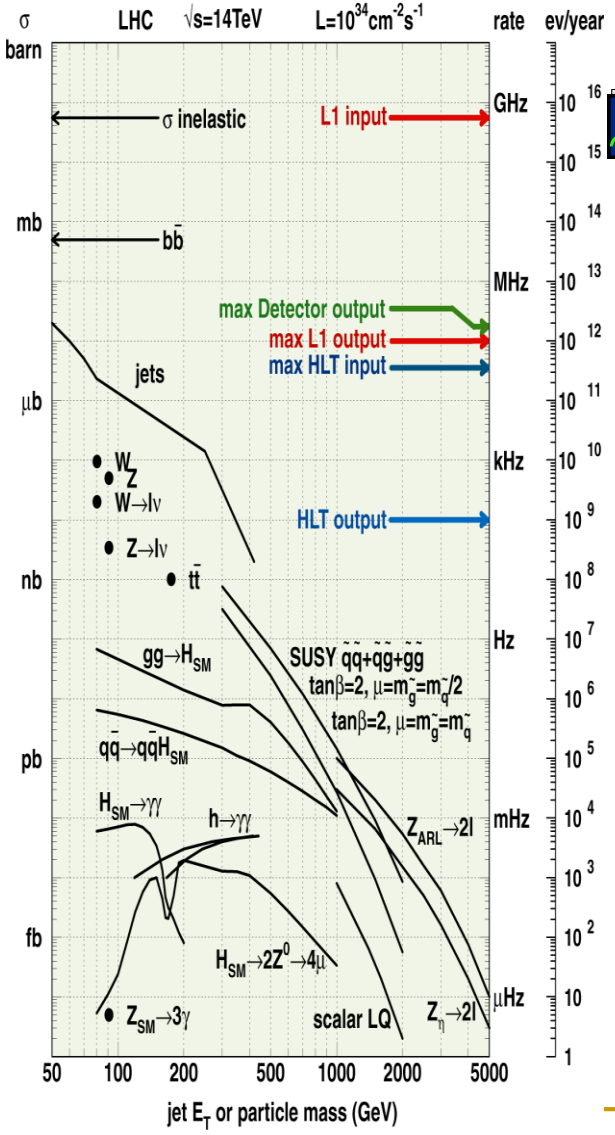
Proton - Proton

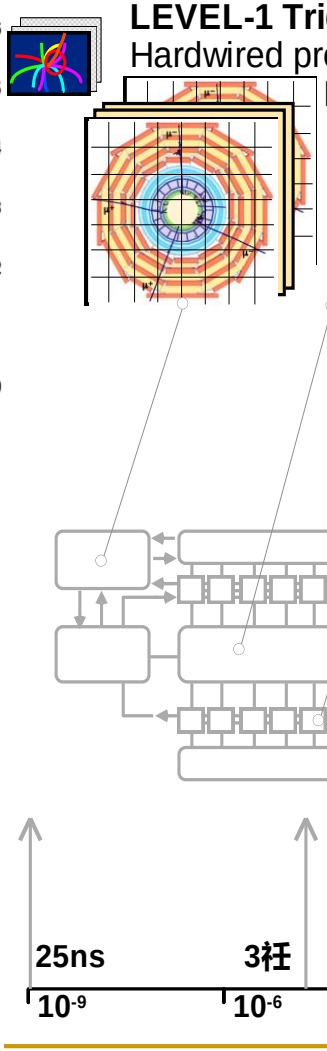
Protons/bunch	2804 bunch/beam 10^{11}
Beam energy	7 TeV (7×10^{12} eV)
Luminosity	$10^{34} cm^{-2} s^{-1}$
Crossing rate	40 MHz(every25ns)
Collision rate ≈	$7 \times 10^{8}\ s^{-1}$
data recording rate	300 Hz
	200-300 MB/sec

New physics rate ≈ .00001 Hz

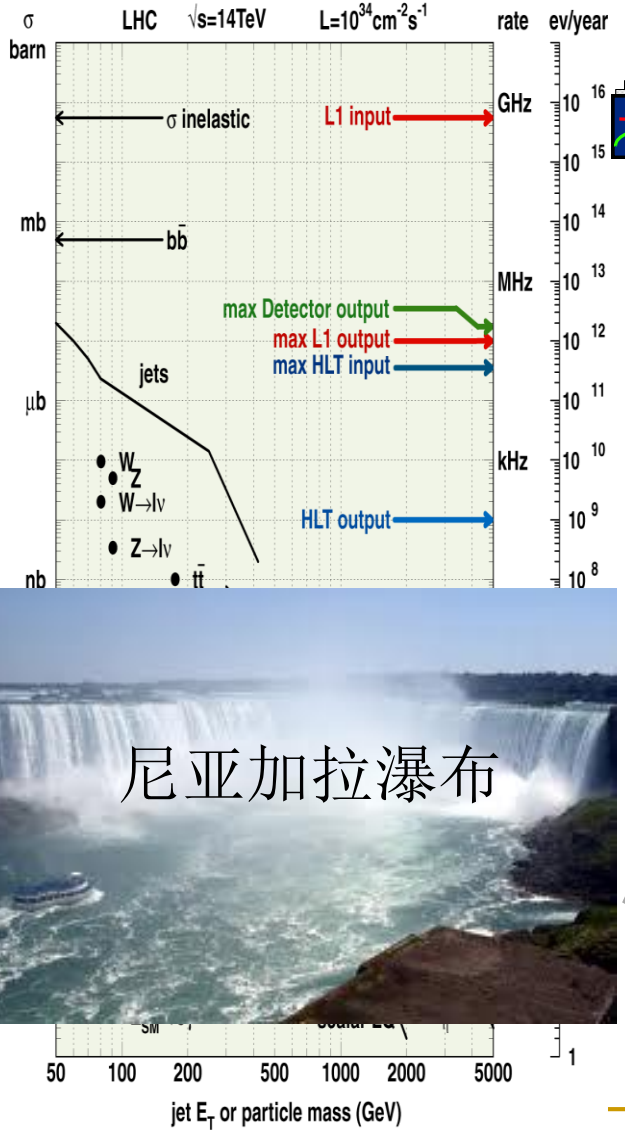
Event selection:
1 in 10,000,000,000,000

LHC实验数据流：物理信号的挑选





触发系统：当代对撞实验的大脑和心脏



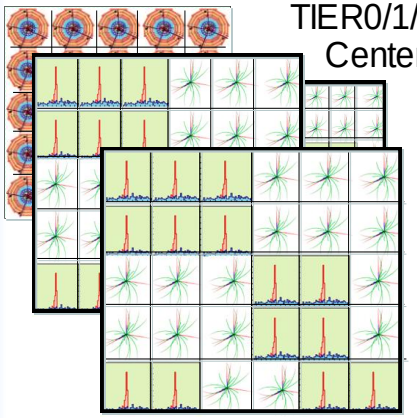
ON-line

OFF-line

LEVEL-1 Trigger
Hardwired processors (ASIC, FPGA)
Pipelined massive parallel

HIGH LEVEL Triggers
Farms of processors

Reconstruction&ANALYSIS
TIER0/1/2 Centers



25ns 3μs ms sec hour year
10^{-9} 10^{-6} 10^{-3} 10^{-0} 10^{3}
Giga Tera Petabit

物理结果

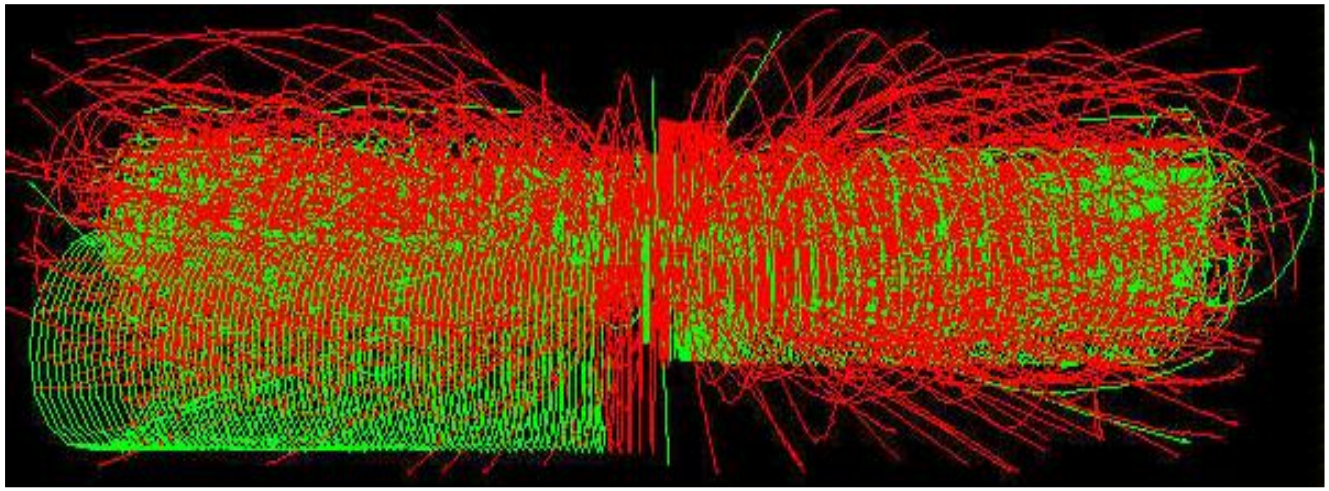
Why Trigger?

pp collisions at 14 TeV at 10^{34} $cm^{-2}s^{-1}$

- ~30 min bias events overlap
- H→ZZ

Z →μμ

H→ 4 muons: the cleanest (“golden”) signature

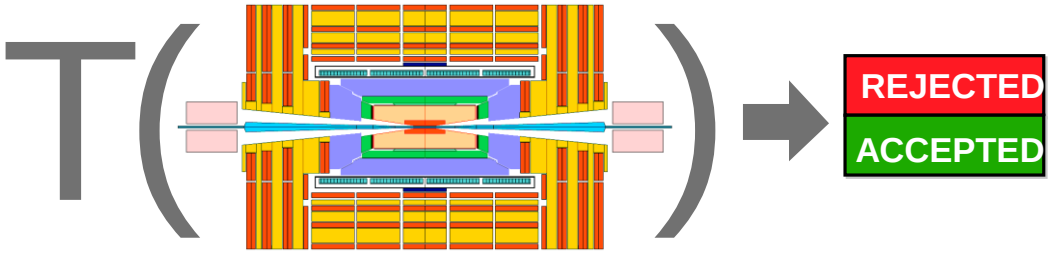


And this (not the H though…) repeats every 25 ns…

Triggering

- Task: inspect detector information and provide a first decision on whether to keep the event or throw it out

The trigger is a function of :



Event data & Apparatus
Physics channels & Parameters

- **Detector data not (all) promptly available**
- **Selection function highly complex**

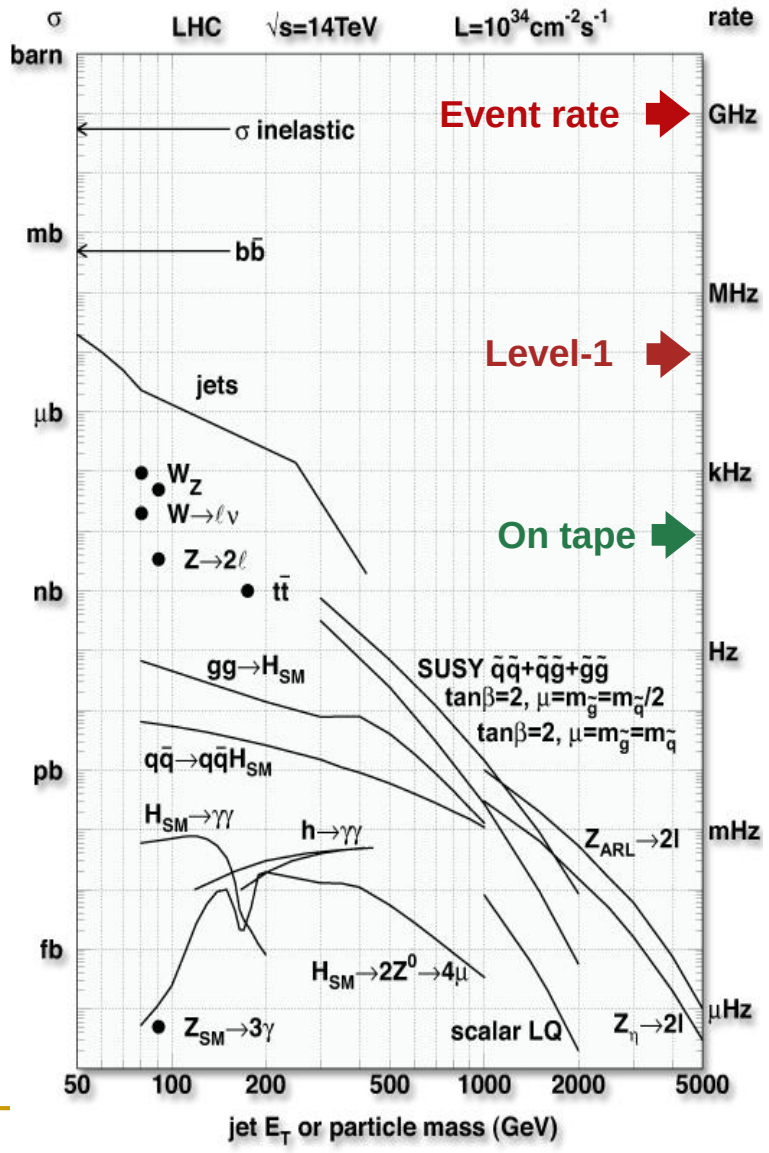
⇒T(...) is evaluated by successive approximations, the

TRIGGER LEVELS

(possibly with zero dead time)

Selectivity: the physics

- Cross sections of physics processes vary over many orders of magnitude
  - Inelastic: 10^9 Hz
  - W→ ● ν: 10^2 Hz
  - t t production: 10 Hz
  - Higgs (100 GeV/c^2): 0.1 Hz
  - Higgs (600 GeV/c^2): 10^{-2} Hz
- QCD background
  - Jet E_T ~250 GeV: rate = 1 kHz
  - Jet fluctuations → electron bkg
  - Decays of K, π, b → muon bkg
- Selection needed: $1:10^{10-13}$
  - Before branching fractions...

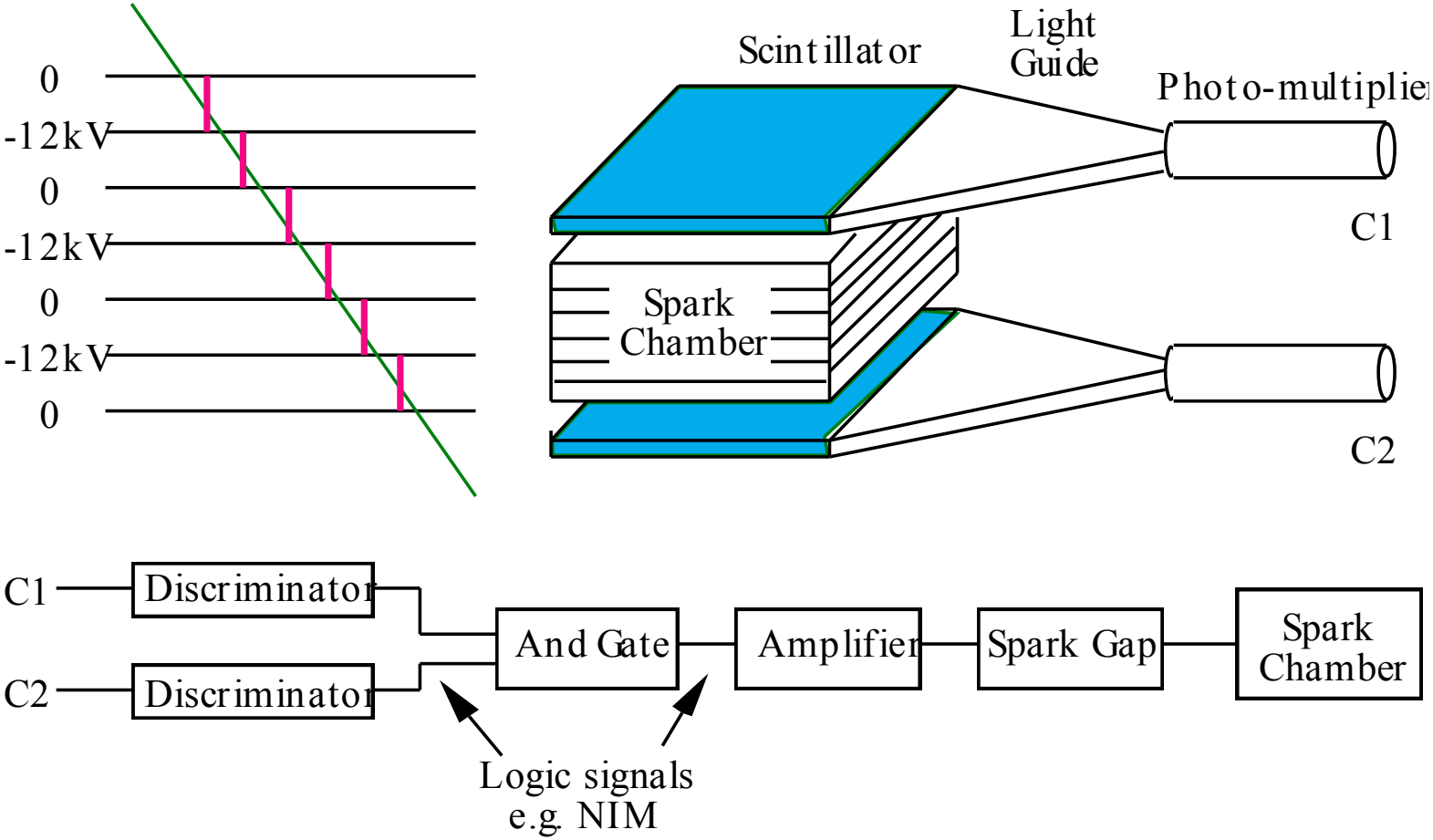


General trigger requirements

- The role of the **trigger** is to make the **online selection** of particle collisions potentially containing interesting physics
- Need **high efficiency** for selecting processes of interest for physics analysis
  - Efficiency should be precisely known
  - Selection should not have biases that affect physics results
- Need **large reduction of rate** from unwanted high-rate processes (capabilities of DAQ and also offline computers)
  - Instrumental background
  - High-rate physics processes that are not relevant for analysis
- System must be **affordable**
  - Limits complexity of algorithms that can be used
- Not easy to achieve all the above simultaneously!
- And never forget that **an event rejected by the Trigger is lost for ever!**

No Trigger, No Physics !

Simple trigger for spark chamber set-up



Trigger systems 1980's and 90's

- bigger experiments $\rightarrow$ more data per event
- higher luminosities $\rightarrow$ more triggers per second
  - both led to increased fractional deadtime

- use multi-level triggers to reduce dead-time
  - first level - fast detectors, fast algorithms
  - higher levels can use data from slower detectors and more complex algorithms to obtain better event selection/background rejection

Trigger systems 1990's and 2000's

- Dead-time was not the only problem
- Experiments focussed on rarer processes
  - Need large statistics of these rare events
  - But increasingly difficult to select the interesting events
  - DAQ system (and off-line analysis capability) under increasing strain - limiting useful event statistics
    - This is a major issue at hadron colliders, but will also be significant at ILC
- Use the High Level Trigger to reduce the requirements for
  - The DAQ system
  - Off-line data storage and off-line analysis

Trigger System Functionality

- For many physics analyses, aim is to obtain as high statistics as possible for a given process
  - We cannot afford to handle or store all of the data a detector can produce!
- What does the trigger do
  - select the most interesting events from the myriad of events seen
    - I.e. Obtain better use of limited output band-width
    - Throw away less interesting events
    - Keep all of the good events(or as many as possible)
  - But note must get it right - any good events thrown away are lost for ever!
- High level trigger allows much more complex selection algorithms

时间延迟的处理

无论在第一、二、三级触发中，都有对一个事例的处理时间大于事例到达的时间间隔这样一个问题。解决这个问题可以采用下面三种并行处理方法：

1) 时间上并行：建立一个专门设计的流水线式的处理器，它分为多个单元。而整个处理算法也分成同样多个计算步骤，每个单元执行一个固定的计算步骤，就把处理结果传给下一个处理单元，自己又从上一个处理单元接受新的数据，就像流水式生产线上每一个工人只做一步工序就把产品交给下一工序，在整个流水线上，多个单元分别地对不同事例进行不同的计算，最后从流水线的出口得到处理的结果。

2) 空间上并行：把整个事例数据分割成许多小块(例如来自某个子探测器的一小部分的数据)，每个处理器处理一个小数据块，再把处理结果送到全局处理器。全局处理器合各个局部处理器的结果，给出全局判选结果.

3)事例间并行：高能物理实验的数据有一个特点，就是各事例的数据之间没有关联；这样可以用多个处理器构成一个大的处理器阵，各个处理器都运行同样的程序，但各别处理不同事例的数据。这种并行计算机称为计算机群(Farm)。当然还有其他类型的并行计算机，如向量机等。但它们不如计算机群更适用于高能物理实验。

第一级触发采用了1).2)相结合的方法，第二级触发往往采用方法2)。方法1)用专门设计的硬件，可以得到最快的速度，但是不容易更改，造价也贵。方法3).灵活性最大，可扩充性也最好，在第三级触发即实时筛选中最适用。

触发判选系统性能的监测

每个事例读出信息和直方图：对一个好事例，可以通过读主触发的输出和输入得知该事例可能是哪一类事例，以及那些触发条件被满足。这些信息在单事例图上同时显示出来。对每一个好事例，还读入各子探测器单元的击中等信息，并积累成直方图可以随时察看，以了解探测器和触发子系统(如寻迹电路等)工作是否正常。触发效率的测量也要用这些数据。

判选过程中各种计数率的监测II定标器

在运行过程中，除了满足触发条件的好事例都被读出并记录到磁带上以外，还必须监测在各级判选过程中产生的各种信号的计数率，如各种触发条件信号的计数率、通过第一、二、三级触发的事例率以及死时间等，以便及时地判断谱仪工作是否正常。

但在上段所述的每个事例读出信息中所记录的都是通过第三级触发后的信息，中间过程均已丢失。所以要另外用定标器来记录这些计数率。一般要用两套并行的定标器。一套在控制台上，手动测量，用数码管显示，可以实时监测。通过这些计数率判断对撞机和探测器工作是否正常，并且可以分析和完善触发条件表。另一套是CAMAC或VME定标器插件，由数据获取系统每1 ~5min读出一次记录到磁带上去，主要用于死时间校正和显示各种计数率随时间的变化。

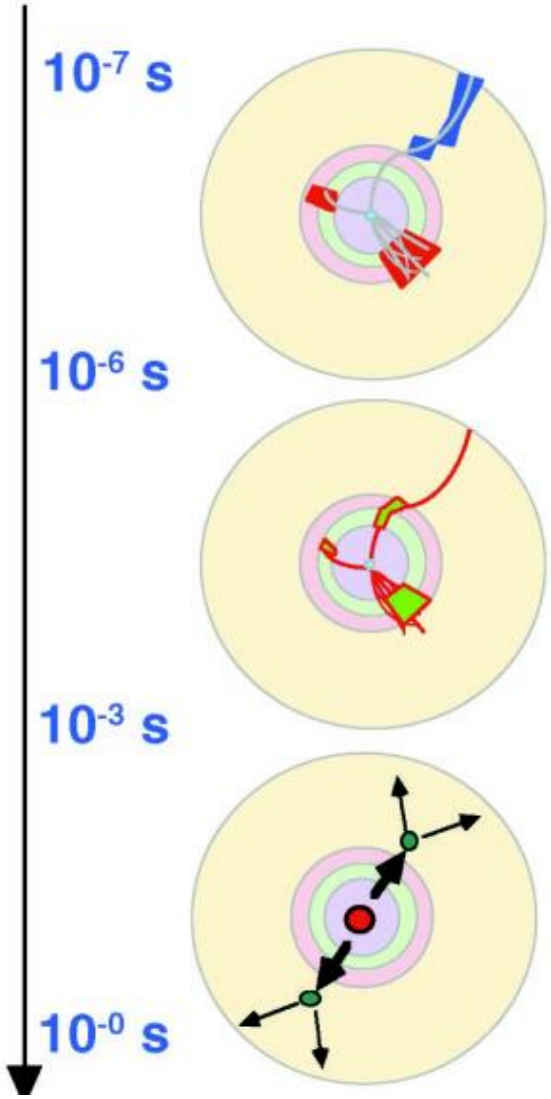
触发效率的测量

触发判选系统的工作状态直接影响谱仪获取数据的数量和质量。本底排斥比和触发效率是触发判选系统的两个重要指标，它们决定系统的触发率和数据的信噪比。其中触发效率还和物理分析密切相关。

离线分析中筛选过程将记录在磁带上的原始数据筛选，选出好事例以进行事例重建。如果因筛选条件不对而发生错误，可以改正条件重新筛选。但触发判选是实时的选择，一个事例只有通过触发判选后，在线系统才开始动作，把该事例的数据读入，记录在磁带上。一旦由于触发效率问题造成一个好事例没有通过触发判选，它就不能被记在带上而丢失，并且无法补救。

在物理分析中，求好事例总数、产生截面、分支比或Monte-Carlo模拟等都要求知道触发效率。

LHC Trigger Levels



Collision rate 10^9 Hz

Channel data sampling at 40 MHz

Level-1 selected events 10^5 Hz

Particle identification (High p_T e, μ, jets, missing E_T)

- Local pattern recognition
- Energy evaluation on prompt macro-granular information

Level-2 selected events 10^3 Hz

Clean particle signature (Z, W, ..)

- Finer granularity precise measurement
- Kinematics. effective mass cuts and event topology
- Track reconstruction and detector matching

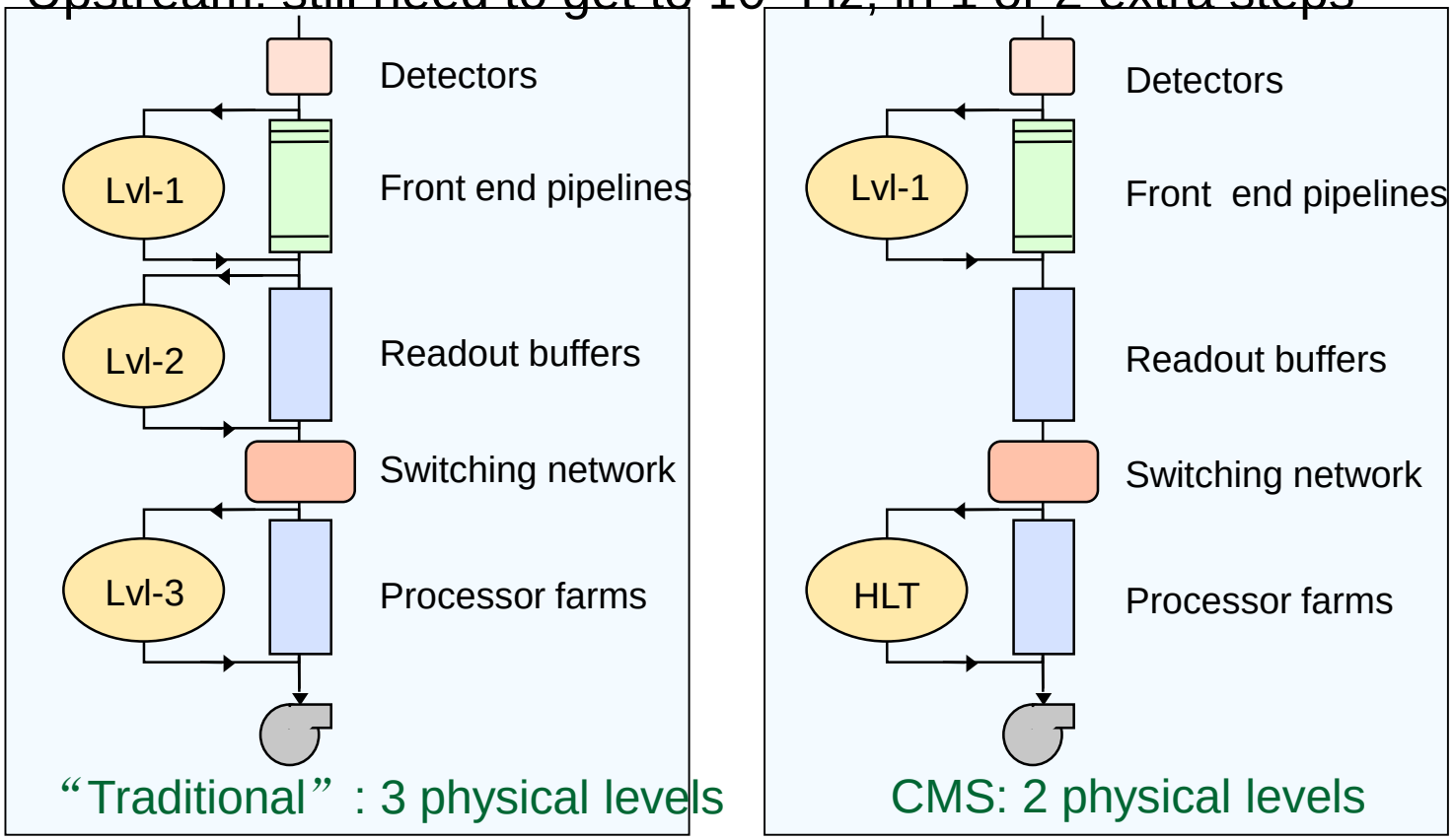
Level-3 events to tape 100- 400 Hz

Physics process identification

- Event reconstruction and analysis

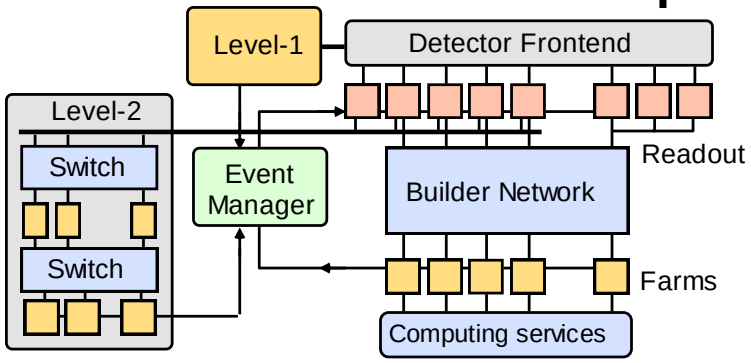
Level-1 Trigger

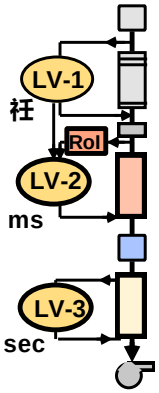
- Level-1 trigger: reduce 40 MHz to 10^5 Hz
  - This step is always there
  - Upstream: still need to get to 10^2 Hz; in 1 or 2 extra steps



Three physical entities(ATLAS)

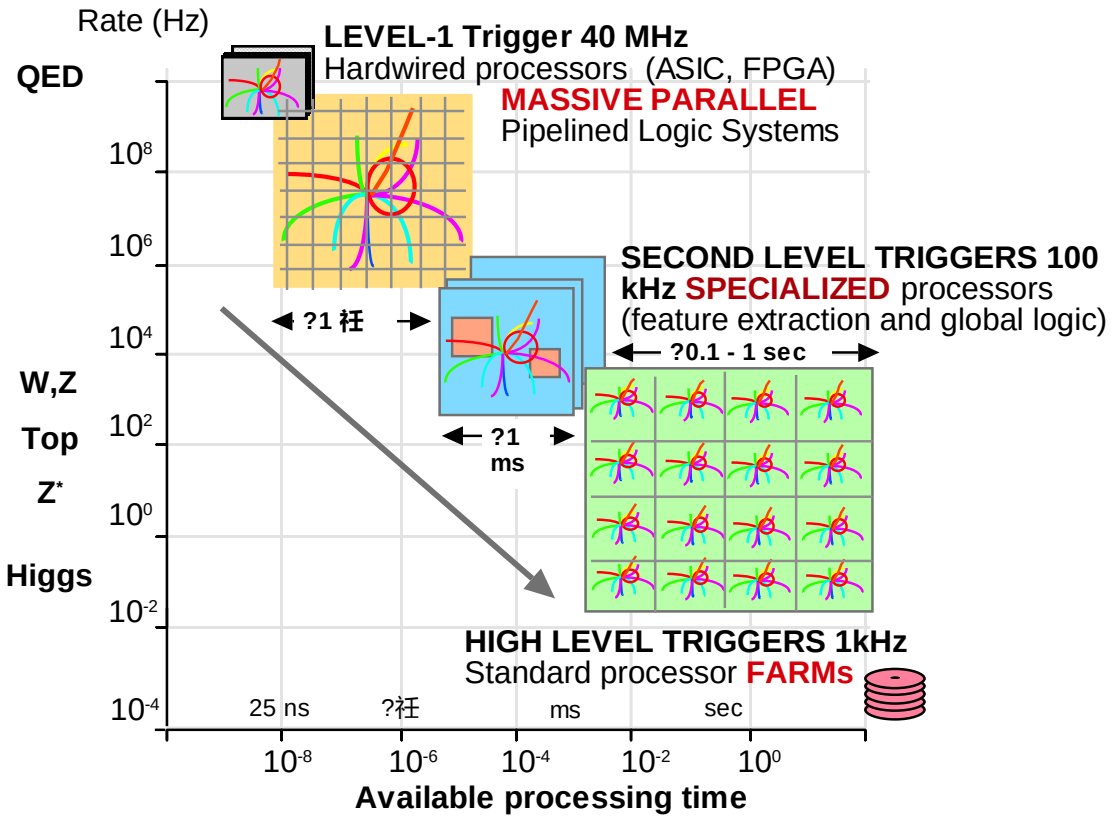
- Additional processing in LV-2: reduce network bandwidth requirements



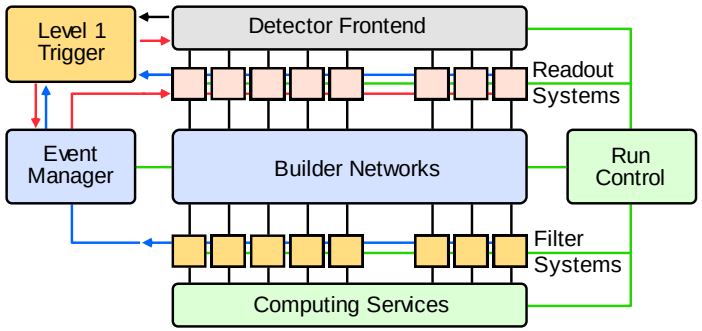


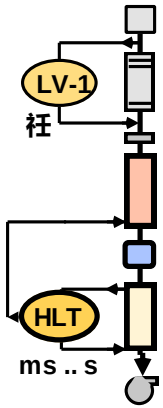
40 MHz
10^5 Hz
10^3 Hz
10 Gb/s

10^2 Hz



Two physical entities(CMS)



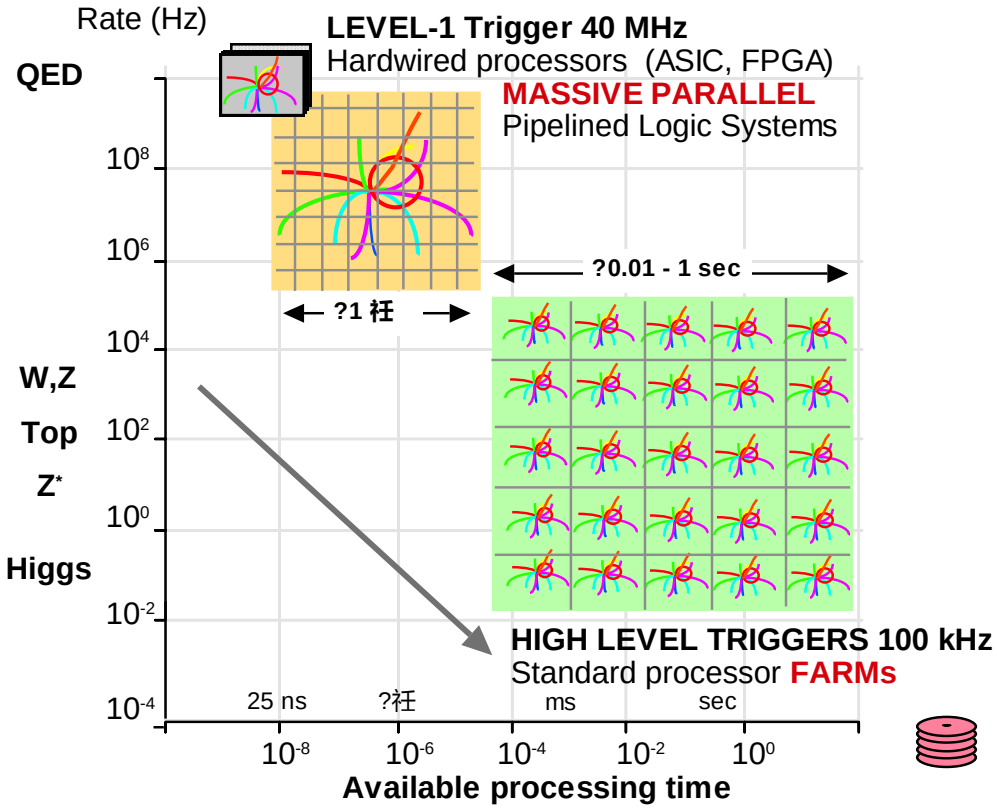


40 MHz

10^5 Hz

1000 Gb/s

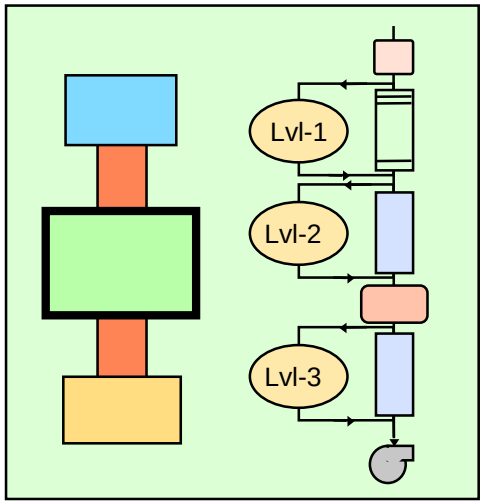
10^2 Hz



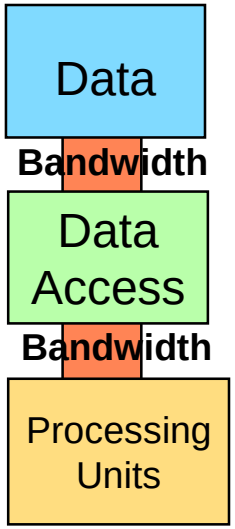
- Reduce number of building blocks
- Rely on commercial components (especially processing and communications)

Comparison of 2 vs 3 physical levels

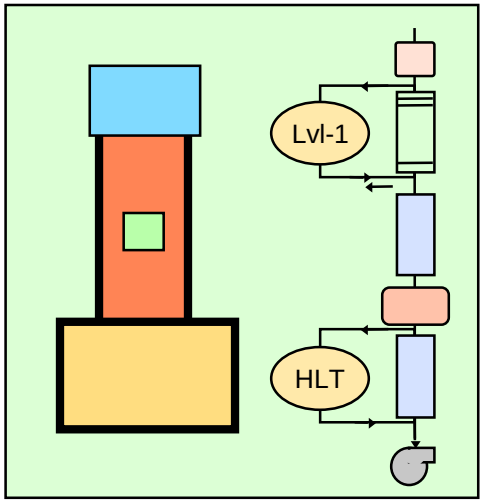
- Three Physical Levels
  - Investment in:
    - Control Logic
    - Specialized processors



Model



- Two Physical Levels
  - Investment in:
    - Bandwidth
    - Commercial Processors

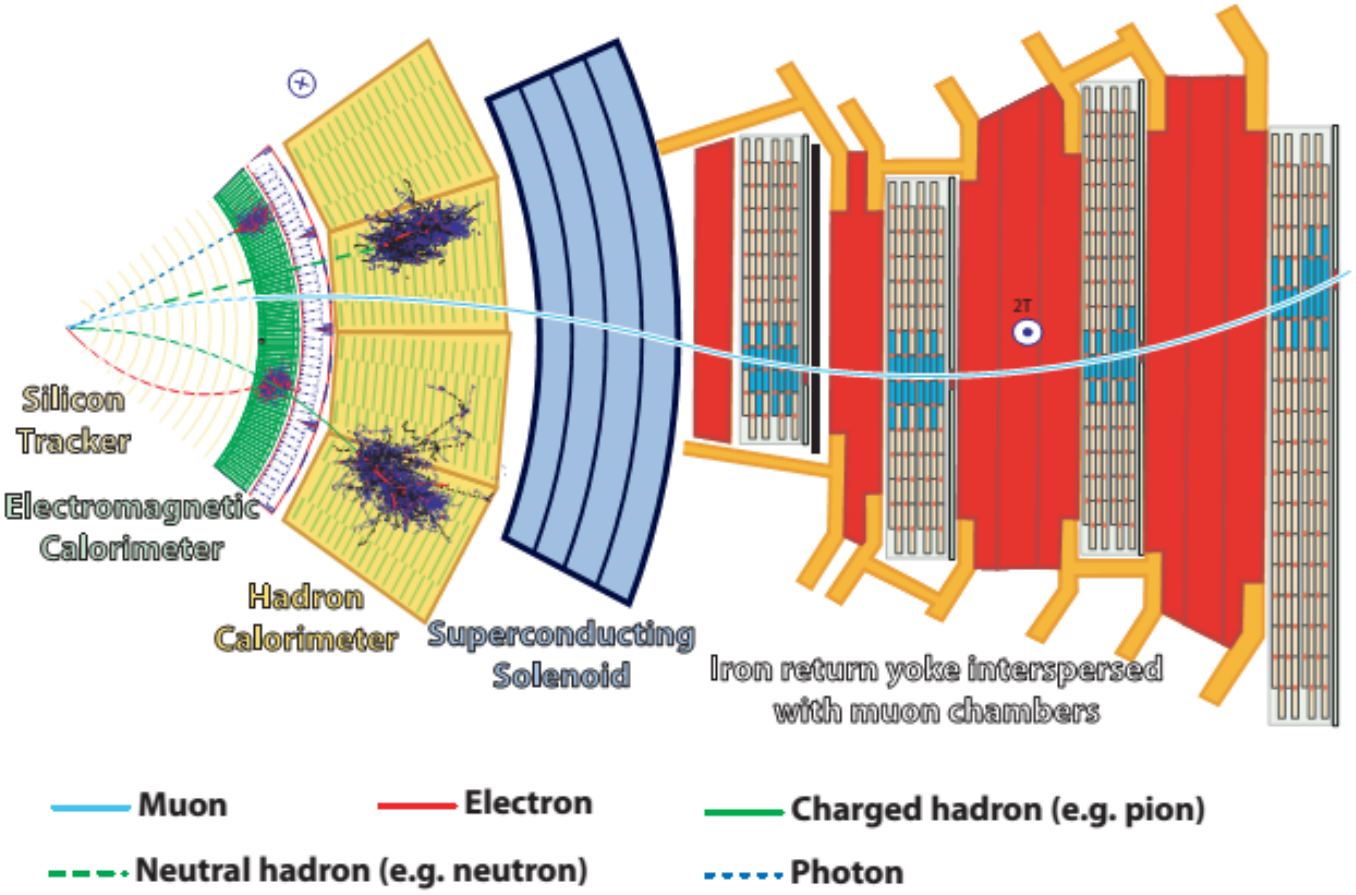


Technologies in Level-1 systems

- ASIC(Application-Specific Integrated Circuit) used in some cases
  - Highest-performance option, better radiation tolerance and lower power consumption (a plus for on-detector electronics)
- FPGA(Field-Programmable Gate Array) used in all systems
  - Impressive evolution with time. Large gate counts and operating at 40 MHz (and beyond)
  - Biggest advantage: flexibility
    - Can modify algorithms (and their parameters) in situ
- Communication technologies
  - High-speed serial links (copper or fiber)
    - LVDS up to 10 m and 400 Mb/s; HP G-link, Vitesse for longer distances and Gb/s transmission
  - Backplanes
    - Very large number of connections, multiplexing data
      - operating at ~160 Mb/s
  - High speed optical links (fibers)
    - Up to 10Gb/s per link

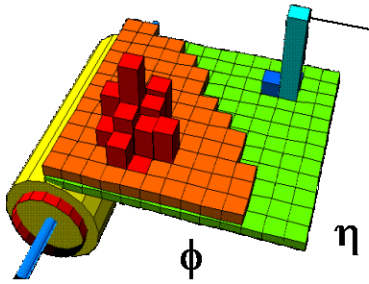
Transverse slice through CMS detector

different features of different particles



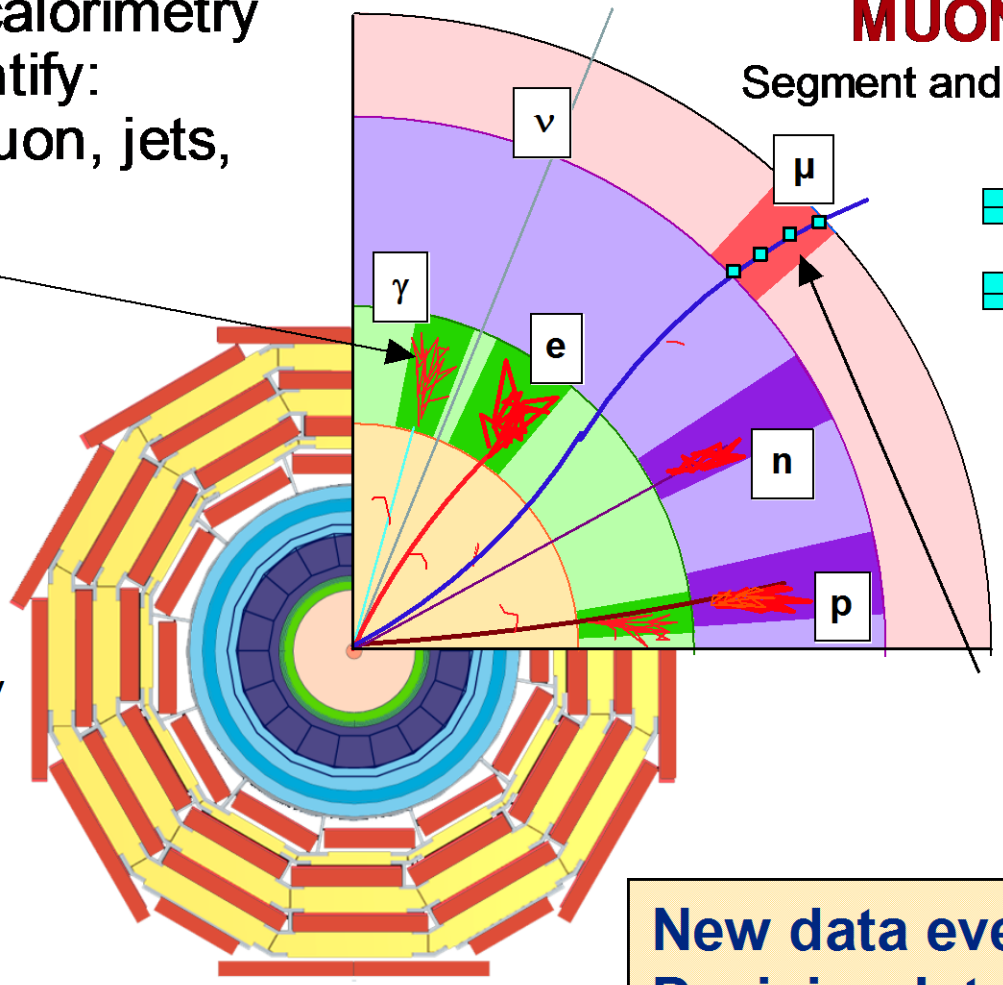
Particle signatures in the detector

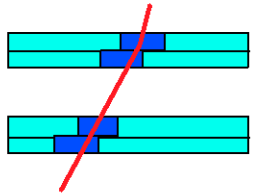
Use prompt data (calorimetry and muons) to identify:
High p_t electron, muon, jets, missing E_T



CALORIMETERs
Cluster finding and energy deposition evaluation

MUON System
Segment and track finding

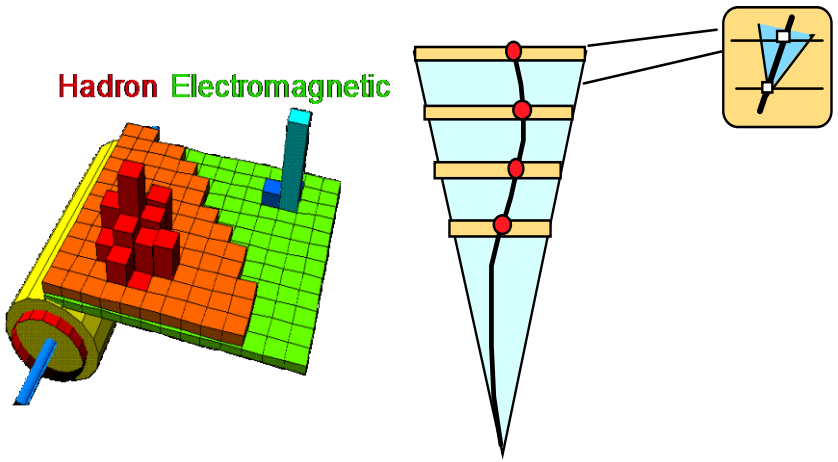




New data every 25 ns
Decision latency ~ 衽

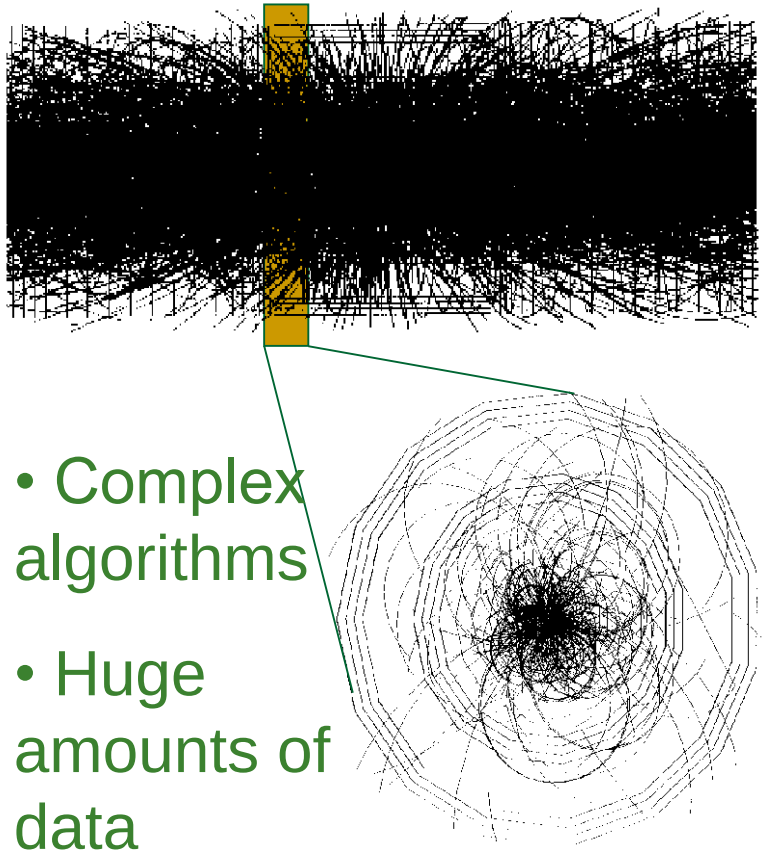
At Level-1: only calo and muon info

- Pattern recognition much faster/easier



- **Simple algorithms**
- **Small amounts of data**
- **Local decisions**

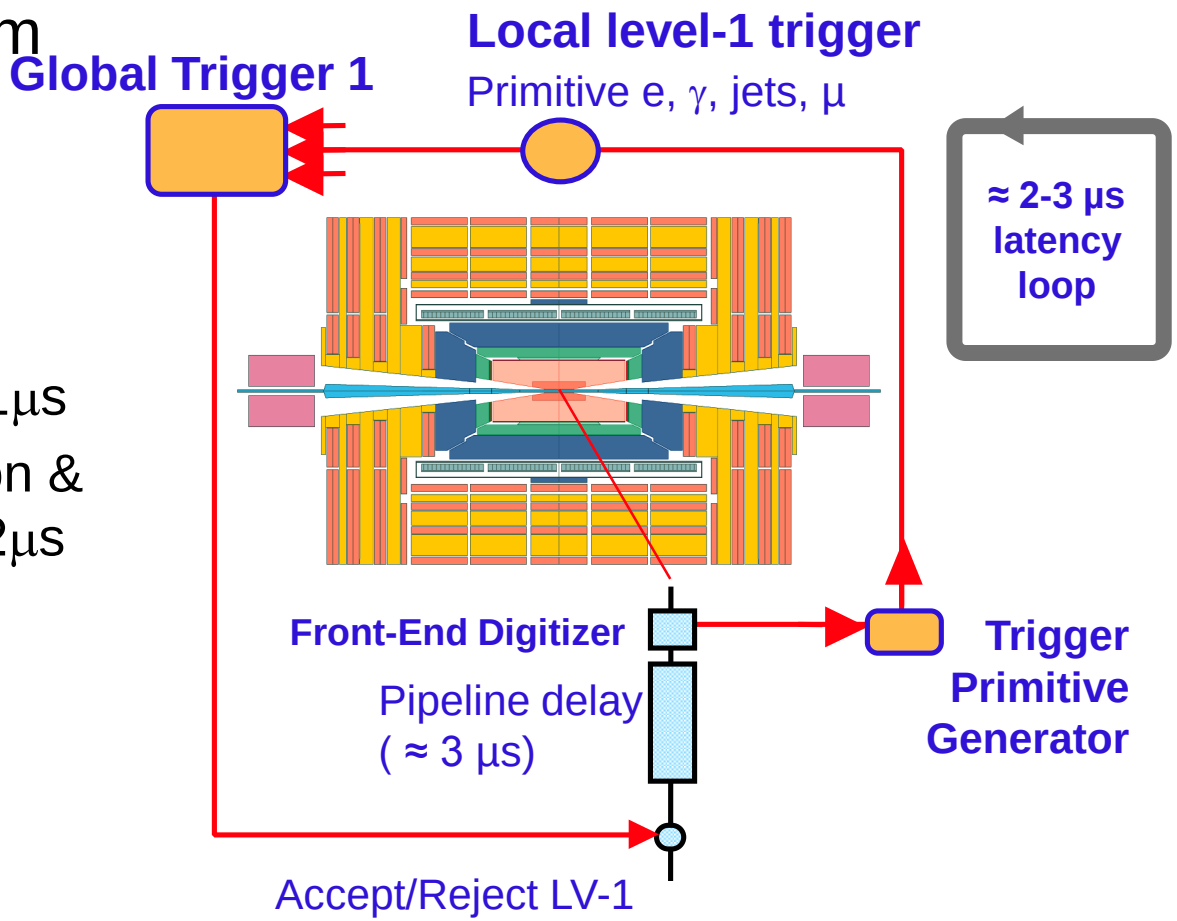
- Compare to tracker info



- Complex algorithms
- Huge amounts of data
- Need to link sub-detectors

Level-1 Trigger: decision loop

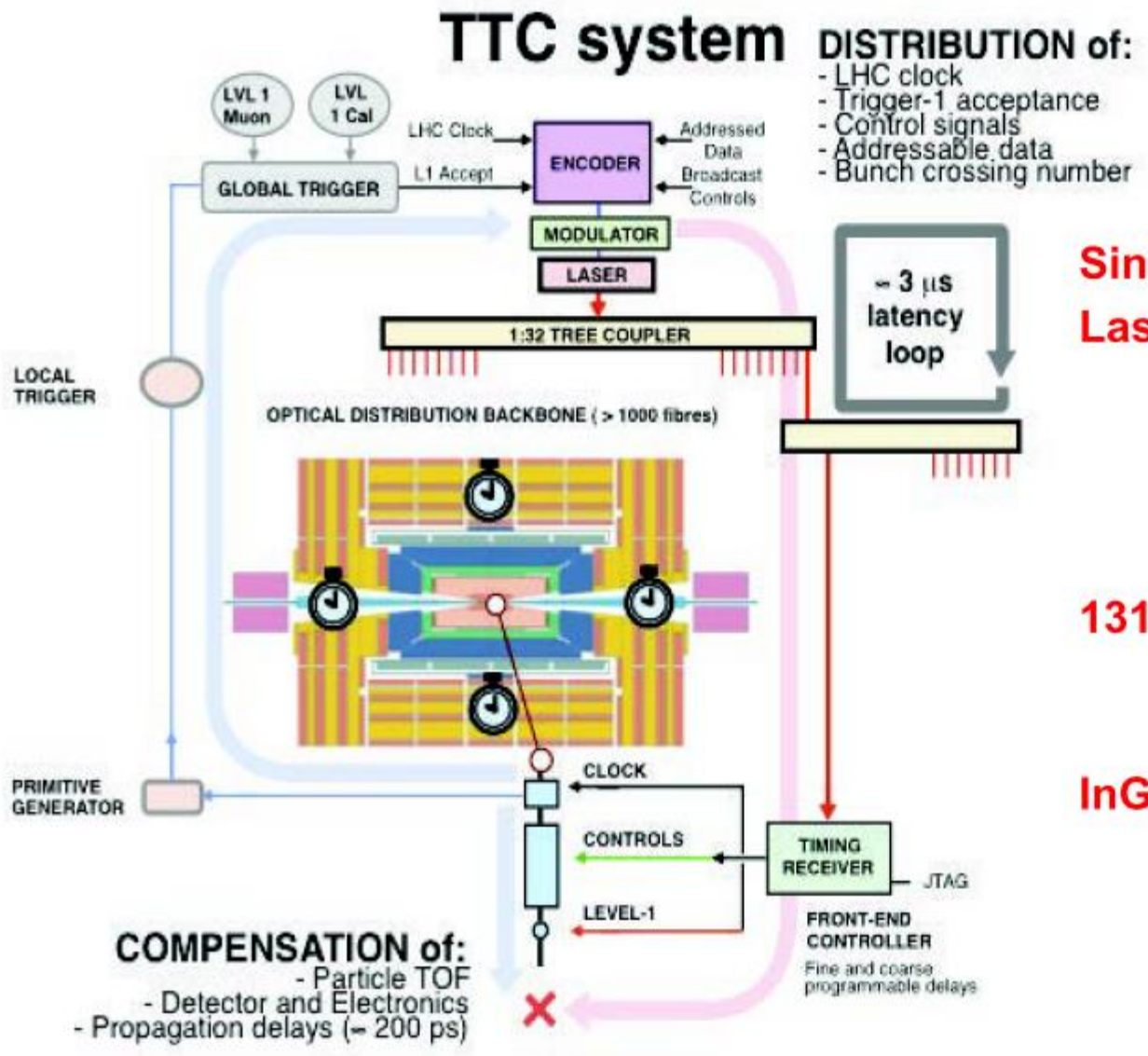
- Synchronous 40 MHz digital system
  - Typical: 160 MHz internal pipeline
  - Latencies:
    - Readout + processing: < 1μs
    - Signal collection & distribution: ≈ 2μs
- At Lvl-1: process only calo+μ info



Global Trigger

- A very large OR-AND network that allows for the specification of complex conditions:
  - 1 electron with P_T>20 GeV OR 2 electrons with P_T>14 GeV OR 1 electron with P_T>16 and one jet with P_T>40 GeV…
  - The top-level logic requirements (e.g. 2 electrons) constitute the "trigger-table" of the experiment
    - Allocating this rate is a complex process that involves the optimization of physics efficiencies vs backgrounds, rates and machine conditions

Trigger timing and control



Optical System:

Single High-Power Laser per zone

- Reliability, transmitter upgrades
- Passive optical coupler fanout

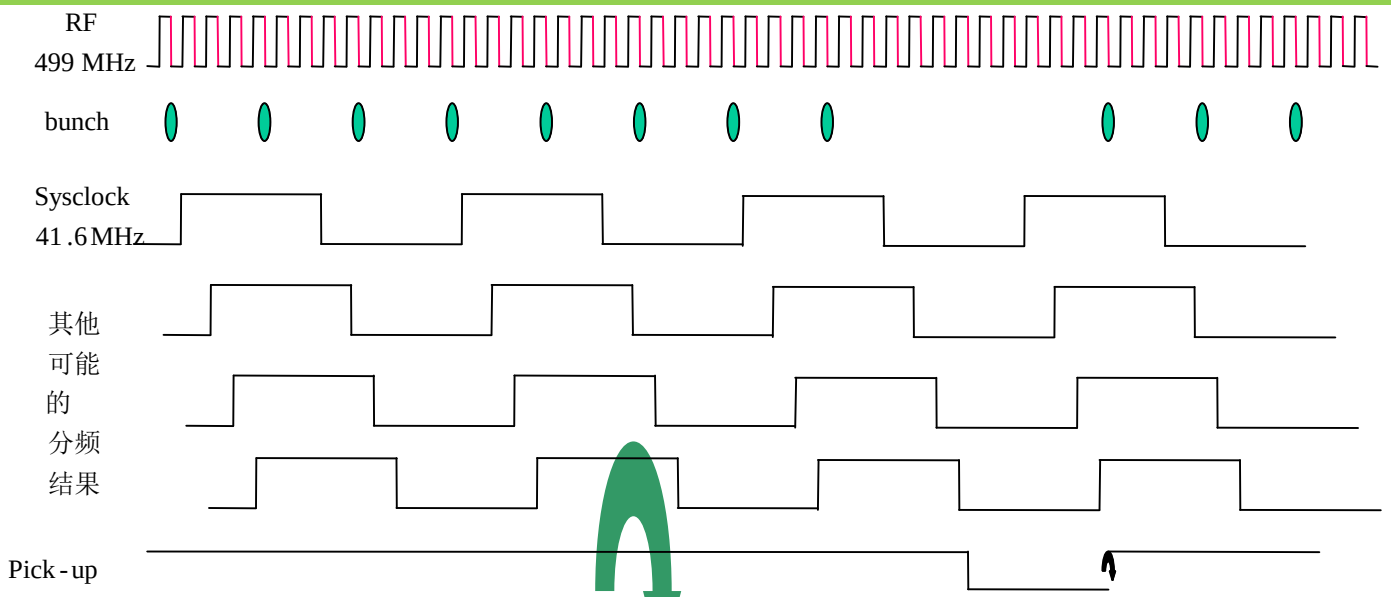
1310 nm Operation

- Negligible chromatic dispersion

InGaAs photodiodes

- Radiation resistance, low bias

System Clock and Fast Control



GLT

L1, Typ0,Typ1,Busy

To trigger

FCS

FCTL

FCSF

TROC

Pickup

Sync

RF
clock

Clock
Generator

optical

CLKF

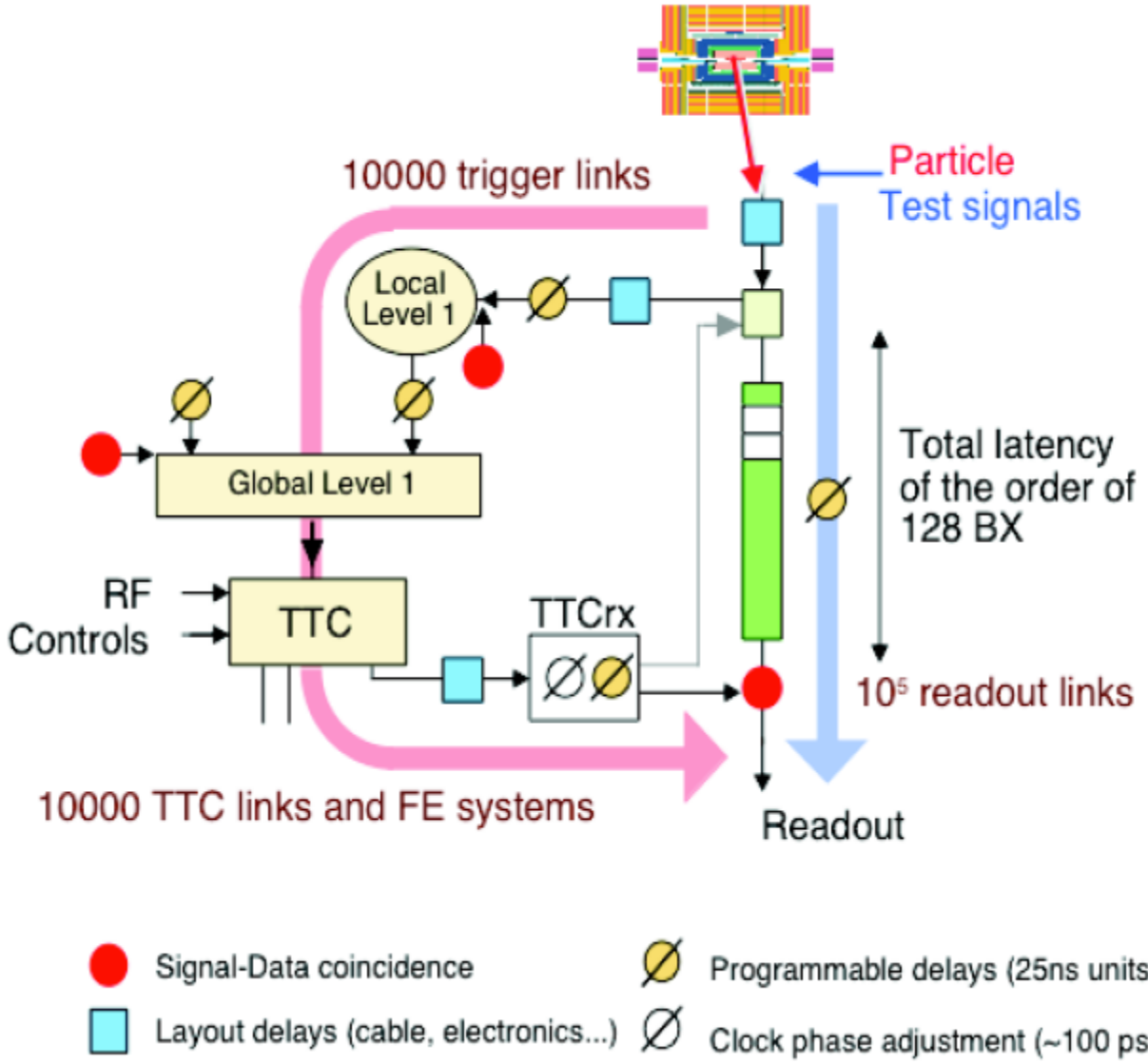
FCDB

Fast control

TOFElectronics

To electronics

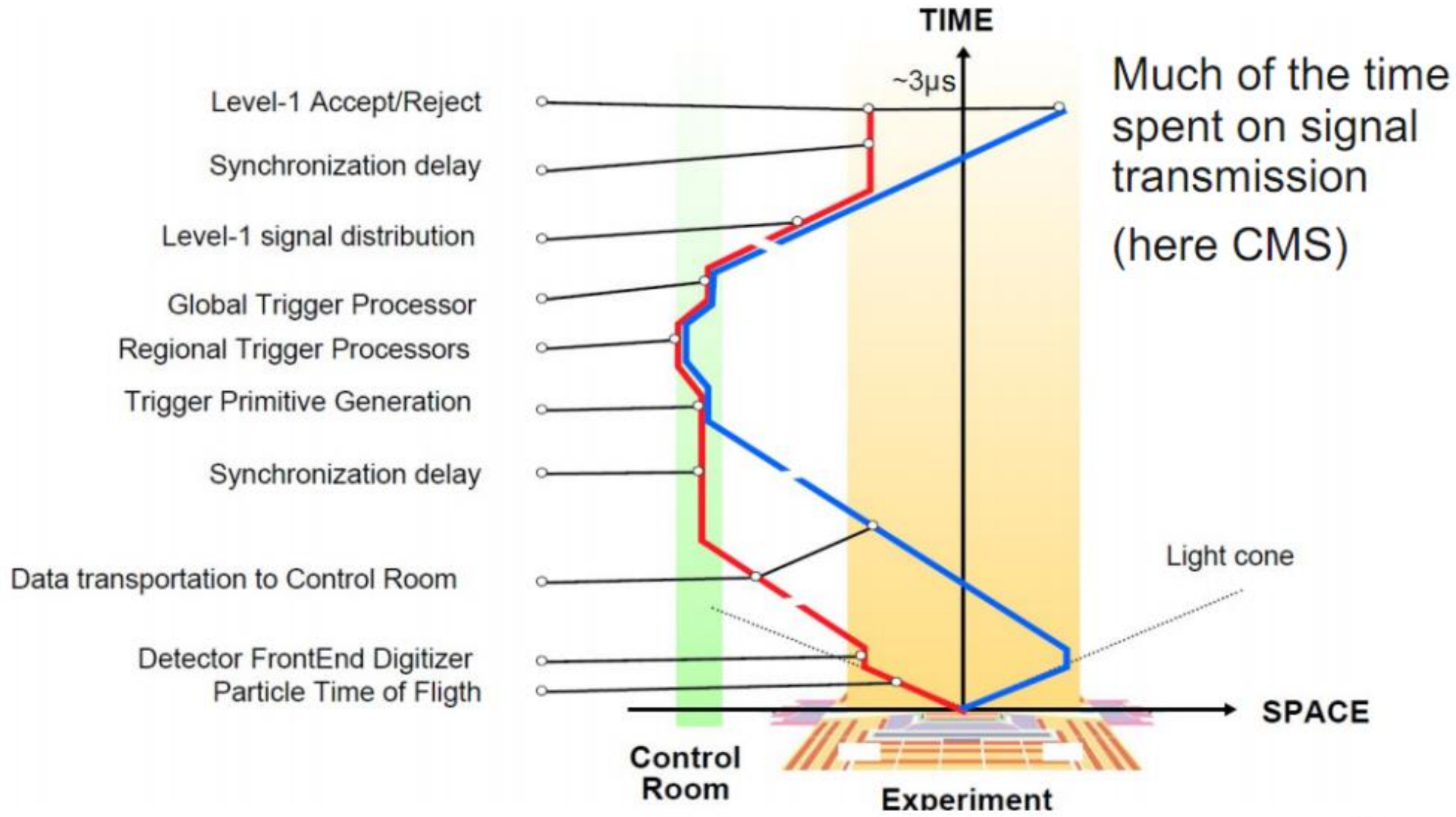
Align the signal



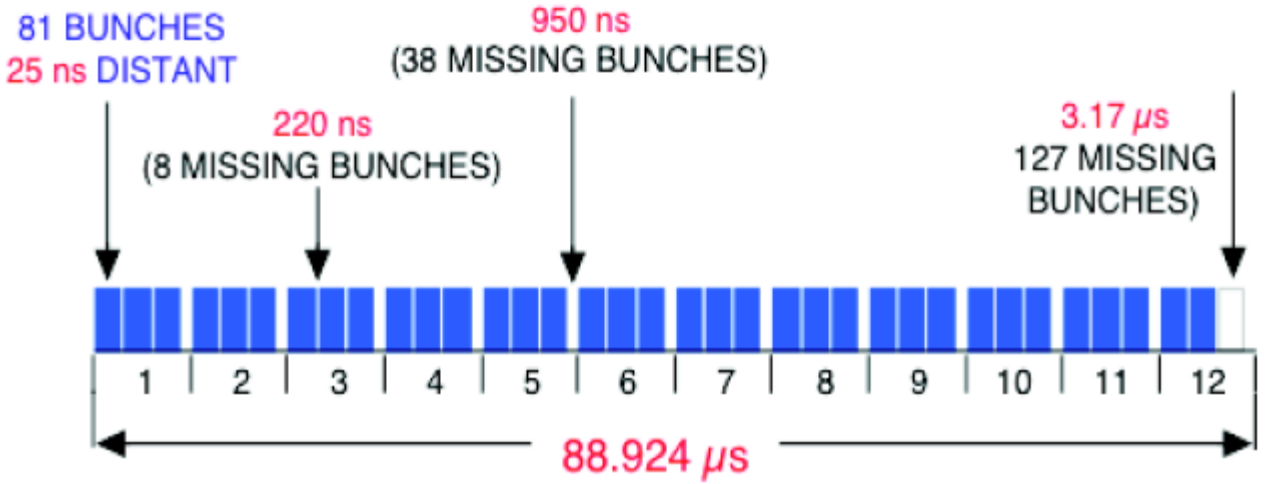
Need to Align:

- Detector pulse w/collision at IP
- Trigger data w/ readout data
- Different detector trigger data w/each other
- Bunch Crossing Number
- Level 1 Accept Number

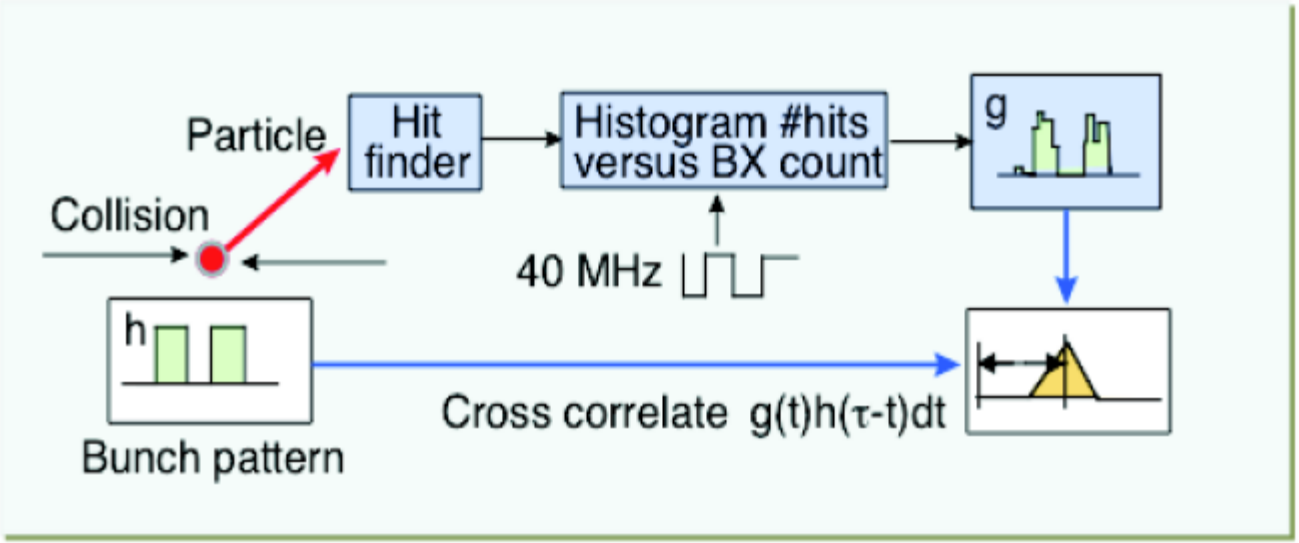
Trigger Latency



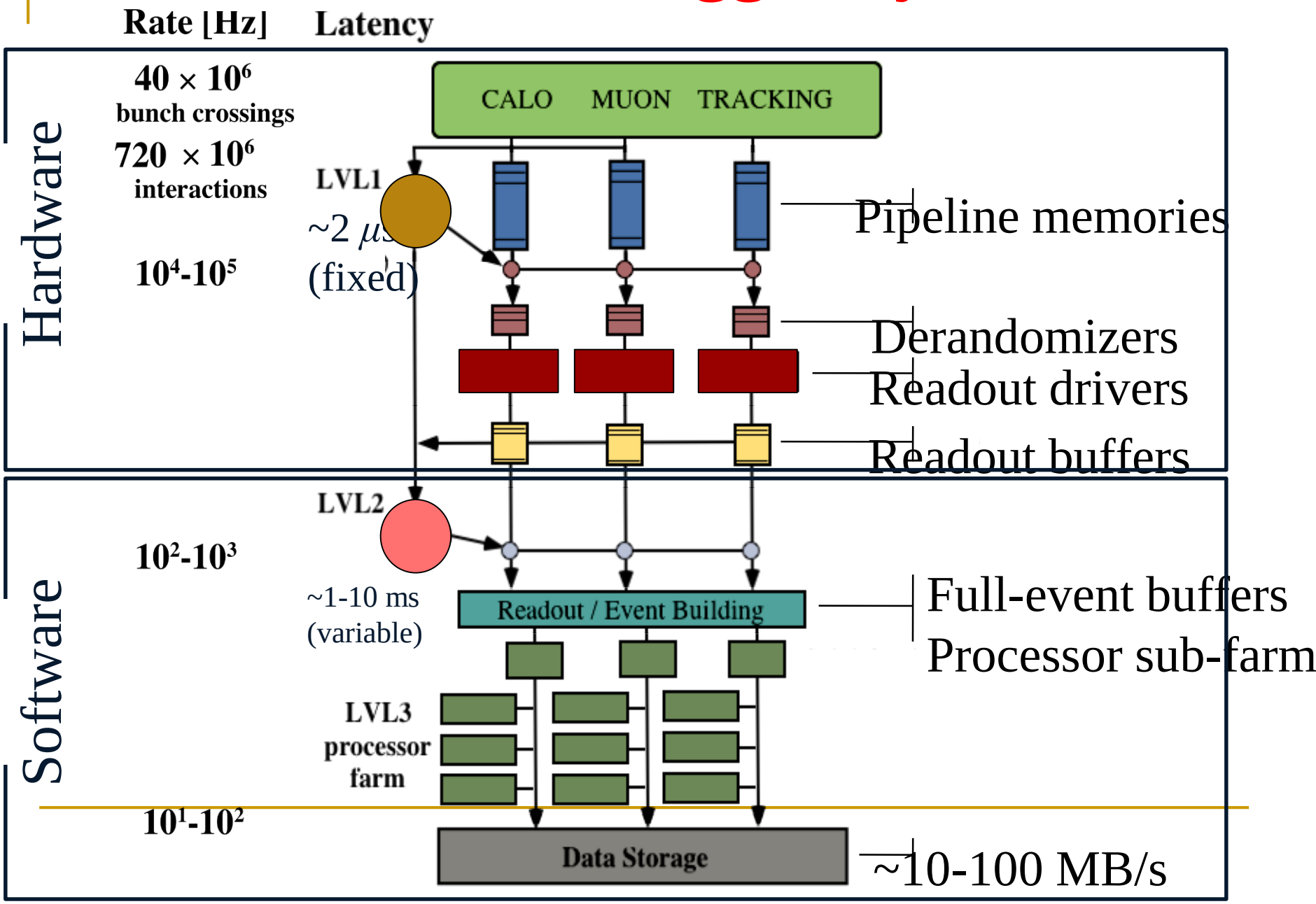
Synchronization



2835 out of 3564 p bunches are full, use this pattern:

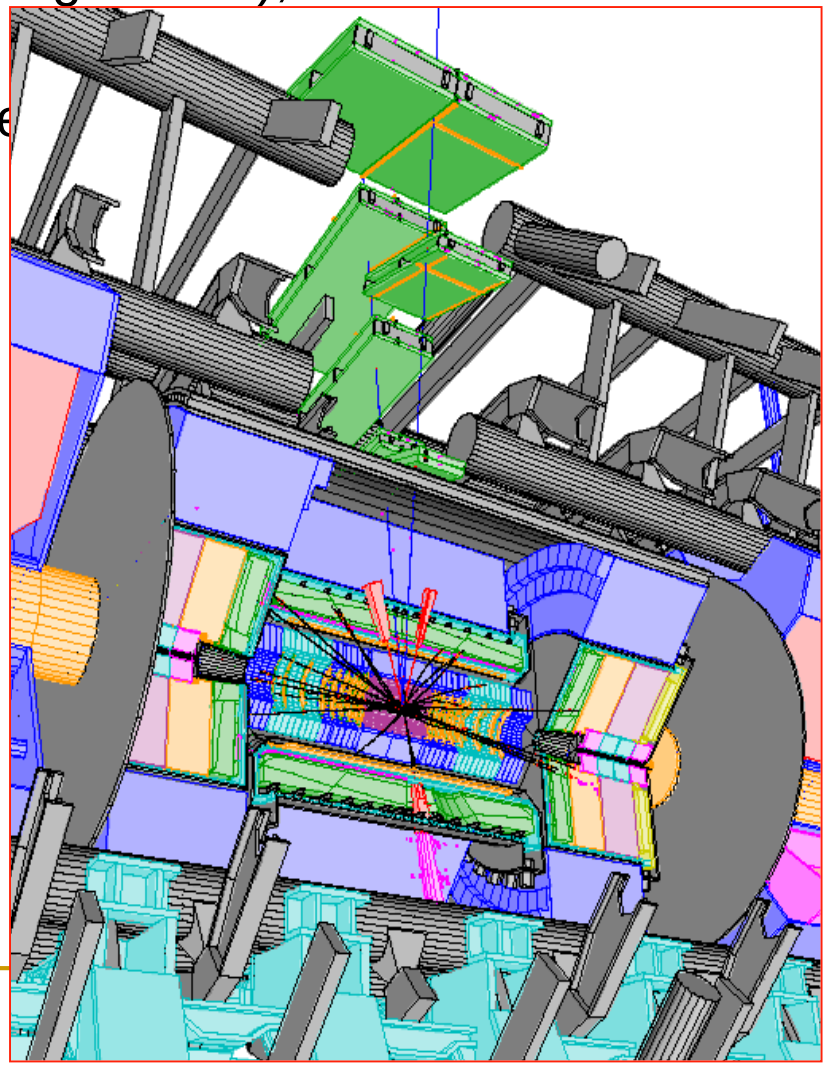


The ATLAS Trigger System

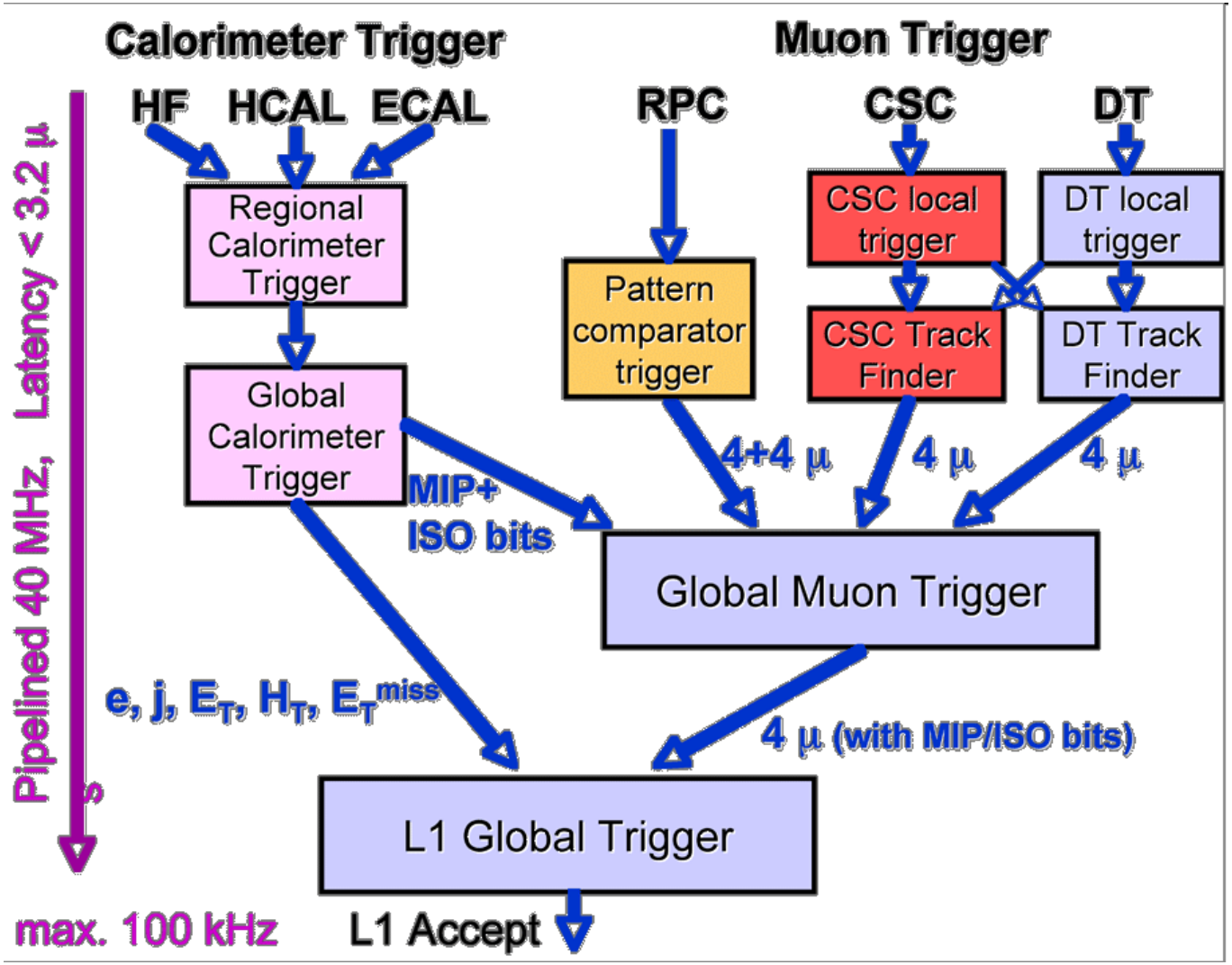


ATLAS: the Region of Interest - Why?

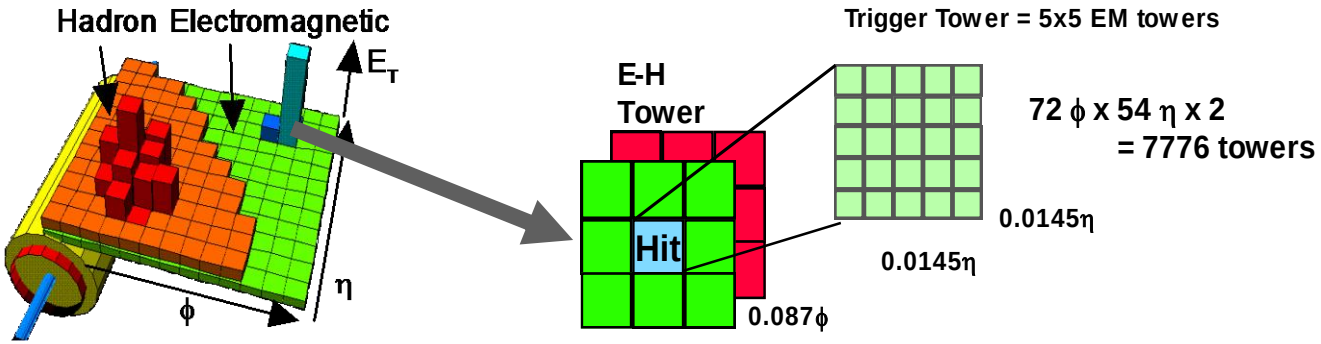
- The Level-1 selection is dominated by local signatures
  - Based on coarse granularity (calo, mu trig chamb), **w/out access to inner tracking**
  - Important further rejection can be gaine with local analysis of full detector data
- The geographical addresses of interesting signatures identified by the LVL1 (Regions of Interest)
  - Allow access to local data of each relevant detector
  - Sequentially
- Typically, there are less than 2 RoIs per event accepted by LVL1
  - <RoIs/ev> = ~1.6
- The resulting total amount of RoI data is minimal
  - a few % of the Level-1 throughput



Overview of CMS L1 Trigger



Lvl-1 Calo Trigger: e/γ algorithm (CMS)



E_T( ) + max E_T( ) > E_T^{min}

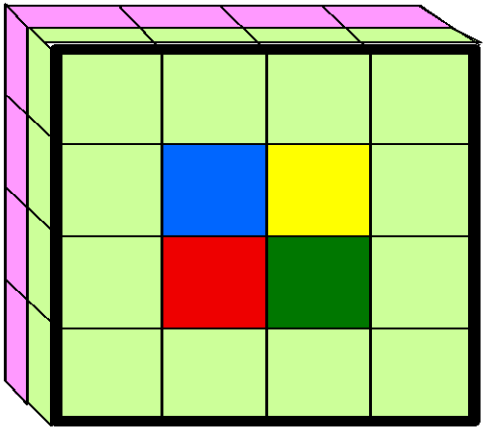
E_T( ) / E_T( ) < HoE^{max}

At least 1 E_T(, , ,) < E_{iso}^{max}

Fine-grain: ≥1( ) > R E_T^{min}

→ Isolated "e/γ"

ATLAS em cluster trigger algorithm



E.M. calorimeter

Hadronic calorimete

$\Delta\eta \times \Delta\phi \approx 0.1 \times 0.$

> E.M. cluster threshold

AND

< E.M. isolation threshold

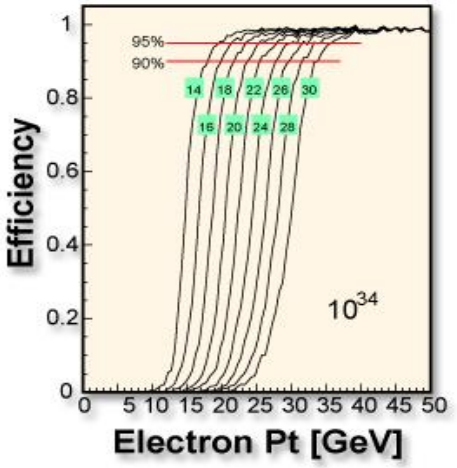
AND

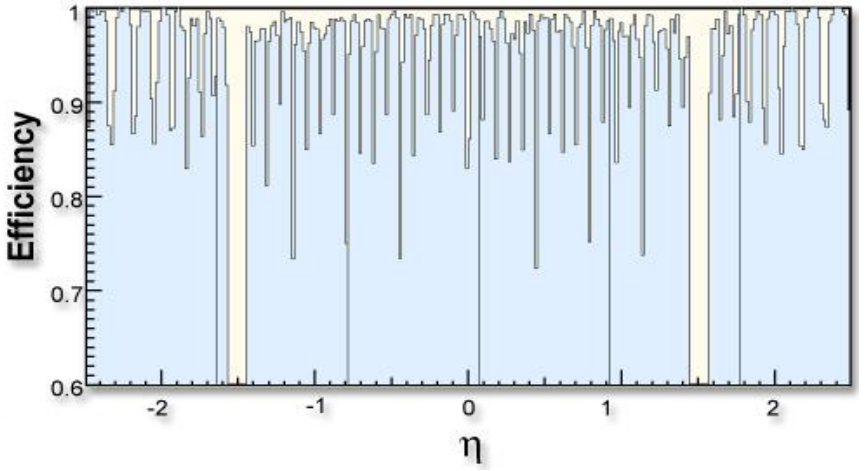
< Hadronic isolation thresh

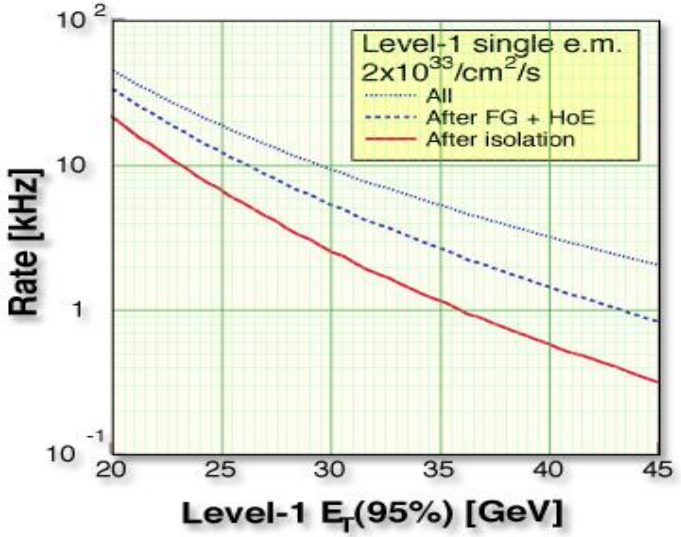
"Sliding window" algorithm repeated for each of ~4000 cells

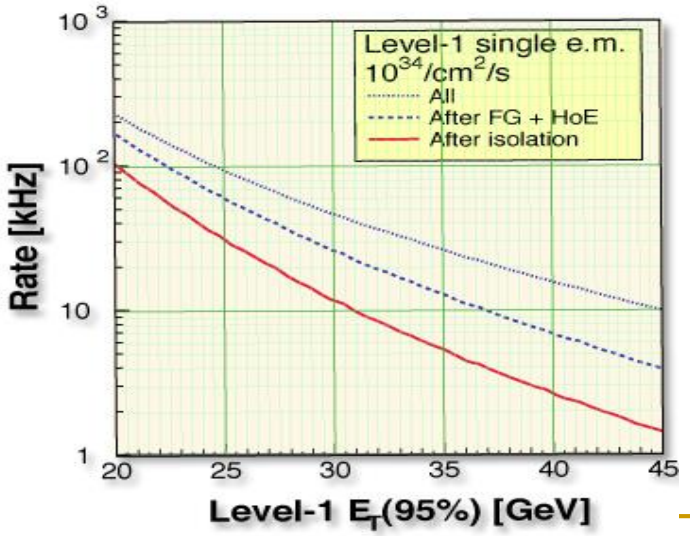
Lvl-1 Calo e/γ trigger: performance

- Efficiencies and Trigger Rates



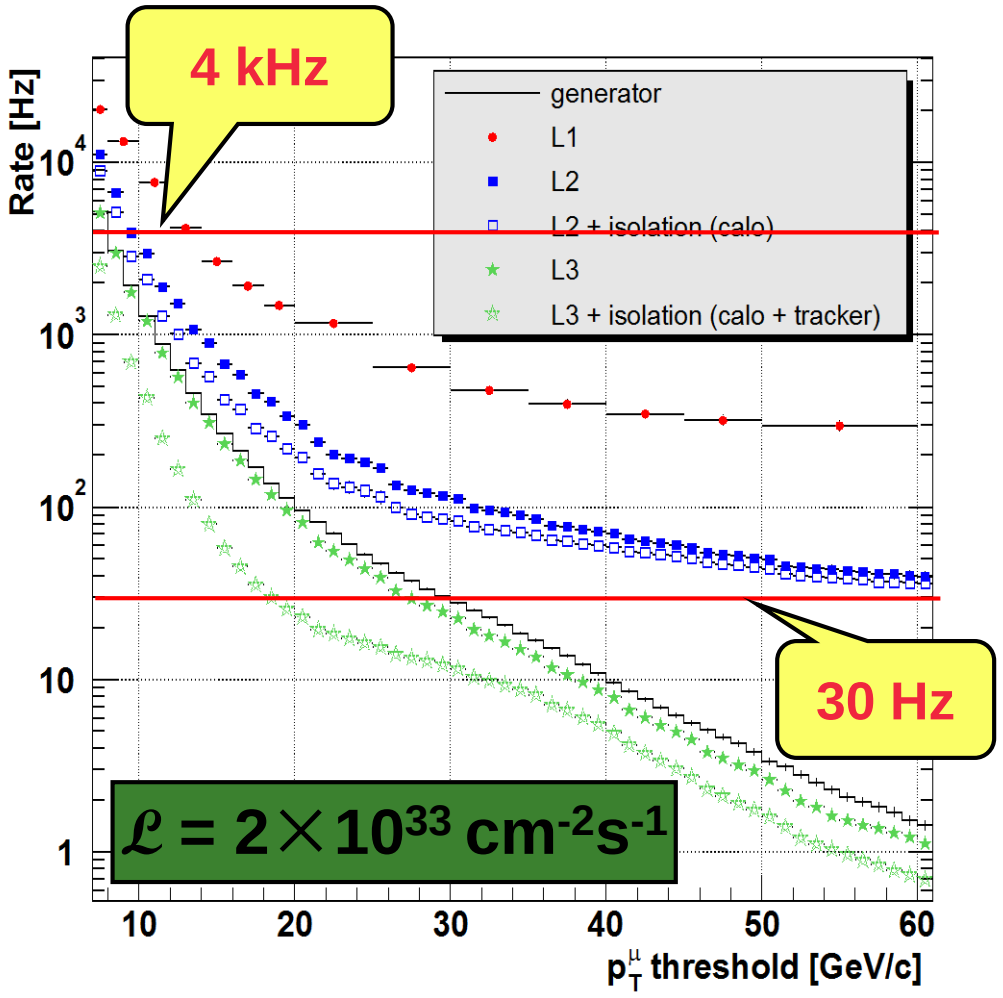






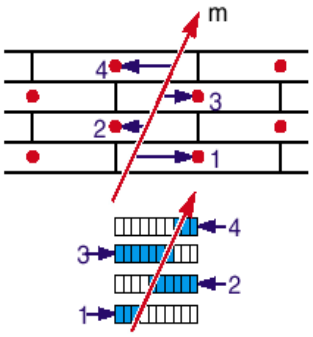
Lvl-1 muon trigger

- The goal: measure momentum online
  - Steeply falling spectrum; resolution costs!
- The issue: speed
  - CMS: RPC added to DT and CSC (which provide standalone trigger)

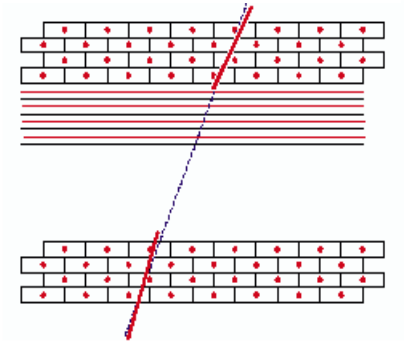


Lvl-1 muon trigger (CMS)

Drift Tubes

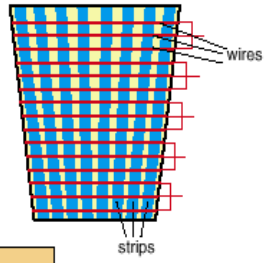


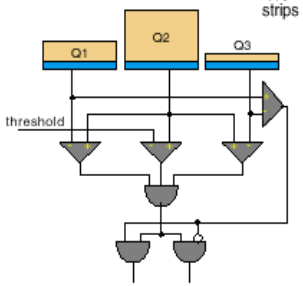
Meantimers recognize tracks and form vector / quartet.



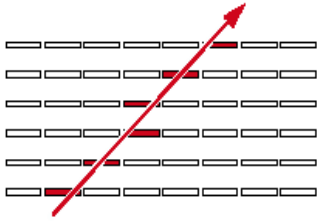
Correlator combines them into one vector / station.

CSC

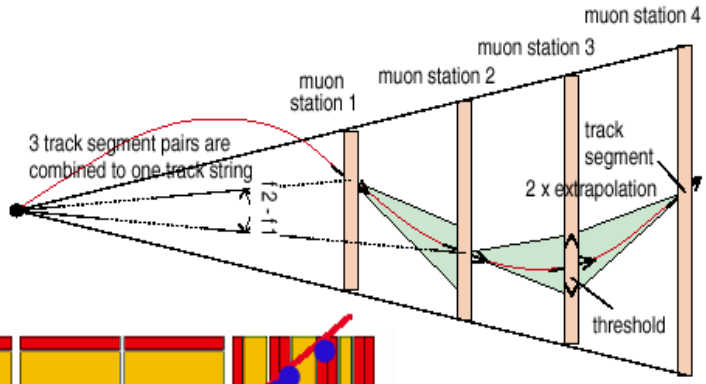


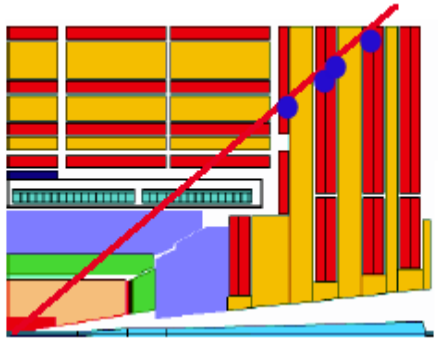


Comparators give 1/2-strip resol.



Hit strips of 6 layers form a vector.





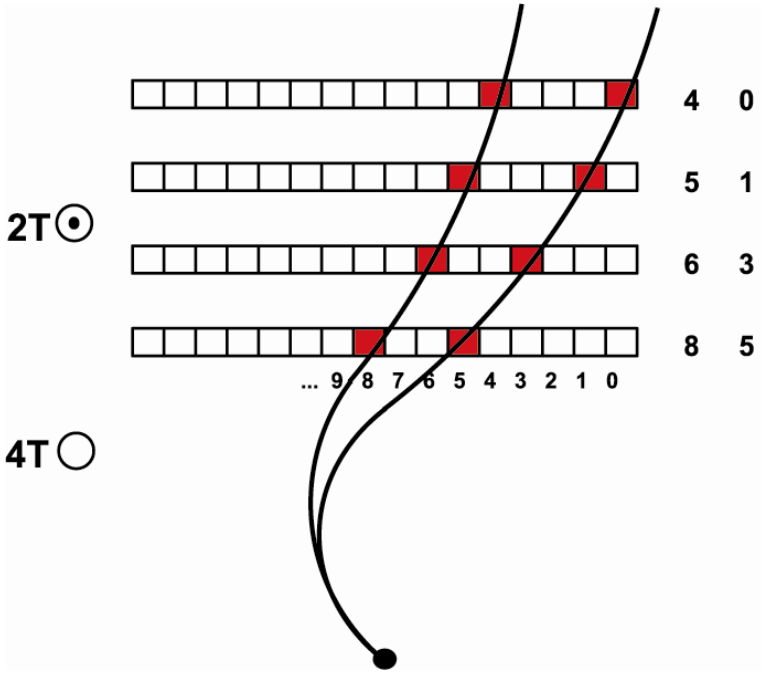
- Extrapolation: using look-up tables
- Track Assembler: link track segment-pairs to tracks, cancel fakes
- Assignment: P_T (5 bits), charge, η (6 bits) , φ(8 bits), quality (3 bits)

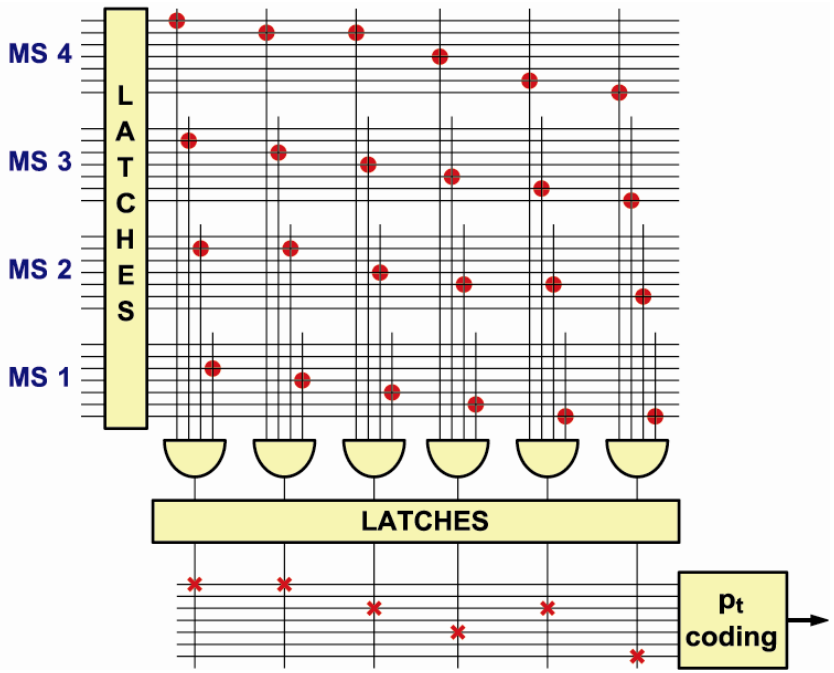
Hardware implementation:

ASICs for Trigger Primitive Generators

FPGAs for Track Finder processors

Lvl-1 muon trigger (CMS)





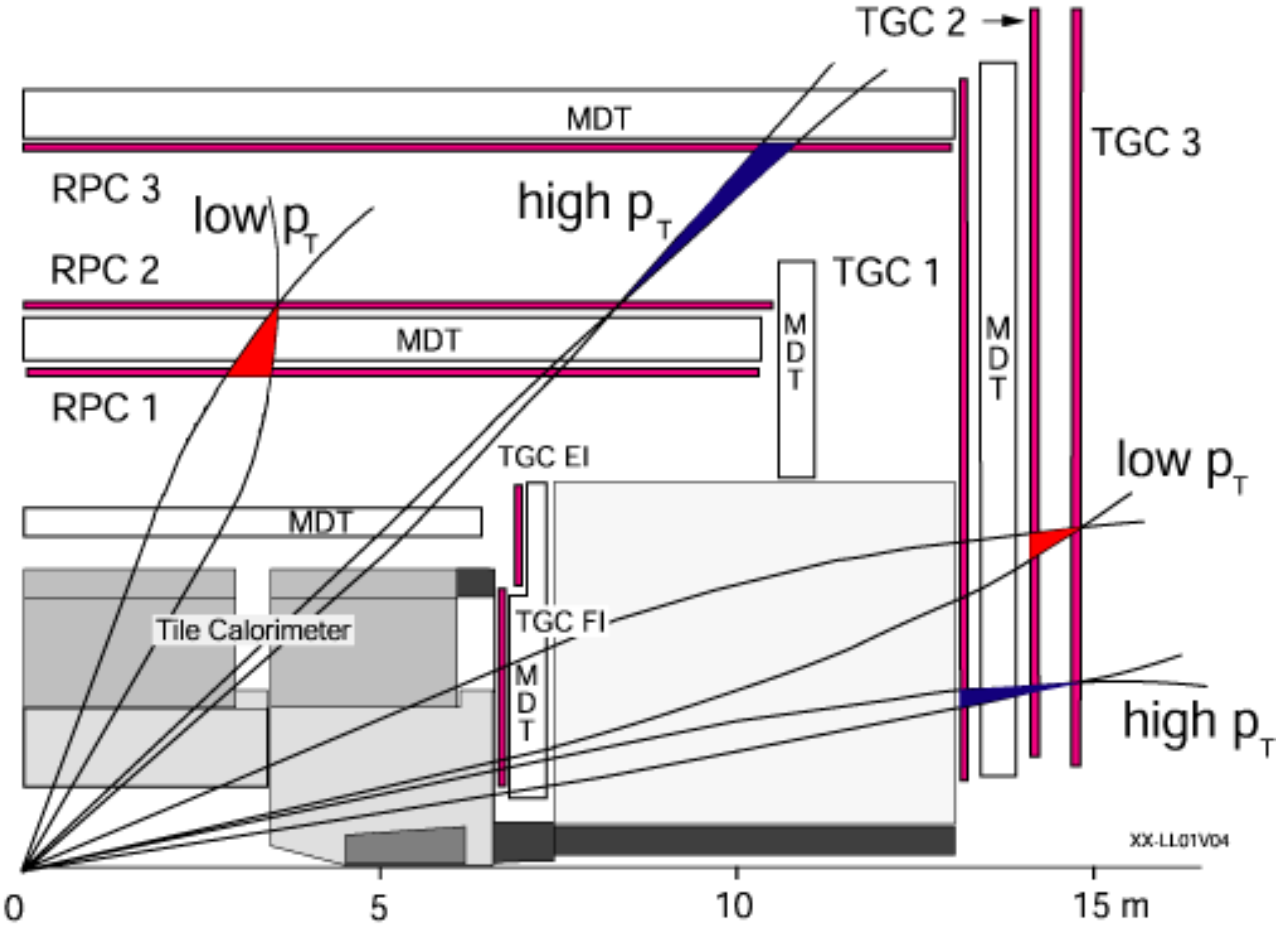
Pattern of strips hit:
Compared to predefined patterns corresponding to various p_T

Implemented in FPGAs

The LVL1 Muon Trigger（ATLAS）

- Safe Bunch Crossing Identification
- Wide p_T-threshold range
- Strong rejection of fake muons (induced by noise and physics background)

→Fast and high redundancy system

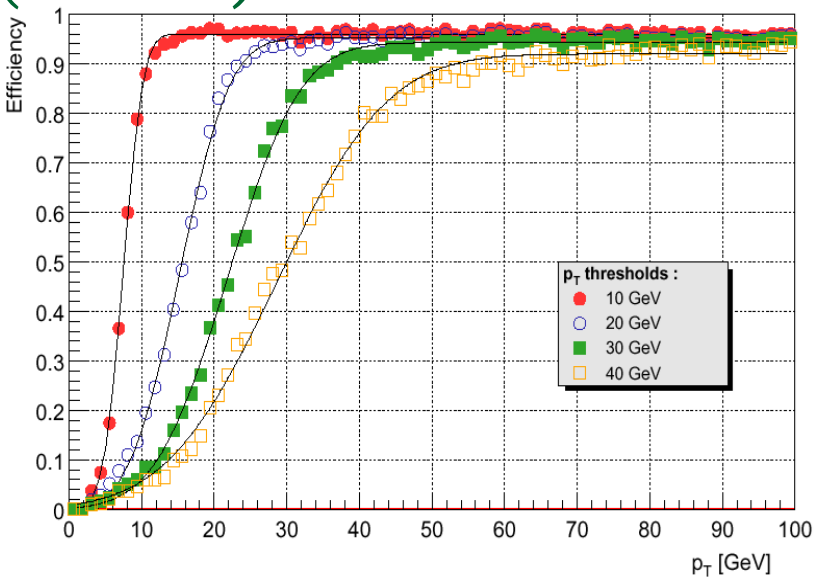


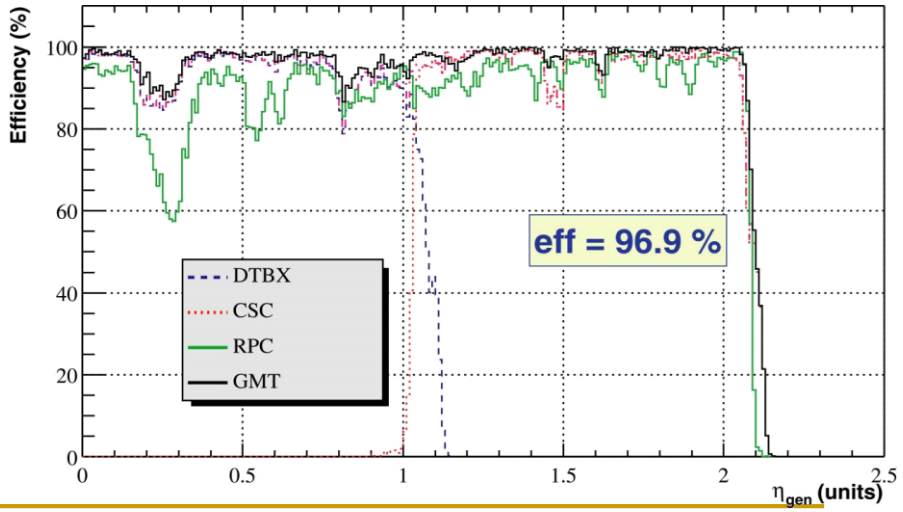
However this system:

1. Looks only for tracks coming from the pp collision point
2. Looks only for ultrarelativistic tracks

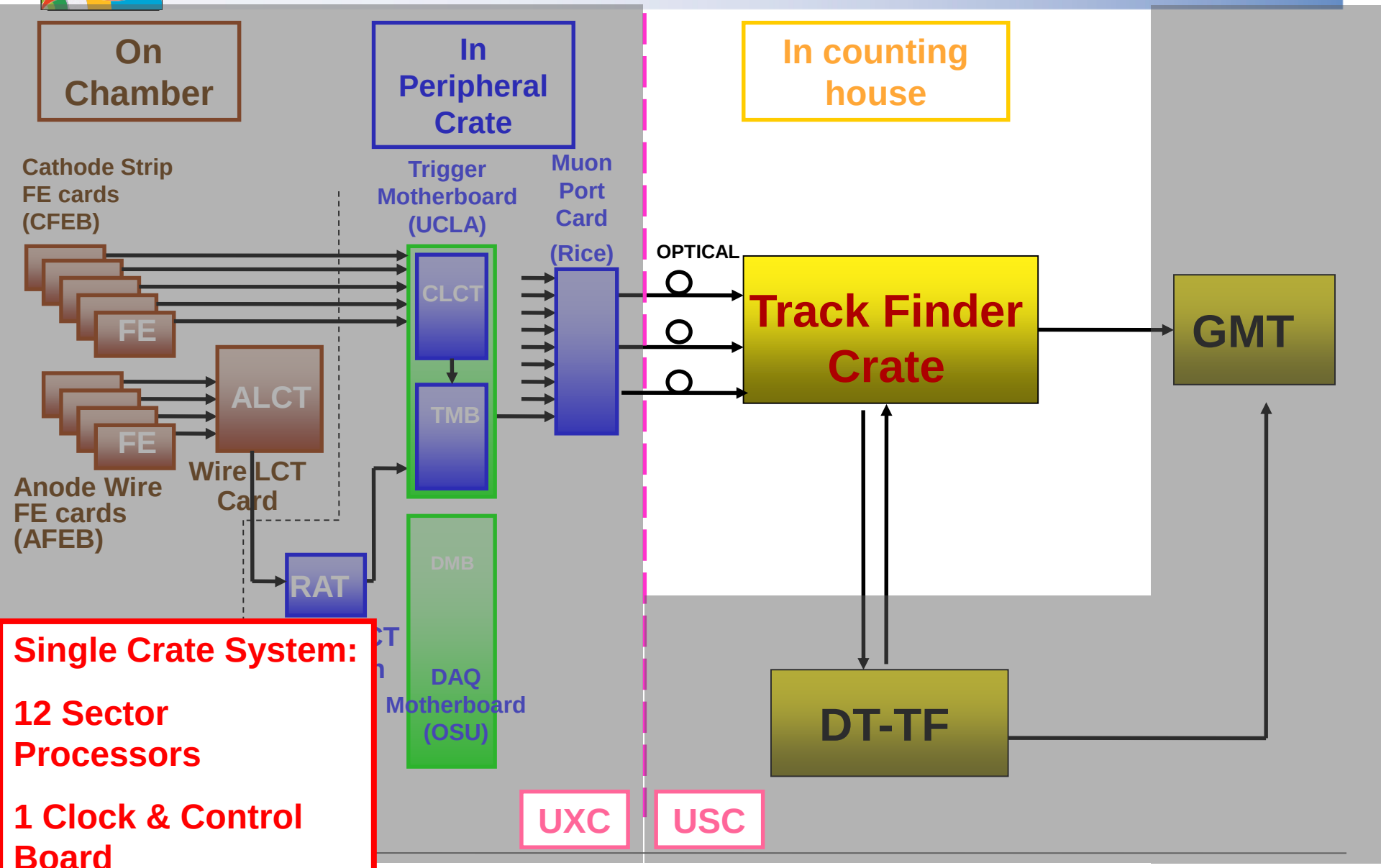
Global muon trigger (CMS)

- Combine results from RPC, CSC and DT triggers
- Match muon candidates from different trigger systems; use complementarity of detectors
- improve efficiency and rate
- assign muon isolation
- deliver the 4 best (highest P_T, highest-quality) muons to Global Trigger
- Pt resolution:
  - 18% barrel
  - 35% endcaps
- Efficiency: ~ 97%





CSC Track Finder: overview



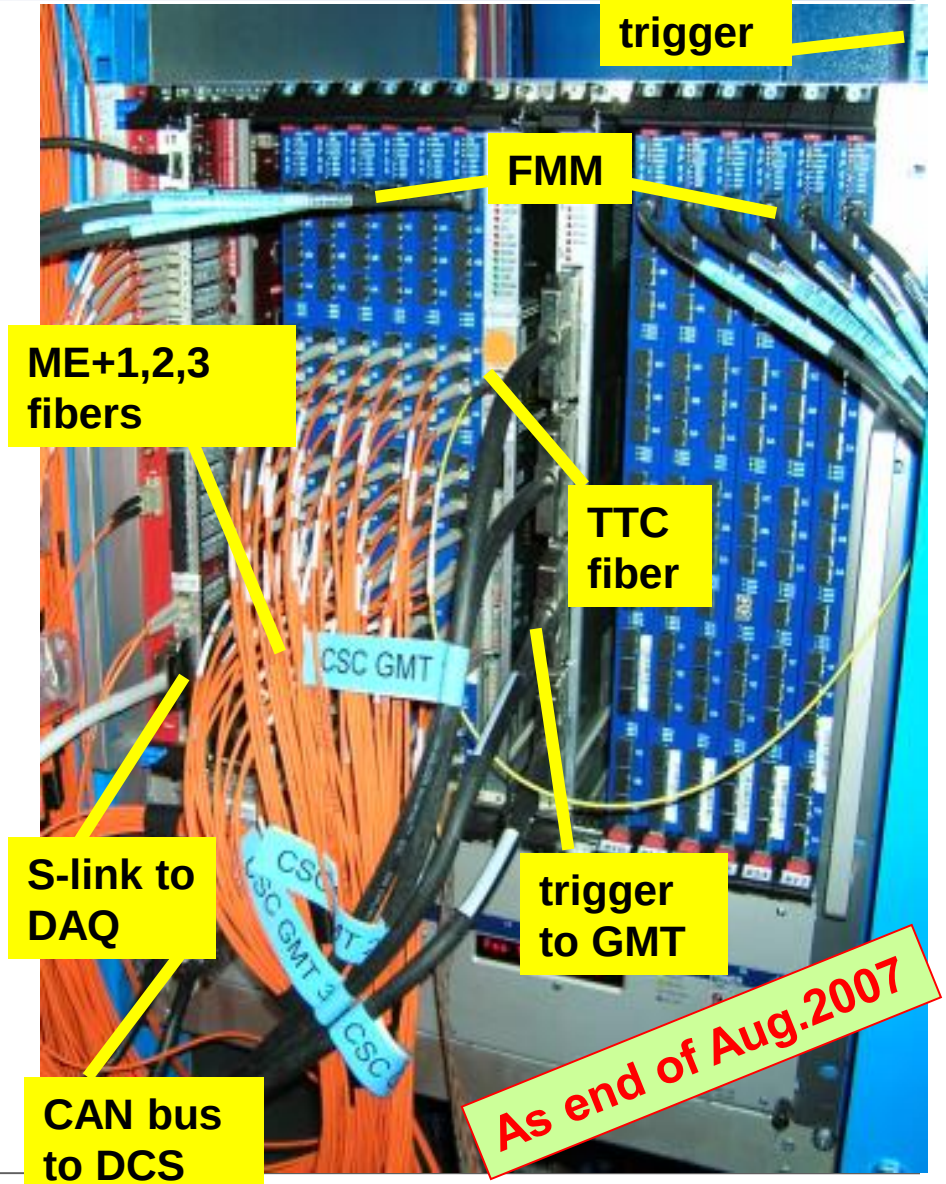
Single Crate System:

12 Sector Processors

1 Clock & Control Board

Hardware Layout

- **Hardware installed in USC S1D04**
  - single crate, filled
  - ME+1,2,3 fibers connected, waiting for fiber length&ME4
  - 13 FMM, DAQ s-link , GMT trigger cables connected
  - Cables to DTTF(72 scsi) connected, tested locally
  - Control PC & s/w installed & basic Hardware re-validation done
  - CSC TTC distribution validated



CSC-DT integration

Information sent out to DTTF

Variable	Function	bits / muon	bits / 2 muons
ϕ	azimuth coordinate	12	24
quality	quality	3	6
BXN	LSBs of bunch i.d.	–	2

Information received from DTTF (from one 30° sector)

Variable	Function	bits / muon
ϕ	azimuth coordinate	12
ϕ_b	ϕ bend angle	5
quality	quality	3
BXN	LSBs of bunch i.d.	2
Synch./Calib.	Special Mode	1
Flag bit	Denote if 2nd muon	1

CSCTF transition board and connection cables to DTTF with Channel-Link LVDS transmission

Mezzanine Card Upgrade

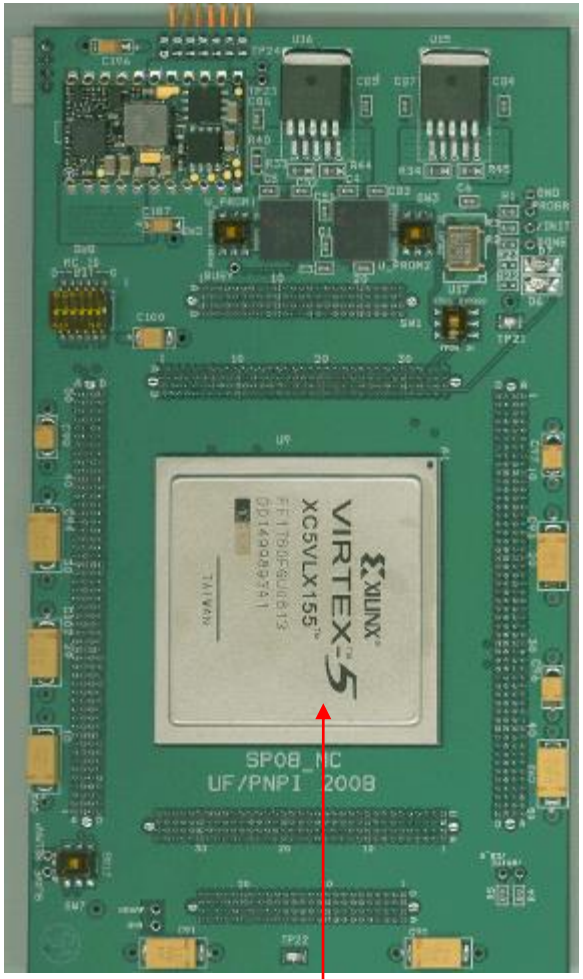
The new mezzanine card for the SP has three times more logic space

- Allows for more robustness in the trigger logic
- Provide more trigger flexibilities for ME4/2 chambers
- accommodate increased functionality for LHC triggering

Timelines

- Prototype mezzanine cards produced and tested successfully in B904 and P5 last November
- Production launched in January, and testing completed in Florida in February
- Data-taking test @ P5 in mid-Feb using prototype was successful
- 17 boards sent to CERN early March
- 4 boards installed and tested successfully in B904
- On Mar.15, upgraded all 12 SPs in the CSCTF crate to the new mezzanine cards

Functioning well since deployment. We will phase in the new features by and by



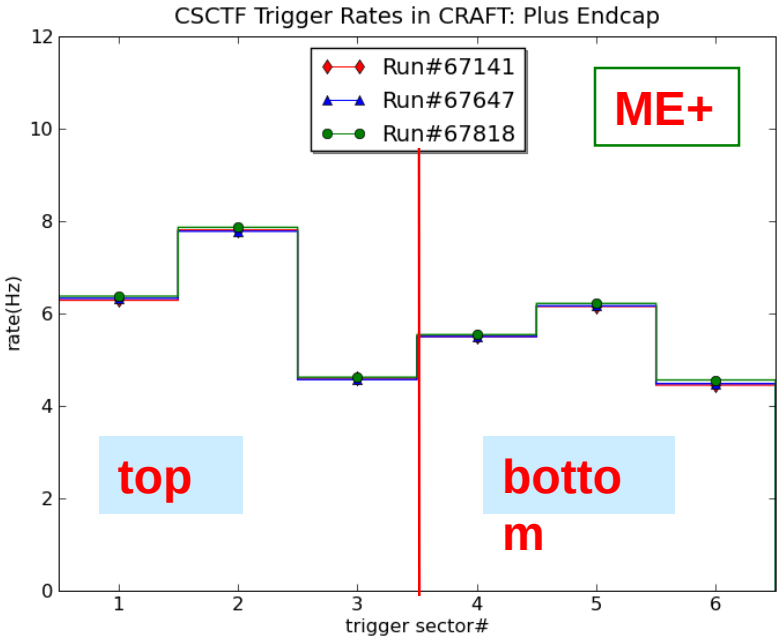
CSCTF Trigger Rates

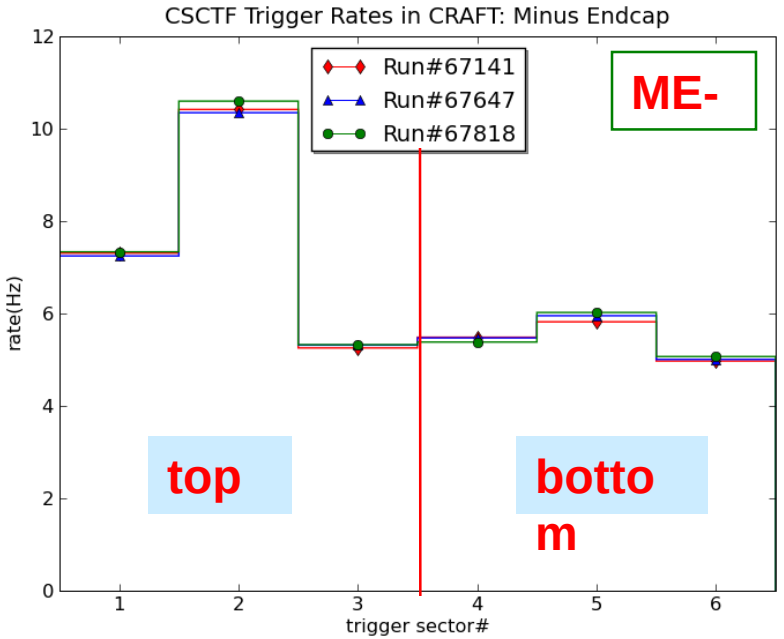
CSC trigger rate history

- Rates from different periods agree quite well
  - Total output rate: 74-75Hz,
- reflectes major activities/incidents

Variations with trigger sectors

- Several typical long runs indicate 4.5-10.5Hz from different sectors
- Features are as expected
  - Phi-dependent; top-bottom difference
  - Effects of the cavern: plus/minus endcap asymmetry





High Level Triggers (HLT)

- **Run on farm of commercial CPUs: a single processor analyzes one event at a time and comes up with a decision**
- **Has access to full granularity information**
- **Freedom to implement sophisticated reco algorithms, complex selection requirements, exclusive triggers …**

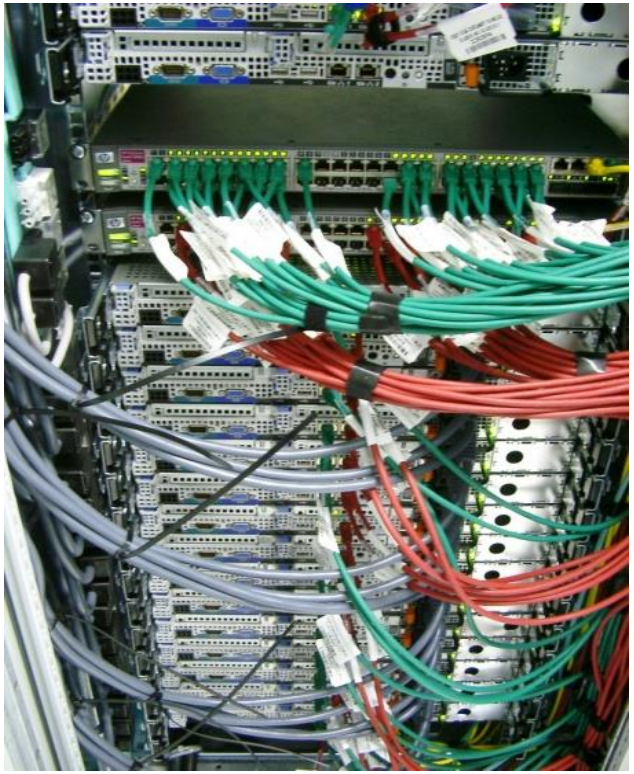
Constraints:

- CPU time (Cost of filter farm)
  - Reject events ASAP: set up internal "logical" selection steps
    - L2: muon+ calorimeter only
    - L3: use full information including tracking
- Must be able to measure efficiency from data
  - Use inclusive selction whenever possible
    - Single/double object above pT/ET, etc.
  - Define HLT selection paths from the L1
- Keep output rate limited (obvious…)

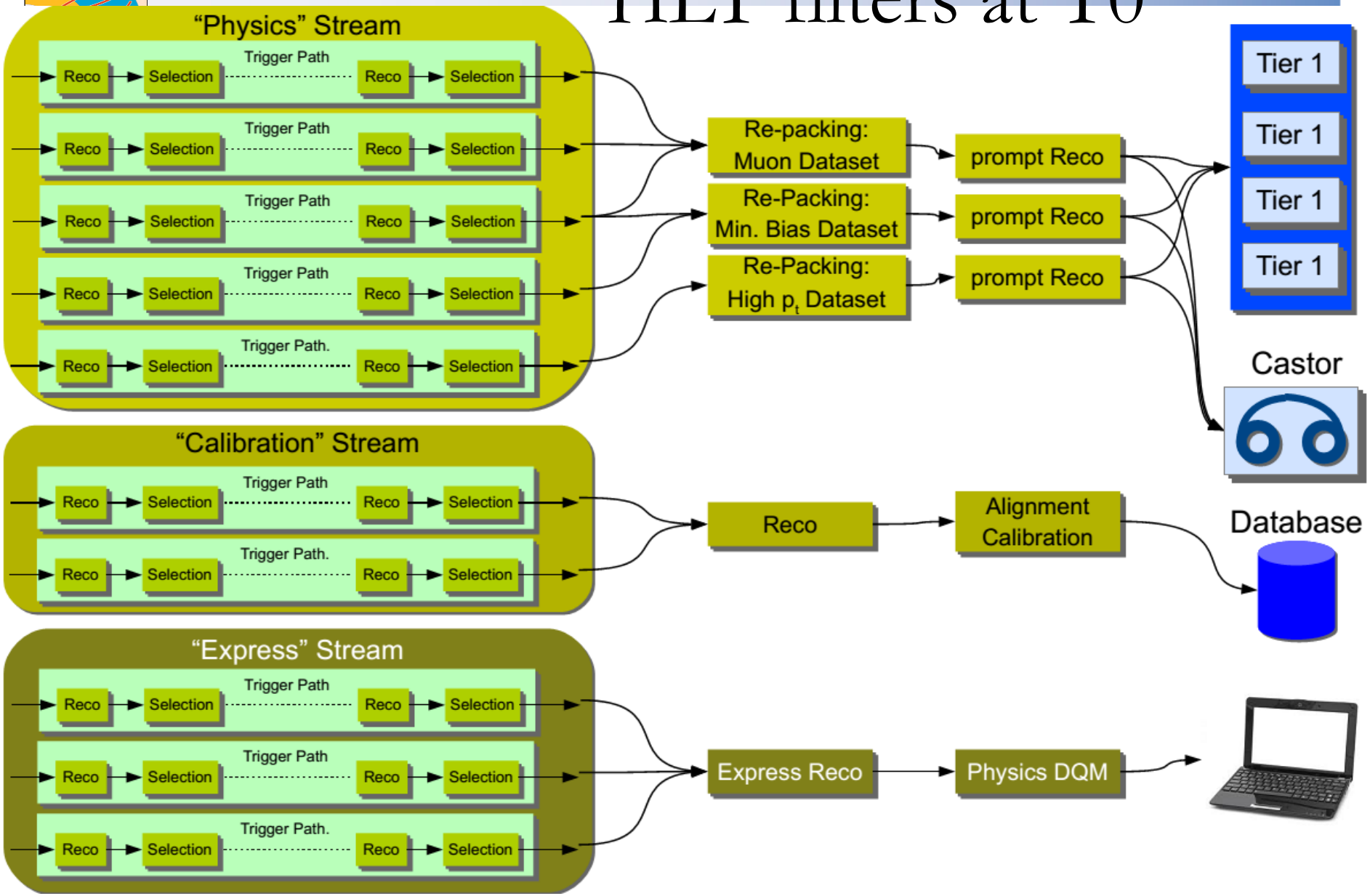
ATLAS HLT Hardware

First 4 racks of HLT processors, each rack contains

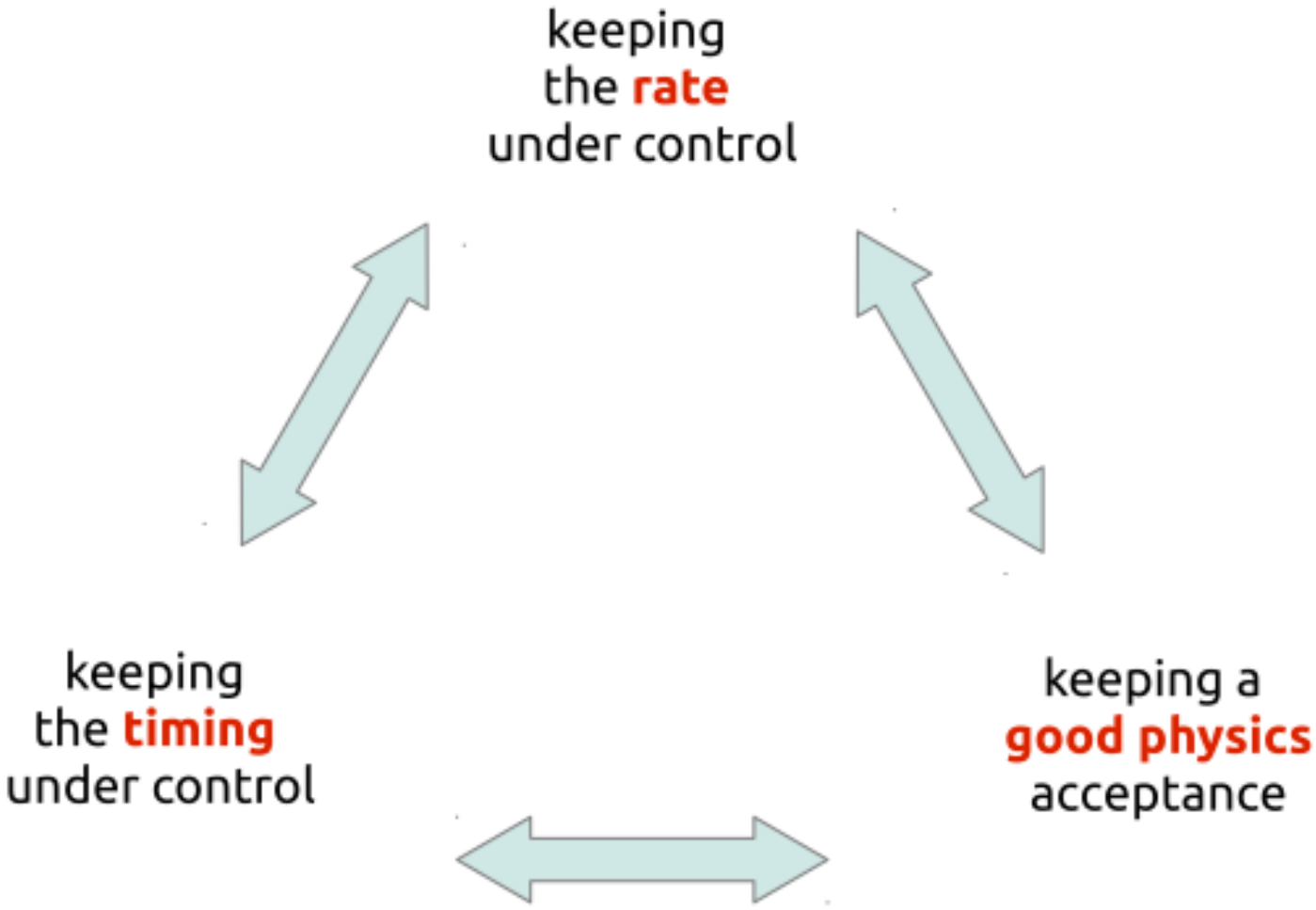
- ~30 HLT PC's (PC's very similar to Tier-0/1 compute nodes)
- 2 Gigabit Ethernet Switches
- a dedicated Local File Server



HLT filters at T0

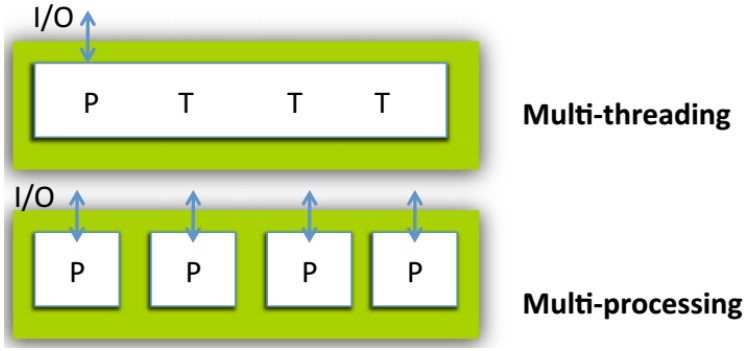


HLT Challenge: Compromise



HLT design principles

- **Early rejection**
  - Alternate steps of feature extraction with hypothesis testing: events can be rejected at any step with a complex algorithm scheduling
- **Event-level parallelism**
  - Process more events in parallel, with multiple processors
  - Multi-processing or/and multi-threading



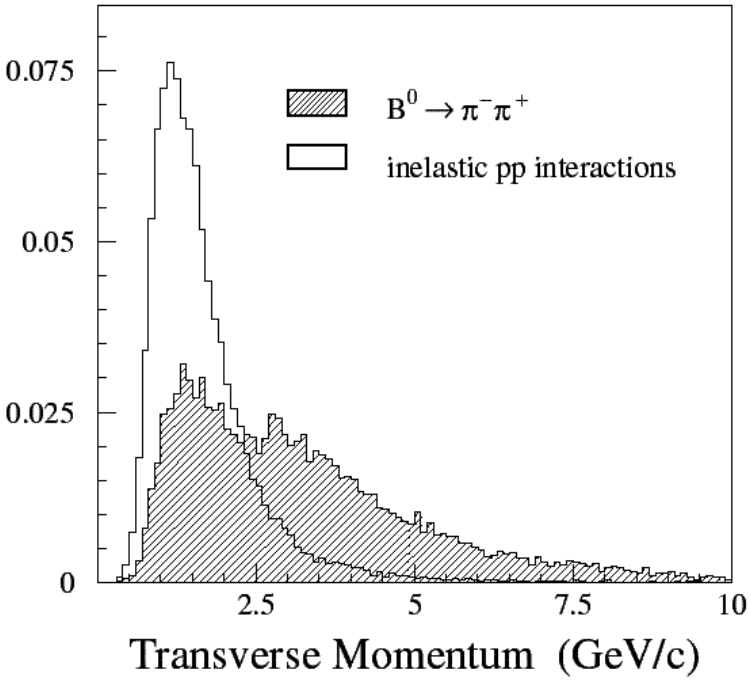
  - Queuing of the shared memory buffer within processors
- **Algorithms are developed and optimized offline, often software is common to the offline reconstruction**

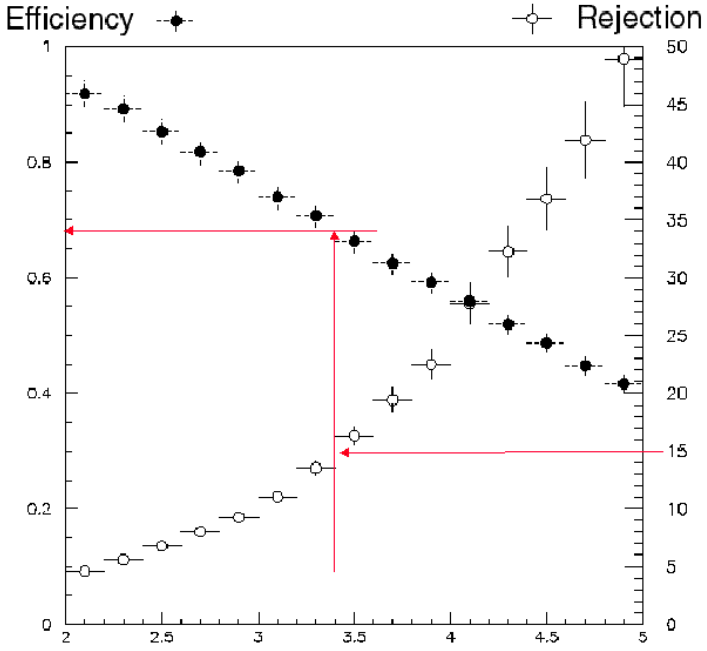
Event reconstruction used in HLT triggers

To optimize their performances, the reconstruction algorithms are designed by following some general rules:

- *regional reconstruction*: the algorithms run on the portion of the detector flagged as interesting by the L1 trigger;
- *partial reconstruction*: the reconstruction of the physics events is performed only to the precision necessary to select each event;
- algorithms are subdivided into sub-levels, to stop the reconstruction as soon as an event has to be discarded;
- the good events undergo all the recosntruction algorithms, to be divided in different physics streams after the selection.

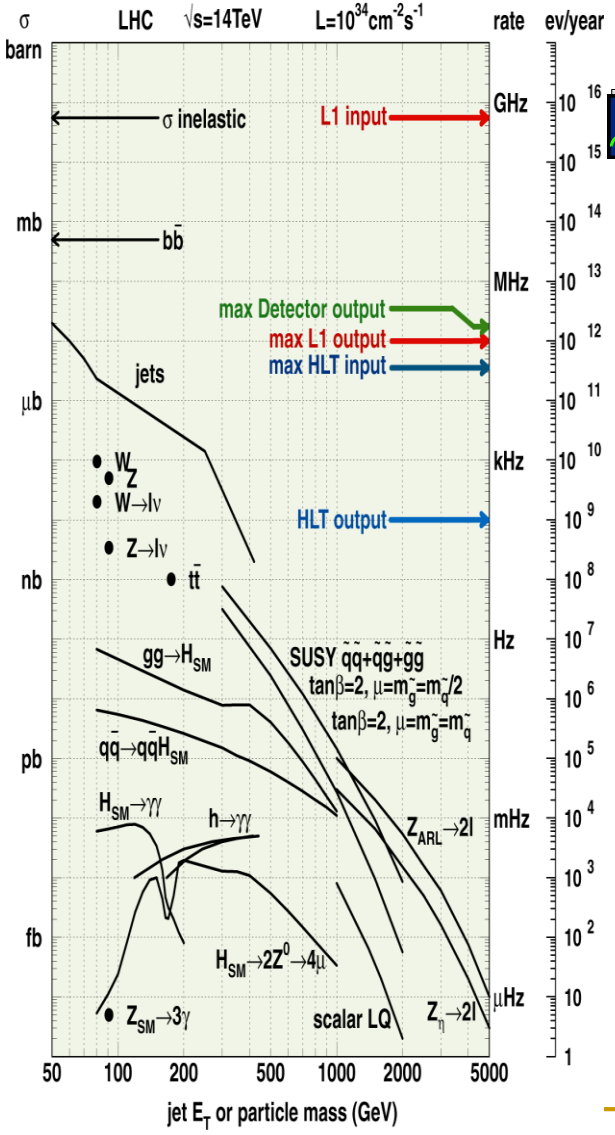
Trigger performance: Efficiency vs background rejection

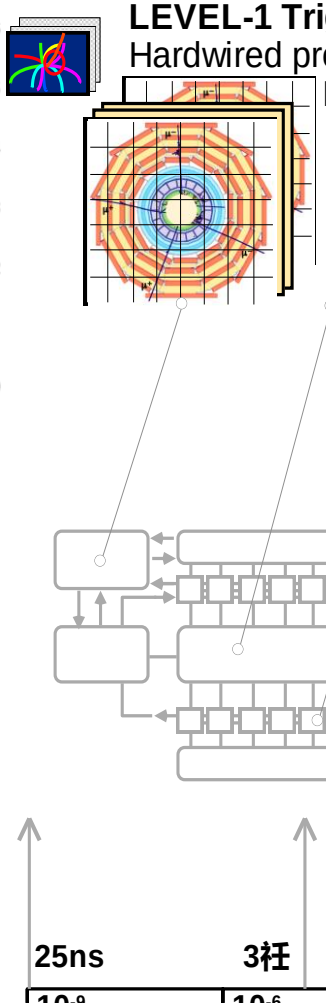




- Example:B meson trigger in LHCb
- **Discriminating variable:** Transverse momentum (P_T)

LHC实验数据流：物理信号的挑选





数据获取系统的任务

一个事例通过触发判选后，数据获取系统开始将该事例的数据读出，进行一定的处理，然后记录到磁带上。此外数据获取系统还有许多其他的任务，可列为：

1) 电子学的刻度和记档。

2) 运行时的初始化，前端各微处理器的加载，各状态控制寄存器的设置。

3) 从前端电子学读数.

4) 数据的预处理和装配。

5) 全事例数据的重建分析(在线事例筛选，即第三级触发)。

6) 数据的记录(磁带或磁盘).

7) 探测器运行情况的监测(抽样分析事例，建立各种直方图并作单事例和直方图显示)。

8) 运行的操作控制，如键盘命令输入、接触屏幕输入和鼠标输入等。

9) 运行条件的显示和记档(如运行RUN号、磁带号、加速器能量和亮度、环境和高压监测等，有时把8和9称为slow control)。

10) 错误显示，报警和记档。

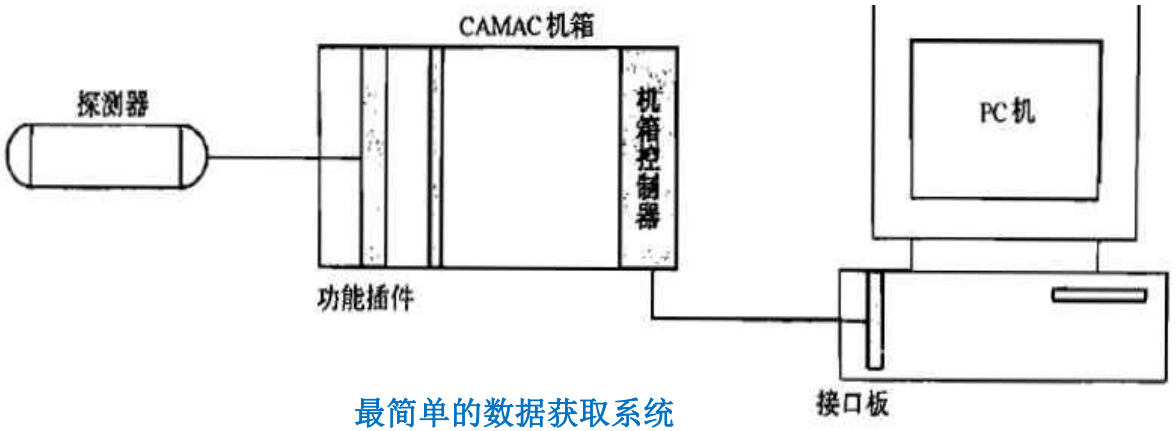
11) 磁带或磁盘上记录的数据的回读及处理。

任务3),4),5),6)是对每一个事例进行的。数据获取系统必须能够按所要求的数据率进行处理，同时又不造成太大的死时间损失。

单CPU的数据获取系统

最简单的数据获取系统由一个CAMAC机箱和一台PC机构成。它可以用于探测器研制，性能测试或束流试验。

CAMAC虽然速度较慢，但技术成熟，插件品种齐全，有现成的硬件和软件，适用于上述规模小，速度要求不高的场合。



最简单的数据获取系统

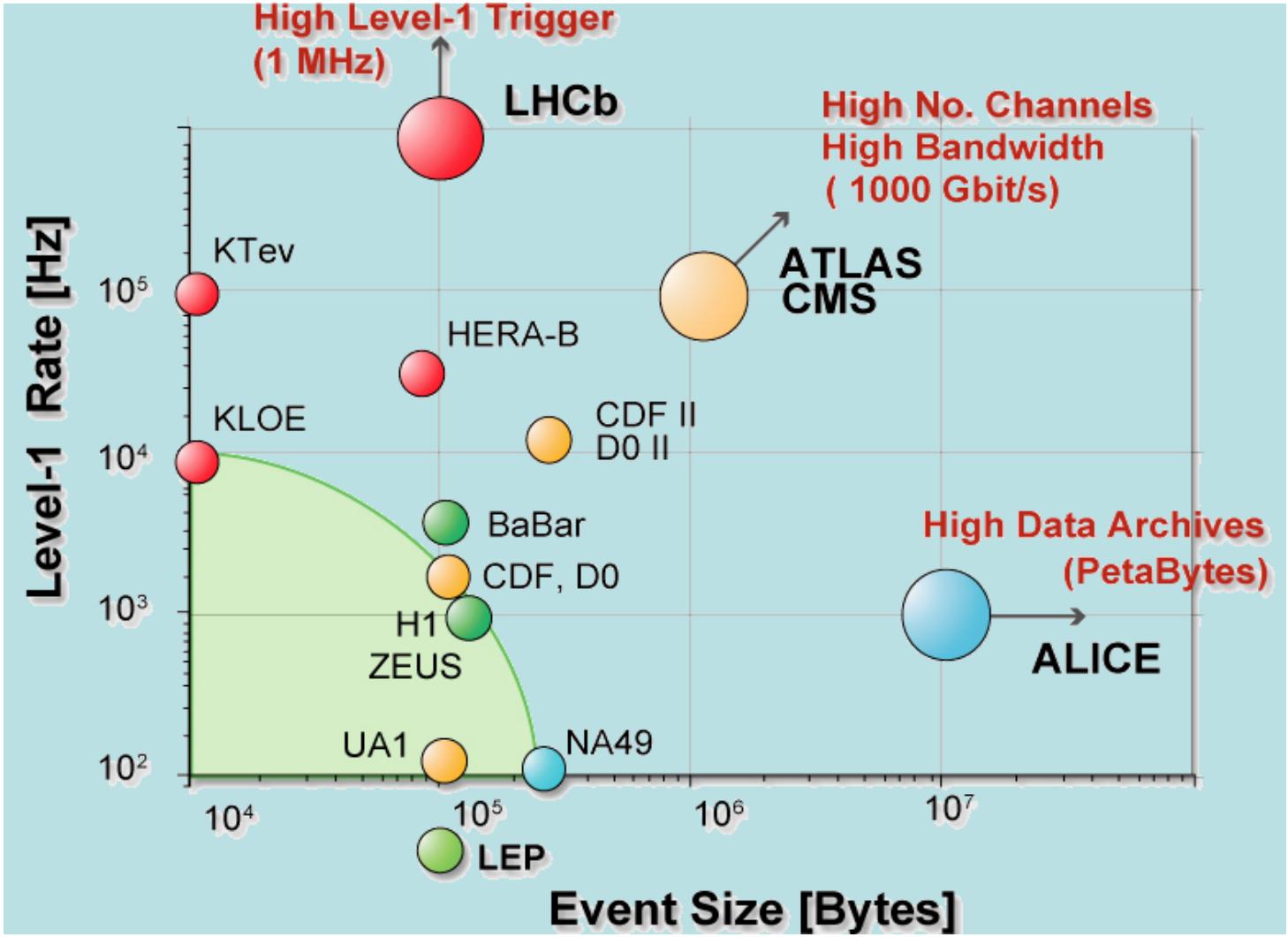
搭建系统时，在CAMAC机箱中插入所需要的ADC、TDC、定标器等功能插件。CAMAC机箱和PC机之间用U型机箱控制器和PC机内的CAMAC接口板相连。

通常在机箱控制器中有站号N寄存器，功能码F寄存器，子地址A寄存器，高、中、低位数据寄存器，H、M、L状态寄存器等，相应于PC机内的一些I/O地址。

PC机将一条CAMAC指令中的N、F、A和要写的数据分别写到相应的寄存器中，然后发出动作命令(可以是以访问另一个I/O地址的形式)，启动一个CAMAC周期。如果是读命令，则在CAMAC周期后，数据已经存入机箱控制器的数据寄存器中，PC机再从数据寄存器中读出。

较大的实验要用多个CAMAC机箱。1-7个机箱构成一个分支(branch)。几个分支都连到惟一的计算机。它执行的程序可分为读数和处理（处理数据，分析和存储）两部分。

Data Rate: TrigDAQ Comparisons

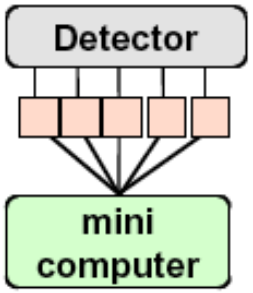


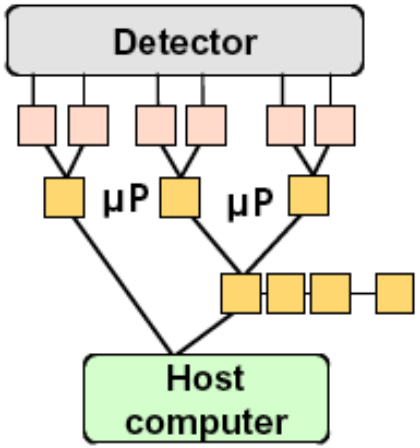
Summary of ATLAS Data Flow Rates

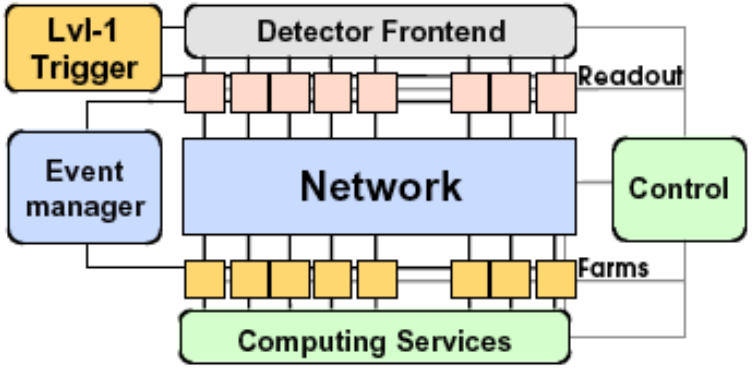
- From detectors > 10^{14} Bytes/sec
- After Level-1 accept ~ 10^{11} Bytes/sec
- Into event builder ~ 10^{9} Bytes/sec
- Onto permanent storage ~ 10^{8} Bytes/sec

$\rightarrow$ ~ 10^{15} Bytes/year

The evolution of DAQ systems







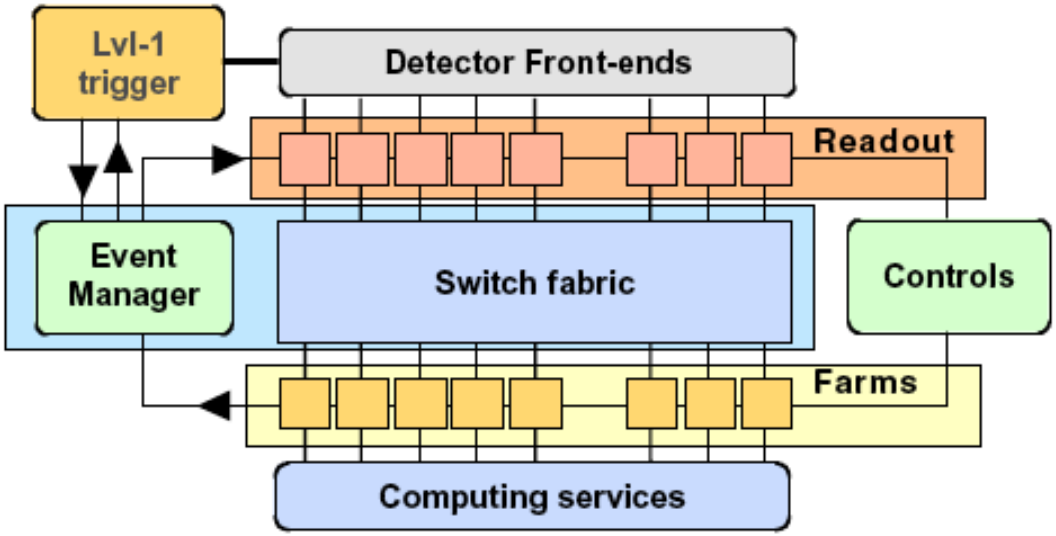
1970-80
MiniComputers
first standard:
CAMAC
•kByte/s

1980-90
Microprocessors
Distributed
systems
•MByte/s

1990-2000+
Communications networks
Control & Data networks
Embedded processors
•GByte/s

Typical architecture 2000+

- Basic Architecture: ~ same for most experiments

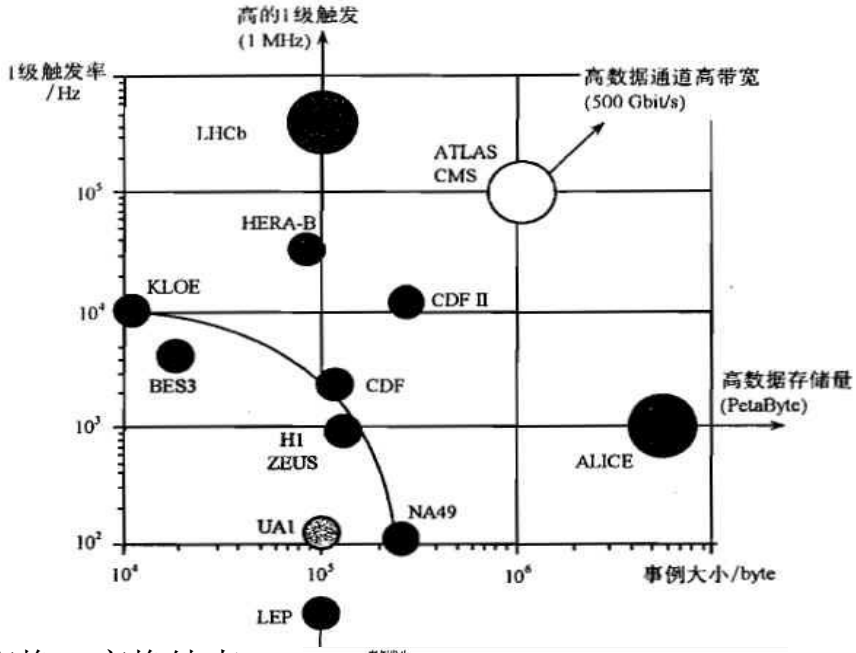


- Readout (units/drivers/buffers/…)
- Switching network
- Processor Farm
- Control & Monitor System

数据量：

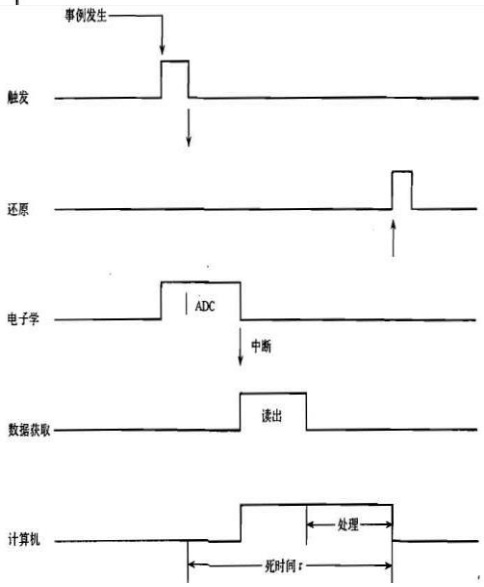
加速器能量提高使平均末态多重数从4增加到几百，且由于高能时的喷注现象，出射粒子往往成组集中在一个小的立体角里，径迹彼此靠得很近。为了区分不同径迹，探测器的单元要分得很细，使电子学的通道数从几万增加到上千万，每个事例的平均数据长度从几kB增加到1MB，大大增加了触发判选和数据获取系统的难度。

事例率和每个事例的数据长度的乘积为数据率。它代表数据获取系统单位时间内所要处理的数据量，是设计数据获取系统的基本参数。



死时间：

如果一个事例通过了触发判选，前端电子学就进行模数变换。变换结束后给出中断信号，使数据获取系统将数据读出并加以一定的处理，最后通知触发系统发出还原信号。这个过程中前端电子学不能再接受从探测器来的信号，这段时间内发生的事例都不能被记录下来，称为死时间。减少死时间造成的损失的方法：**改进触发判选；加快数据获取系统的速度或压缩数据量；使用缓冲存储器**。



实验中事例发生的时间是随机的，数据获取系统的处理能力应比要处理的平均计数率快一个数量级。如果有缓冲存储器，则只要把事例数据读入缓冲存储器即可还原，减少了计算机处理时间。

只要缓冲存储器没有满，就可以快速写入并低速读出。这样经过缓冲器后事例的随机性降低，后面的设备就可以按接近平均事例率的速度进行处理，降低对它们的要求。

Dead time

- Experiments frozen from trigger to end of readout
  - Trigger rate with no deadtime = R per sec.
  - Dead time / trigger = τ sec.
  - For 1 second of live time = 1 + Rτ seconds
  - Live time fraction = 1/(1 + Rτ)
  - Real trigger rate = R/(1 + Rτ) per sec.

Rate in Hz	Dead time ms.	Live time %	Trigger rate Hz
10	10	91	9.1
1000	10	9.1	91

数据获取系统中减少死时间损失的方法：

1) 较快的总线。如用FASTBUS或VME代替CAMAC。

2) 压缩数据量。高能物理实验一个事例中大多数探测器单元没有击中，排除掉输出为零的各道数据可大大减少每事例的数据量。例如BES有2万多道信号，平均一个事例只有700多个击中单元，经过零压缩后，事例数据长度从40kByte减小到3kByte，大大减少了数据传输时间。

3) 用多个智能的控制器并行读出，这时每个控制器所要读出的数据量减少，而且有些智能的控制器能储存对每个事例所要做的一组操作命令，接到启动信号后立即按顺序执行，并有一定的缓冲存储器暂存数据，不必将数据传送到计算机即可还原。

4) 各级间加缓冲存储器。

5) 存储器直接访问(DMA)或数据块传输((block transfer)。数据传输速率的主要限制往往不在于硬件的快慢，而在软件的时间开销。一般每条计算机指令只能传送一个数据。如果要连续传送多个数据，可以用存储器直接访问或数据块传输.。CPU只要指定要访问的首地址和传送数据长度，就可以让数据传输在DMA控制器或其他硬件的控制下进行，以硬件允许的最高速度传送多个数据。

6)用嵌入式的单板机和实时操作系统代替集中式的多用户多进程的主机，以加快中断响应。后者的中断响应时间长达几个毫秒，而前者仅十微秒。

读出单元

由于新实验的高数据率，VME总线已不敷需要，一般把VME总线用作初始化等慢控制，而将数据读出任务交给一个专用的读出单元。在接到L1触发信号后，它把N个前端电子学模块(如FADC等)的数据集中，储存在缓冲存储器内，当接到控制器的发送命令时，再把集中了的数据发送到事例组装器。它可以看做为局部的事例组装器。它的输入连接方式可以是总线连接或点对点(如光纤)连接，输出一般为点对点连接。

事例组装器(event builder)

在新的高能物理实验装置中，数据获取系统要有partitioning(分块)和scalability(可扩充)的能力。Partitioning就是指每一个子探测器的数据获取都构成一个相对完整的小系统，这样便于对各子探恻器进行预制研究和束流试验，并在这过程中完善这小系统的硬件构成和软件程序。当整机联调时，只要把各个小系统连接在一起构成一个大系统，在调试时只要关心有关总体连接的问题就可以了。Scalability是指当探测器分阶段扩充时，新增加部分的数据获取仍可以纳入原有的框架，不会对原有的系统有大的影响。

事例组装器的任务就是把从各个子探测器来的数据段装配成完整的事例数据，送到第三级触发进行在线事例筛选。按照其结构分成三类：

(1)基于单个总线的事例组装器：　各个子探测器前端电子学来的数据段先分别储存在相应的缓冲存储器中，这些缓冲存储器都插在同一个背板总线(如VME或Fastbus)机箱内。事例组装器也连到这个总线，通过总线顺序地从各缓冲存储器中读取属于同一个事例的各个数据段，合并装配成一个事例 。

(2)双总线矩阵式事例组装器：由A,B两种总线纵横构成一个矩阵，在每个交点上有一个双口缓冲存储器因为有n条B总线可以并行地进行事例装配，允许的最高数据传送率也提高了n倍。但这种事例组装器的结构比较复杂，扩充较难。

(3)交换机式事例组装器：　网络交换机是计算机和通信技术的最近发展，在现有的总线和以太网中，多个源设备和目标设备都连到同一个传输介质。某一个设备按照一定的协议获得这个介质的使用权，进行数据传输。每一时刻只能在一对源设备和目标设备之间进行数据传输，其他设备只好等待。当数据率很高时，这单一的传输介质就成为瓶颈。网络交换机就是为解决这个问题而发展起来的。它可以在n个输入端和n个输出端之间建立任意的一对一连接。

在线事例过滤机群(farm)

在线排事例过滤机群完成第三级触发，采用大的处理器阵进行实时事例筛选。这种处理器阵也可以用于全事例重建、Monte-Carlo模拟和离线分析。

控制检测系统

由于新数据获取系统的巨大规模、采用多种最新技术、数据多级缓冲、高带宽又有秒级的延迟以及昂贵的运行费用，使新数据获取系统必须在各种商用软件的基础上建立运行控制系统，并检测系统各部分的运行情况，保证实验的可靠运行。

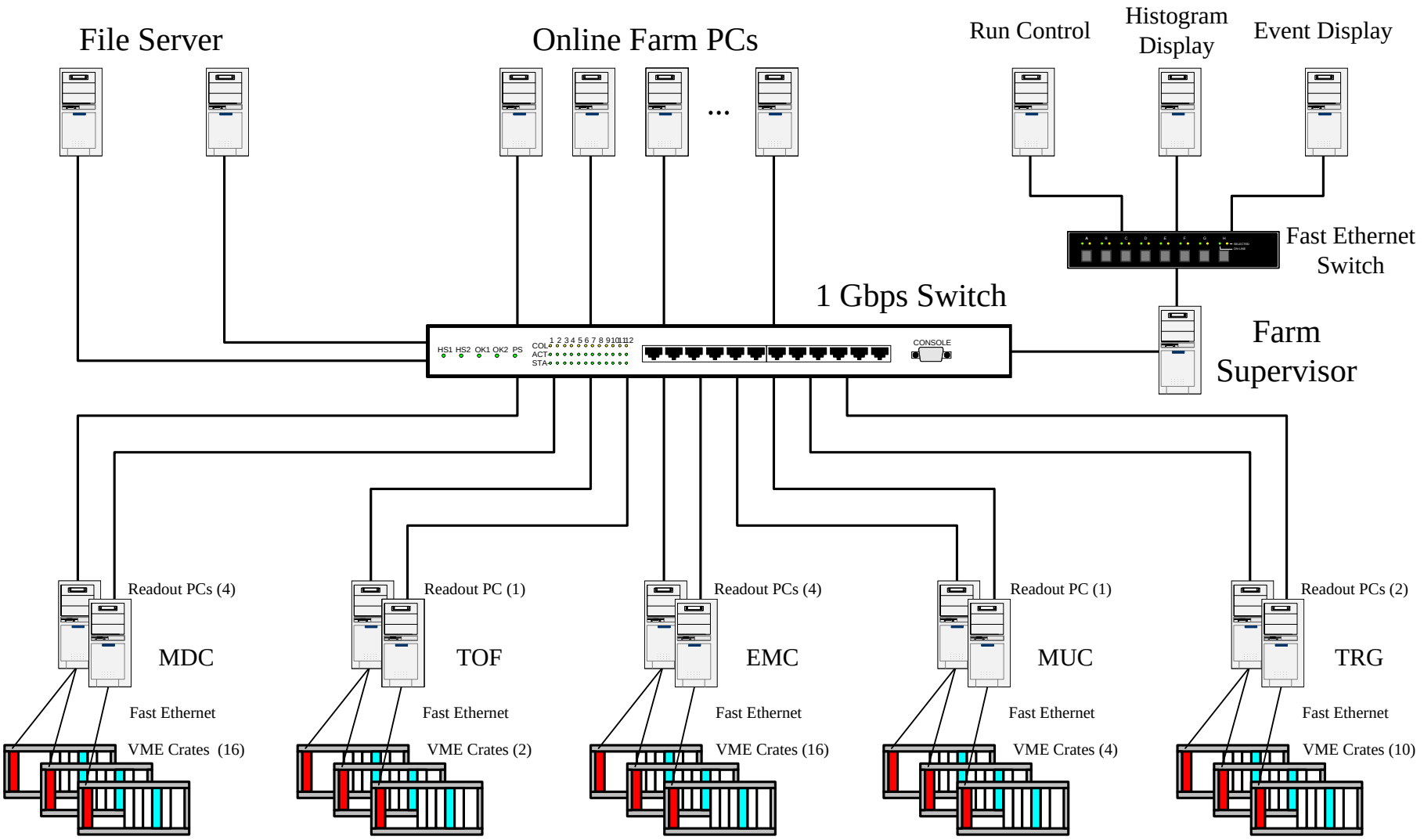
新一代的高能物理实验对数据获取系统提出了极高的要求。以CMS数据获取系统为例：

- 当L1触发率为100kHz时，CMS数据获取系统控制系统的消息流量为10^6msg/s；
- 要有约5万个前端驱动器每秒钟从1亿路电子学通道中取得1Tbit数据，存放在200GByte的缓冲存储器中，然后经过一个500X 500的交换器网络(带宽要求500Gbit/s)送到事例筛选机群组装。
- 经过筛选的事例率为10~100Hz，事例长度为1 MByte，每天将产生I丁Byte的数据。
- 为了处理这些数据将要动用全世界各大实验室的计算机，构成一个全球网络(Grid)，共享CPU和存贮器资源。

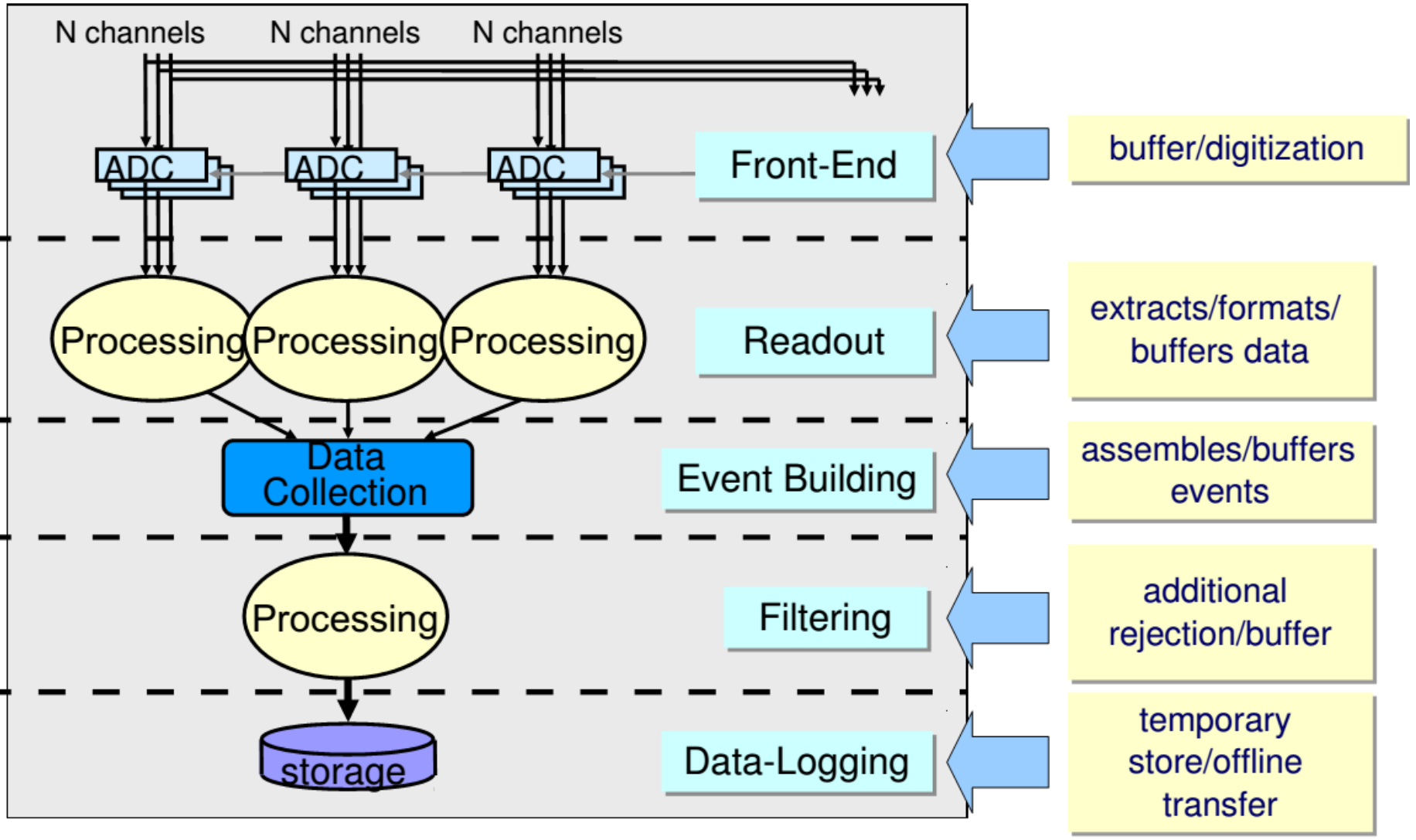
设计指标细化

- 数据读出和组装，并发送到计算中心；
  - 最大触发率：
  - 单机箱数据处理能力：
  - 单台读出计算机（连4台机箱）数据处理能力：
  - 在线机群总处理能力和本地数据暂存能力：
- 数据处理，格式转换、直方图填充和在线过滤框架等；
- 提供运行控制和实时探测器状态监测；
- 多级数据缓冲、并行处理以及网络技术；
- 模块化设计，易扩展和升级。

网络DAQ基本框架

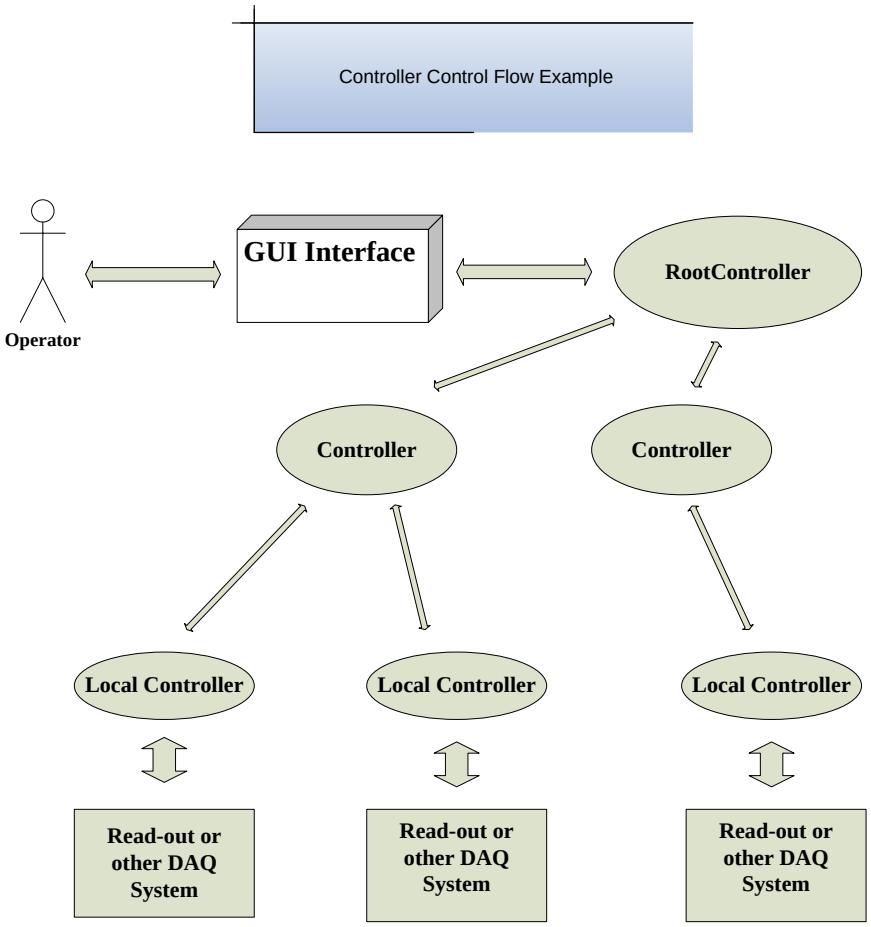

File Server
Online Farm PCs
...
Run Control
Histogram Display
Event Display
Fast Ethernet Switch
1 Gbps Switch
Farm Supervisor
HS1 HS2 OK1 OK2 PS
COL
ACT
STA
CONSOLE
Readout PCs (4)
MDC
Fast Ethernet
VME Crates (16)
Readout PC (1)
TOF
Fast Ethernet
VME Crates (2)
Readout PCs (4)
EMC
Fast Ethernet
VME Crates (16)
Readout PC (1)
MUC
Fast Ethernet
VME Crates (4)
Readout PCs (2)
TRG
Fast Ethernet
VME Crates (10)

Large DAQ system



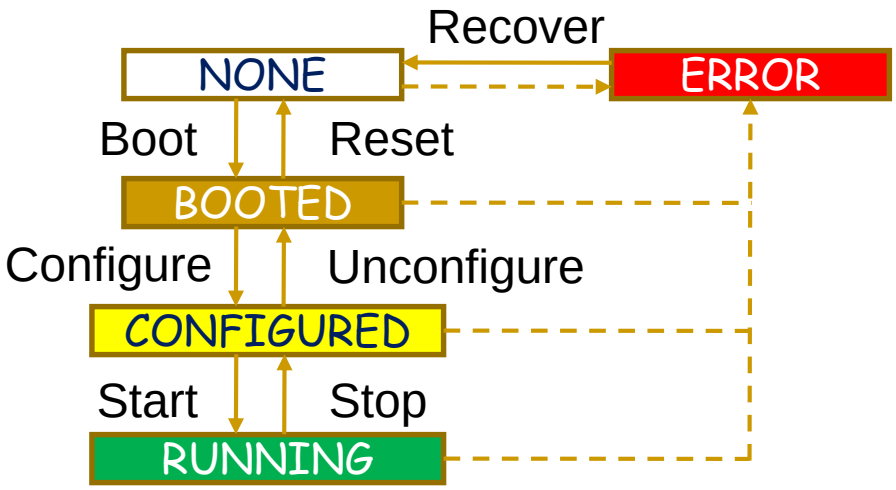
运行控制（RunControl）

- 负责系统运行控制，提供DAQ系统的状态管理，控制数据获取的动作行为
- 遵循有限状态机模型的控制器分级系统，组织成树型层次结构，避免单一控制结点产生消息瓶颈



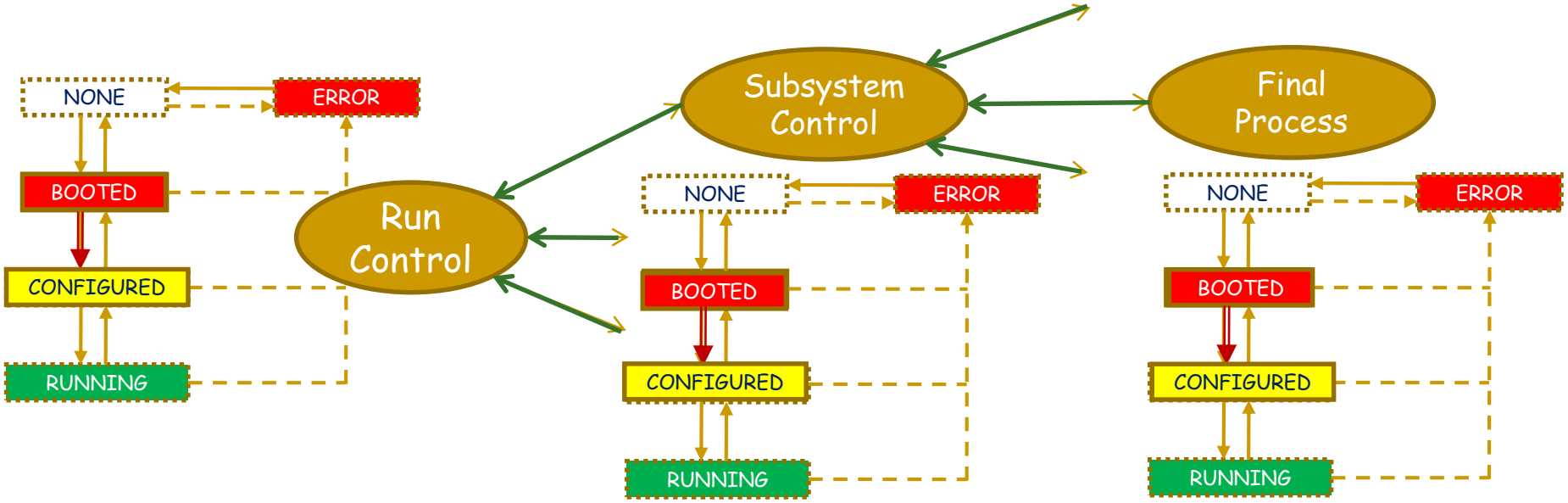
DAQ software: Finite State Machines

- Models of the behaviors of a system or a complex object, with a limited number of defined conditions or modes
- Finite state machines consist of 4 main elements:
  - States which define behavior and may produce actions
  - State transitions which are movements from one state to another
  - Rules or conditions which must be met to allow a state transition
  - Input events which are either externally or internally generated, which may possibly trigger rules and lead to state transitions



Propagating transitions

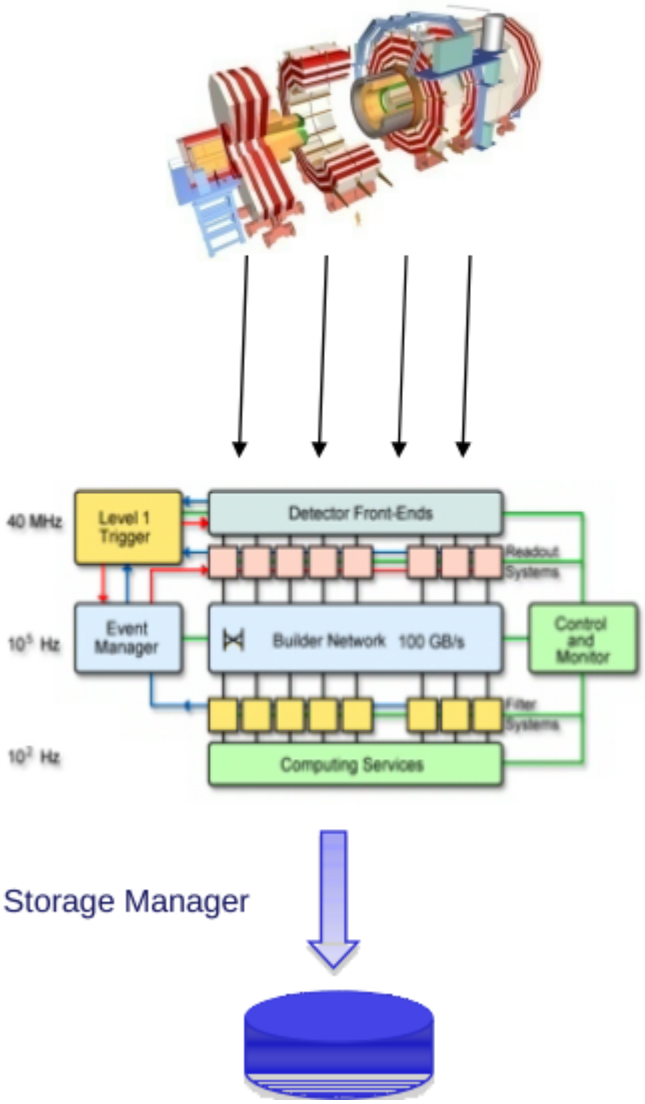
- Each component or sub-system is modeled as a FSM
  - The state transition of a component is completed only if all its sub-components completed their own transition
  - State transitions are triggered by commands sent through a *message system*



FSM implementation

- State concept maps on object state concept
  - OO programming is convenient to implement SM
- State transition
  - Usually implemented as callbacks
    - In response to messages
- Remember:
  - Each state MUST be well-defined
  - Variables defining the state must have the same values
    - Independently of the state transition
- There is also a message system
  - Based on some network protocal (SOAP, TCP/IP)

Overview of the CMS DAQ and useful terminology



- Detector signals are collected through individual data acquisition systems (cables and boards) that end up at the FEDs: the first element of Global Data Acquisition system (DAQ)

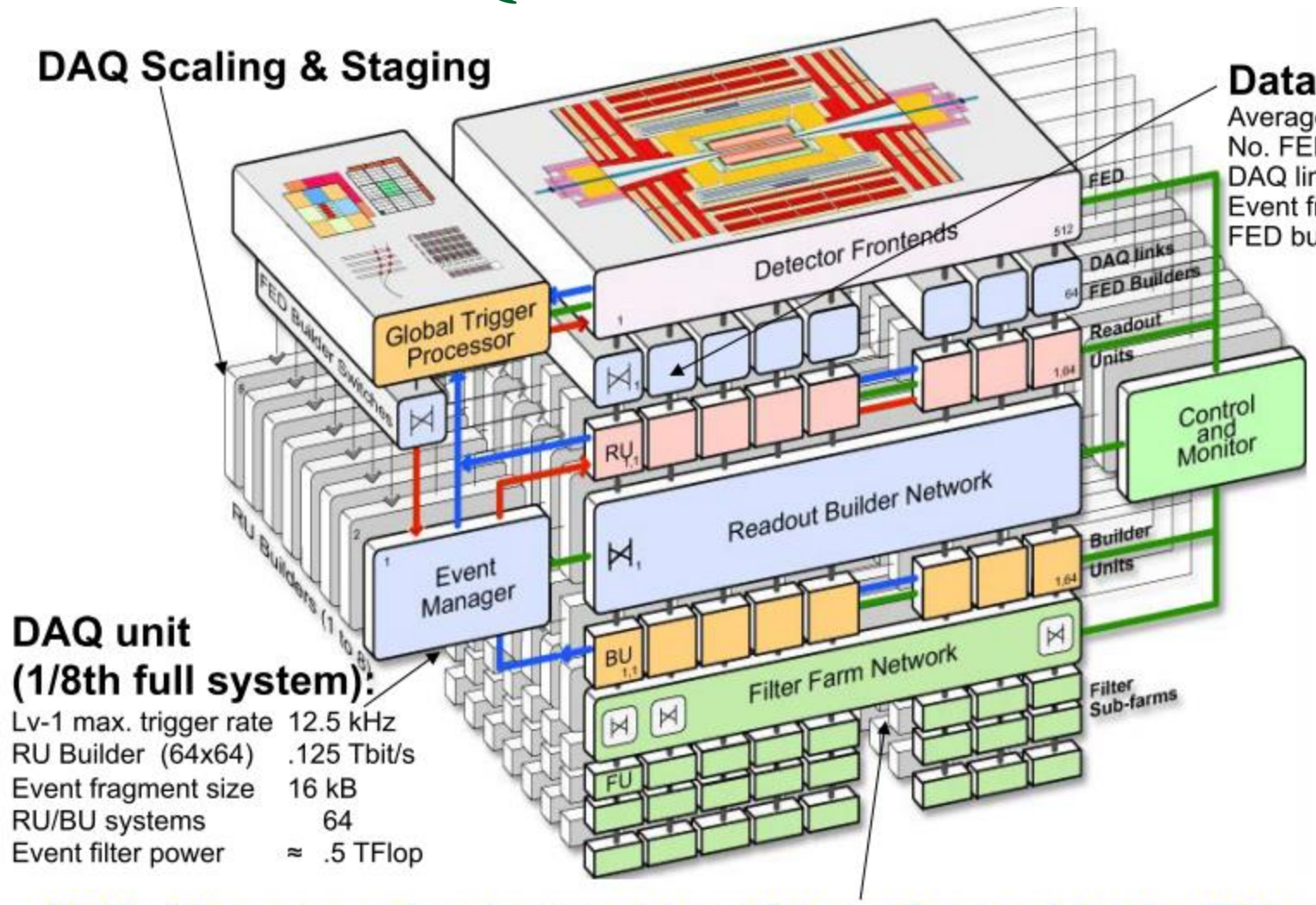
- **FED** (detector FrontEnd boards): multiple FEDs per detector collect event fragments that are sent to the online event processing farm

- **Builder Units:** Computing farm that collects event fragments from all FEDs and merge them to produce full event information

- **Filter Units:** Computing farm where the High Level Trigger (HLT) is run to filter interesting events

- **Storage Manager:** application that saves to local disks events selected by the HLT

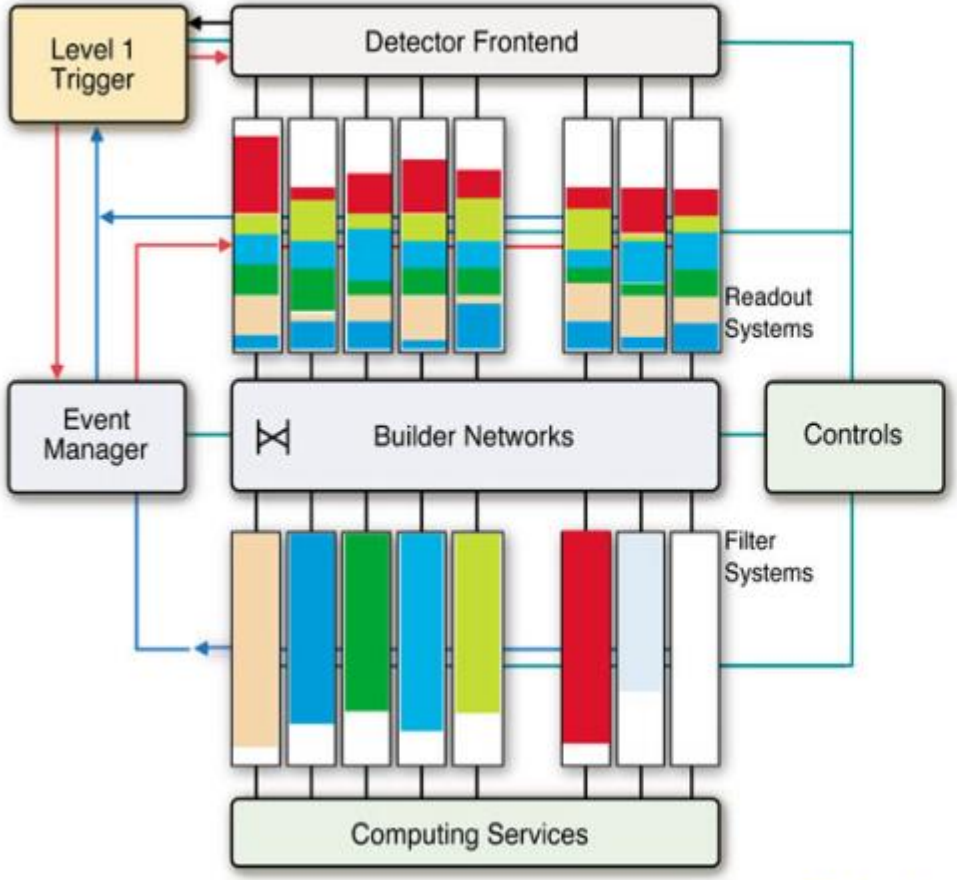
CMS DAQ



HLT: All processing beyond Level-1 performed in the Filter Farm
Partial event reconstruction "on demand" using full detector resolution

Event Building

Event builder :
Physical system interconnecting data sources with data destinations. It has to move each event data fragments into a same destination

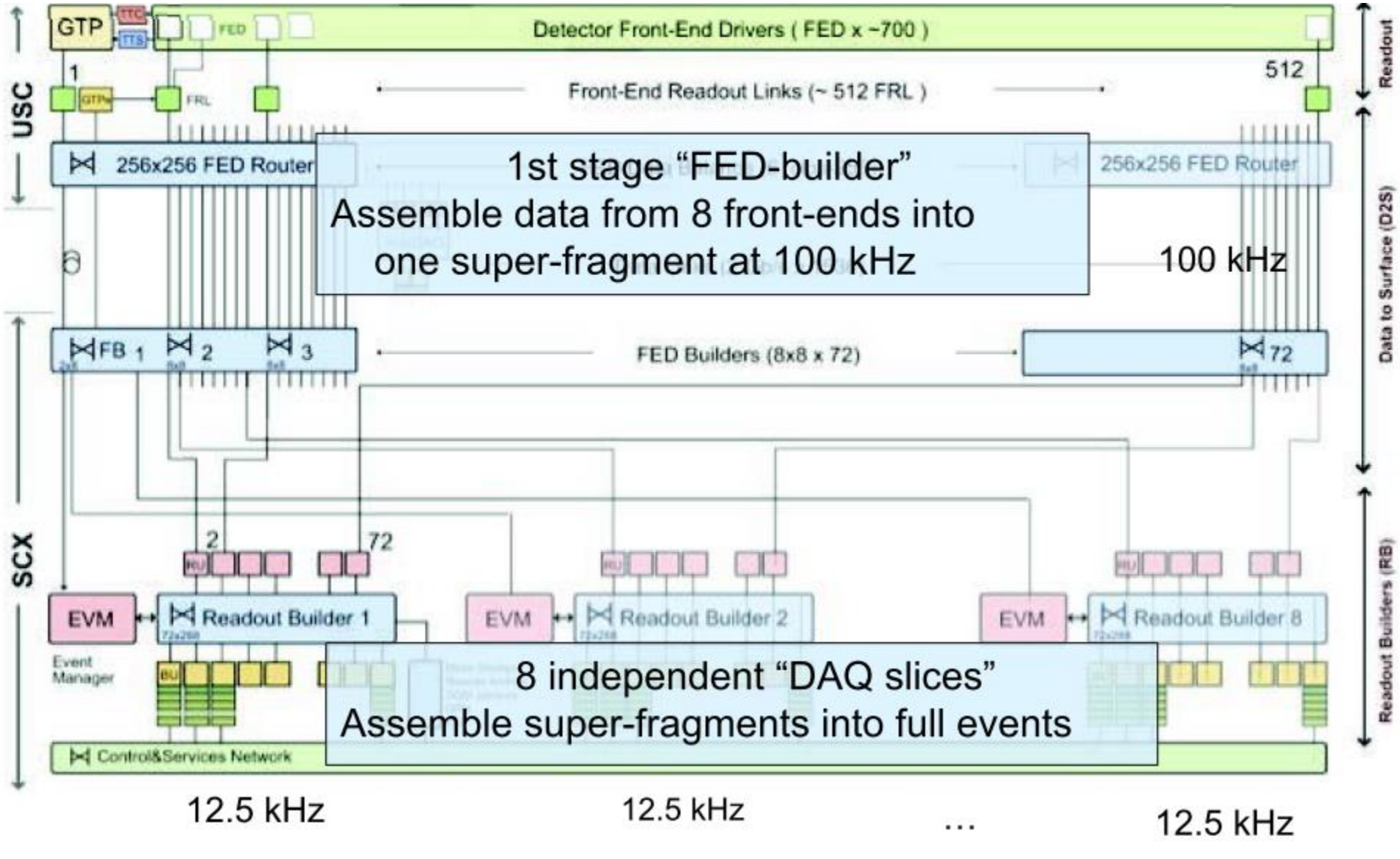


Event fragments :
Event data fragments are stored in separated physical memory systems

Full events :
Full event data are stored into one physical memory system associated to a processing unit

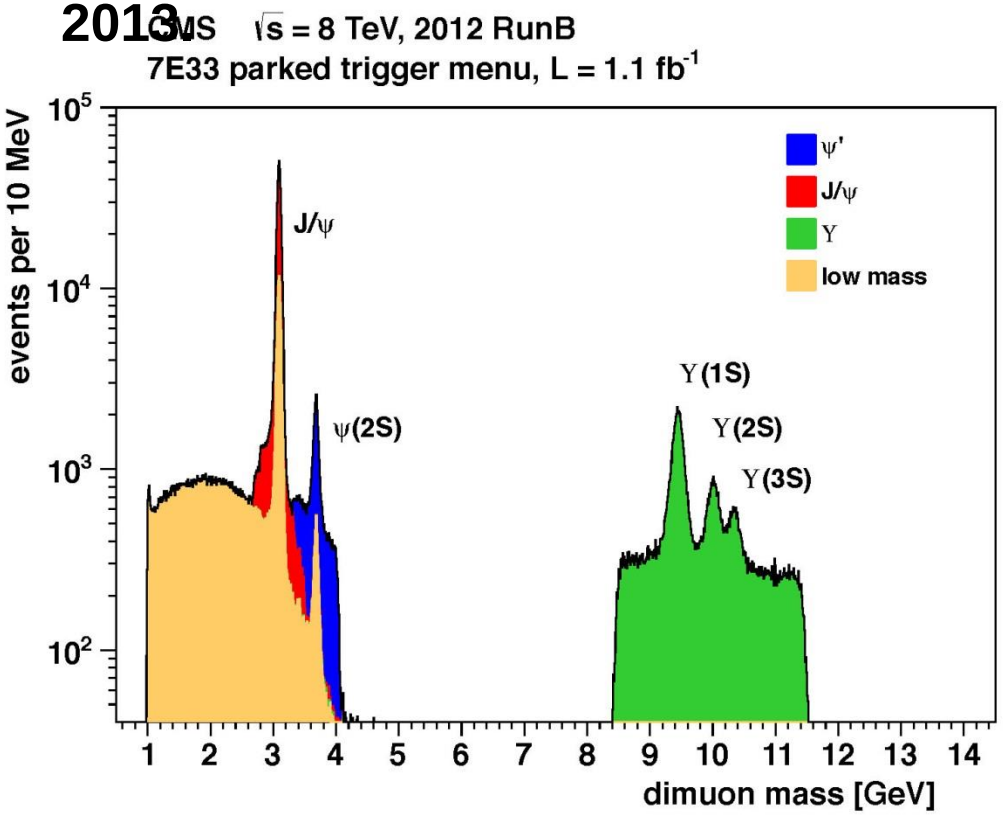
Hardware:
Fabric of switches for builder networks
PC motherboards for data Source/Destination nodes

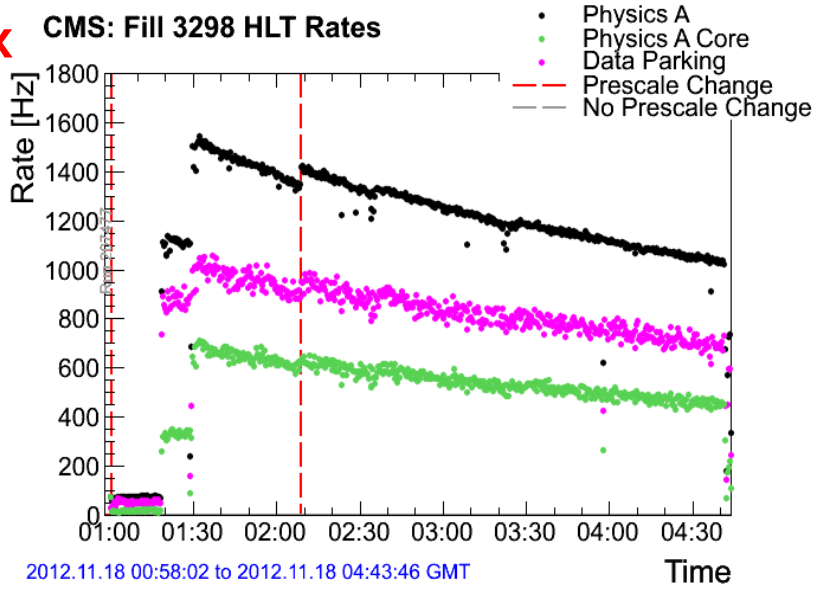
Two-Stage Event Builder

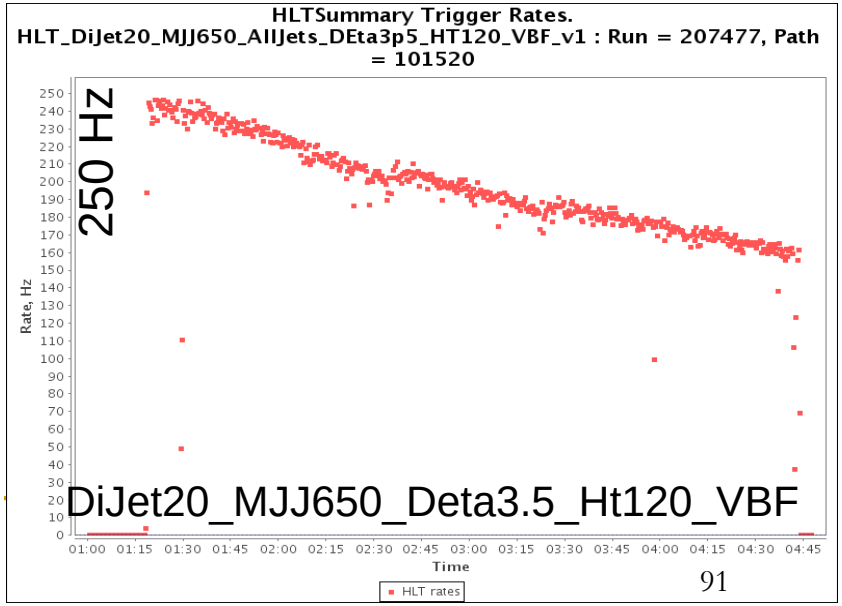


Data Parking

- **We don’t want CMS physicists to relax after summer 2013**
- ***Keeping almost x2 of additional data that before was discarded!***
- **“Parked data”: to be reconstructed 2013**



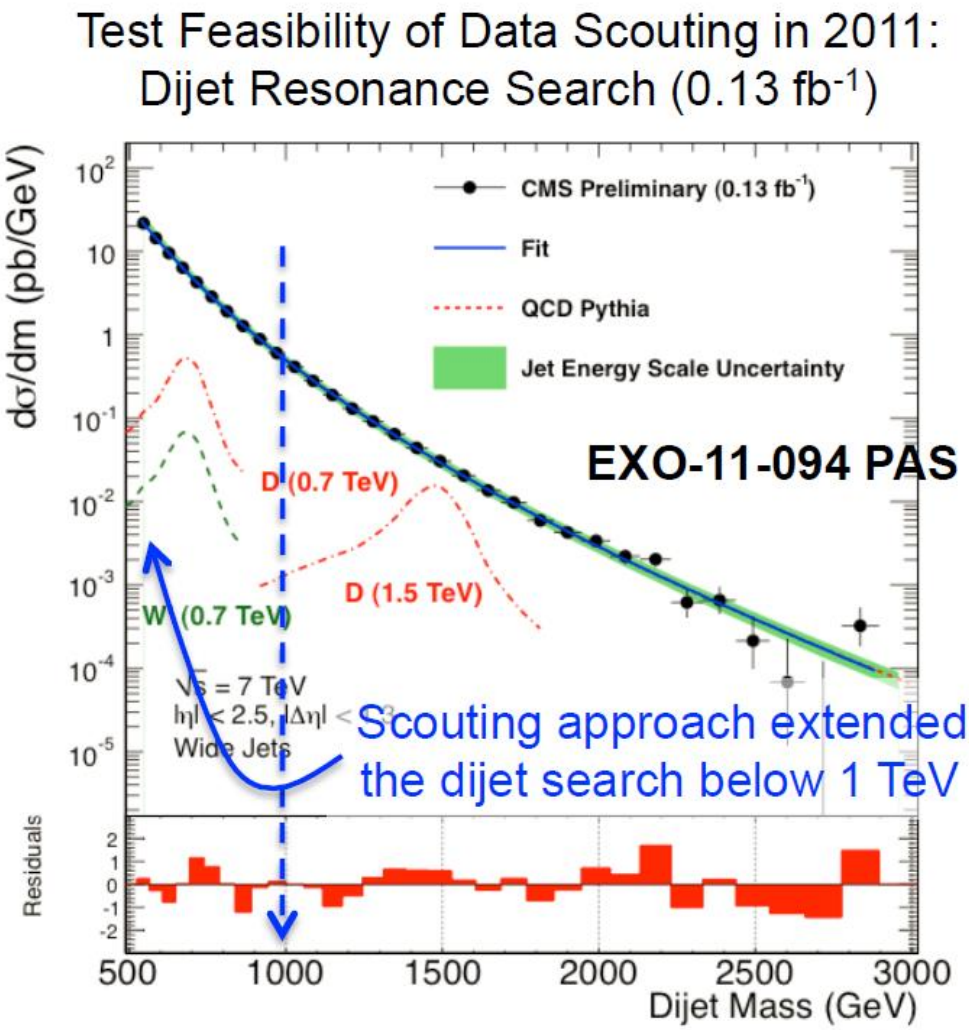




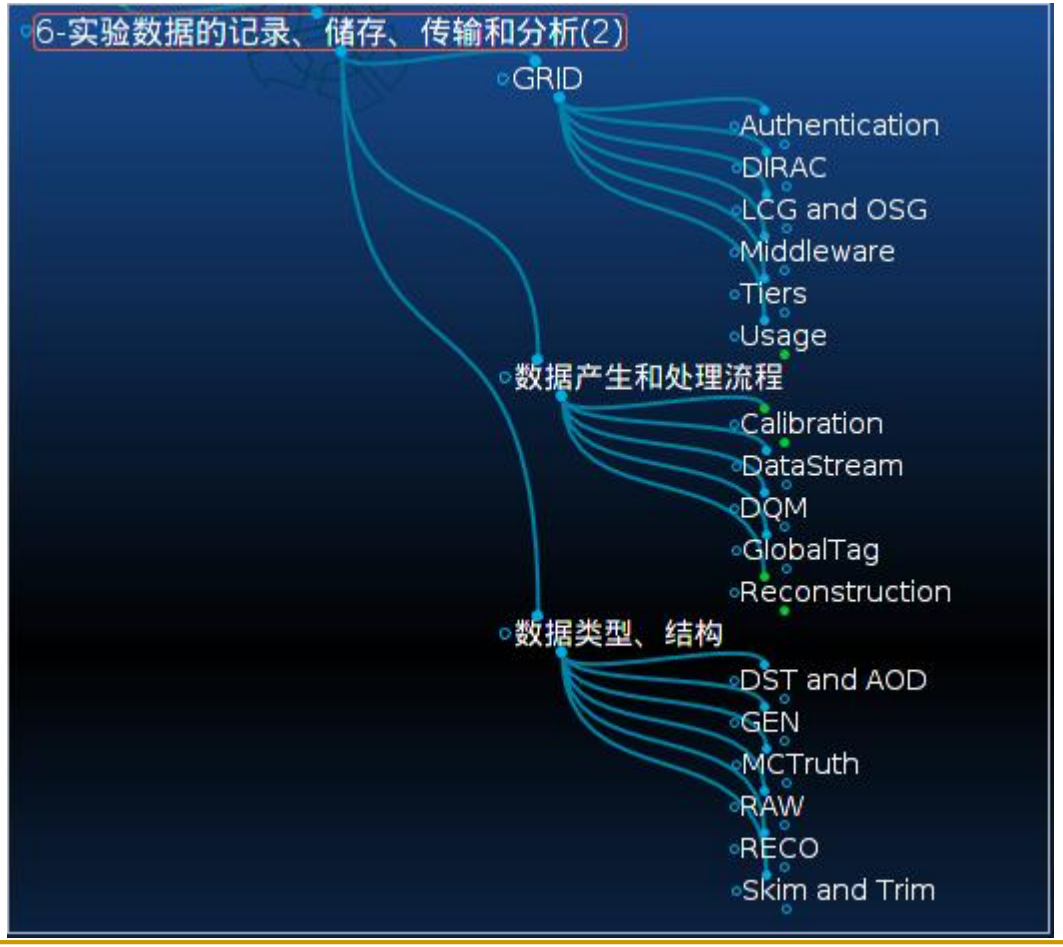
Data Scouting

Take stream of data at very large rat
but storing minimal information:
reduced event content

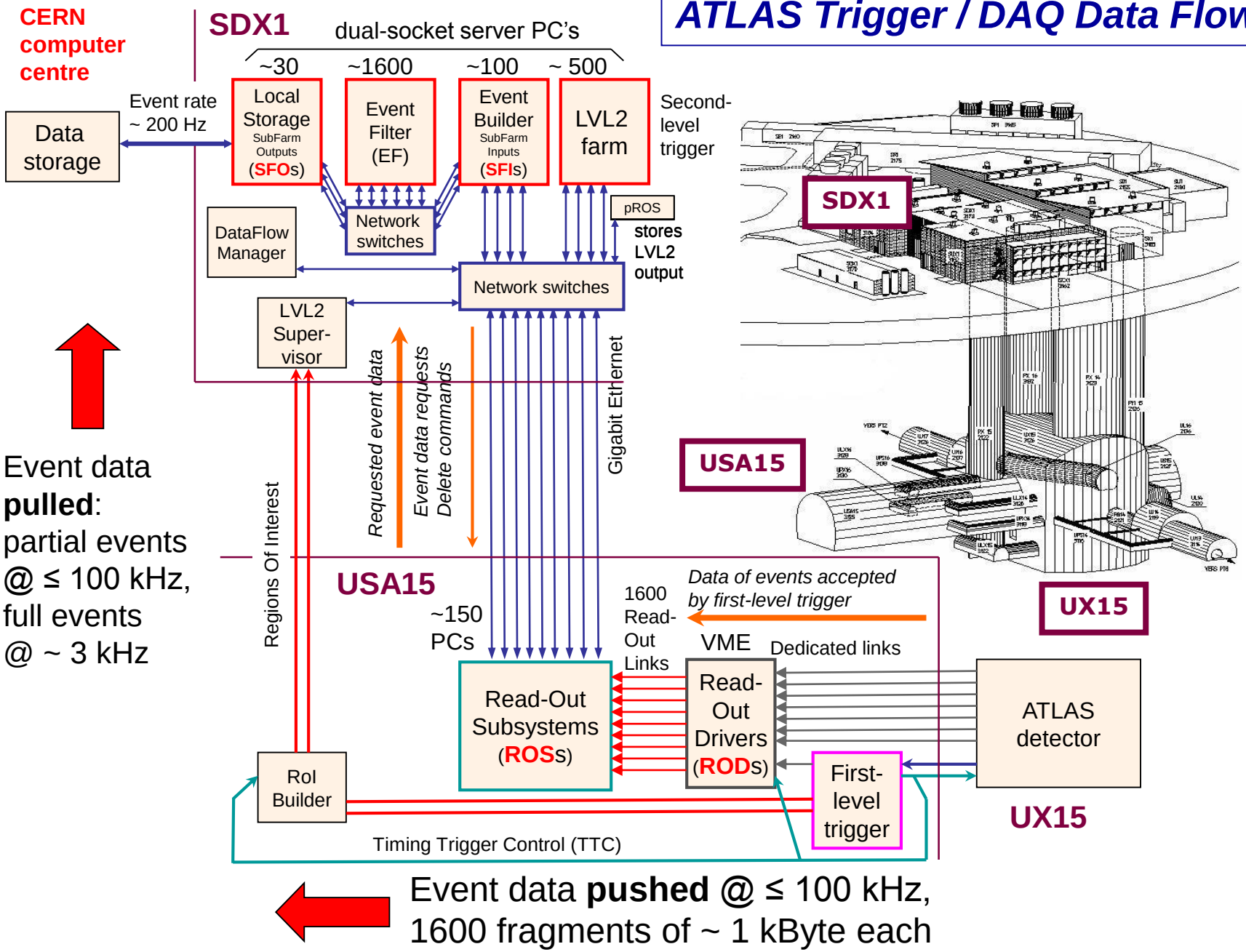
If something really interesting is found on data scouting, may conside to implement looser trigger on main stream



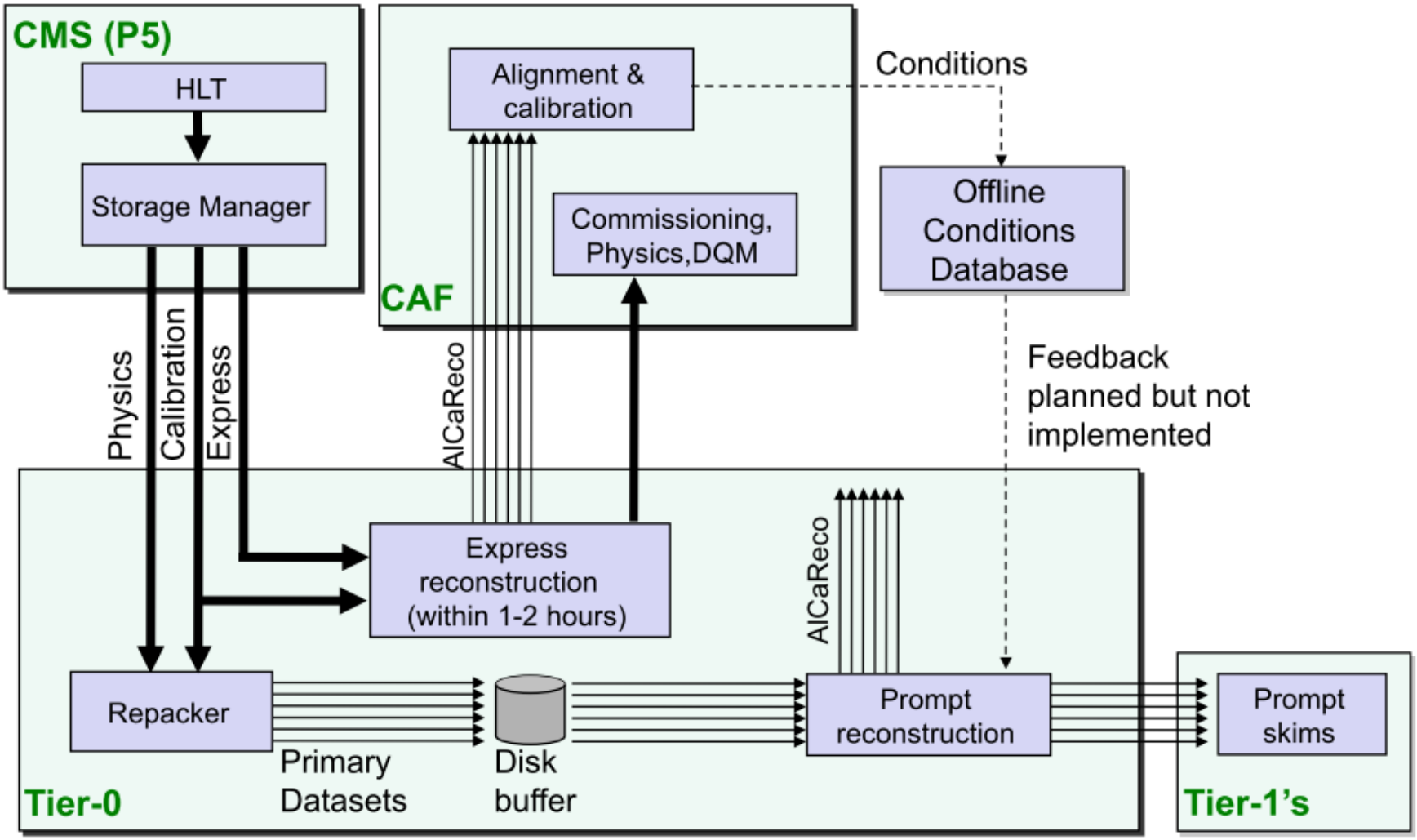
实验数据流：
监测，记录、存储、传输



ATLAS Trigger / DAQ Data Flow

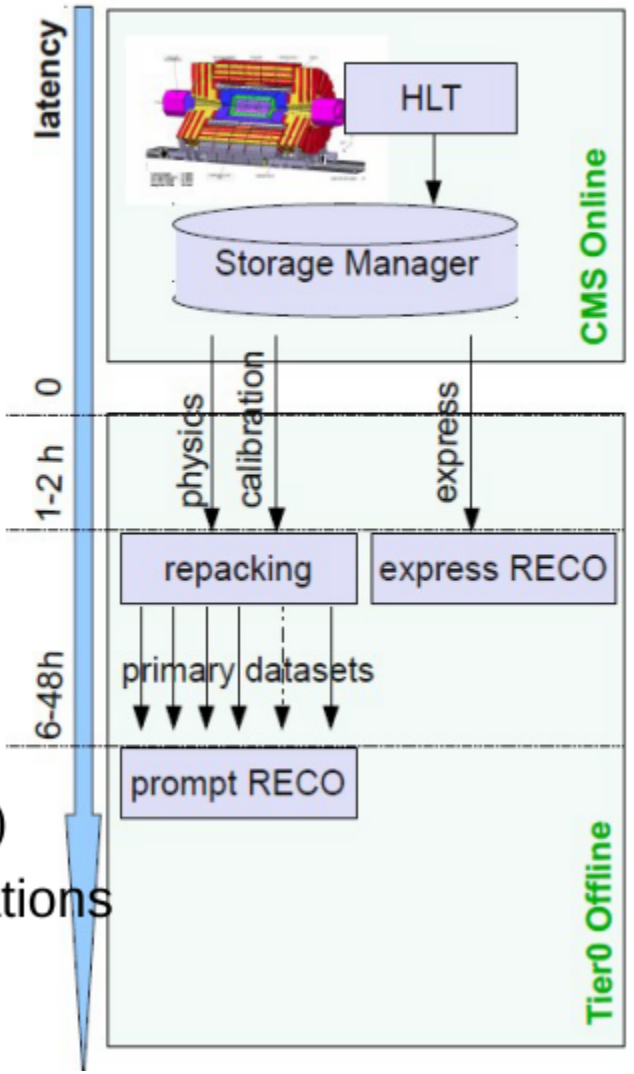


Central Processing @ CERN



Data Streams and Tier0 workflow

- Data streams & Tier0 workflows → specialized for different tasks
- Depending on the latency
  - express → prompt feedback & calibrations
    - short latency: 1-2 hours
    - ~40Hz bandwidth shared by:
      - calibration (½)
      - detector monitoring (¼)
      - physics monitoring (¼)
  - Alignment & Calibration (AlCa) streams
  - bulk data → sample for physics analysis (prompt reconstruction)
    - split in Primary Datasets (using High Level Trigger (HLT) decision)
    - will be delayed of 48h → get latest calibrations
    - writing ~300Hz



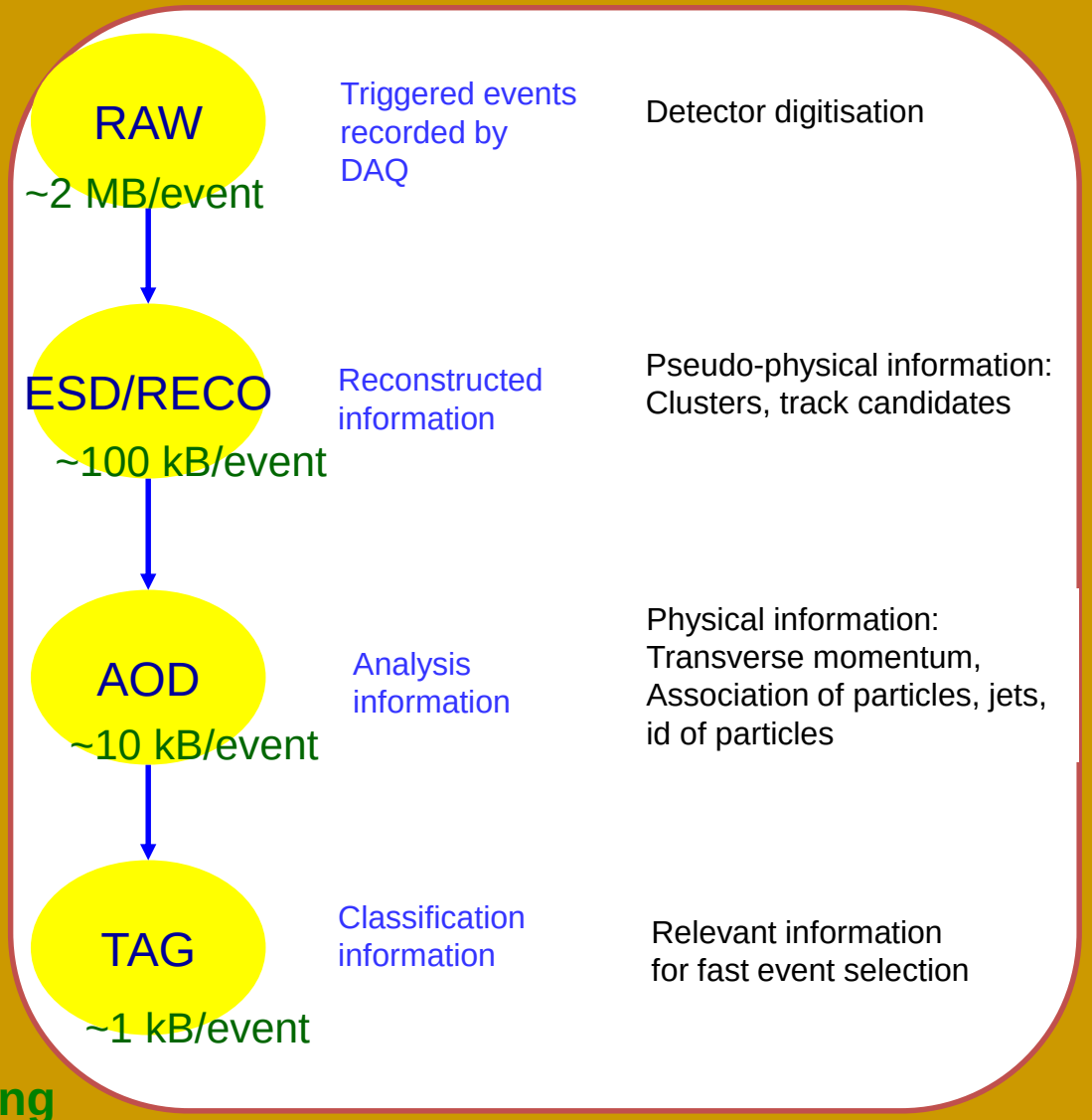
Data Tiers and Algorithms

HEP data are organized as *Events* (particle collisions)

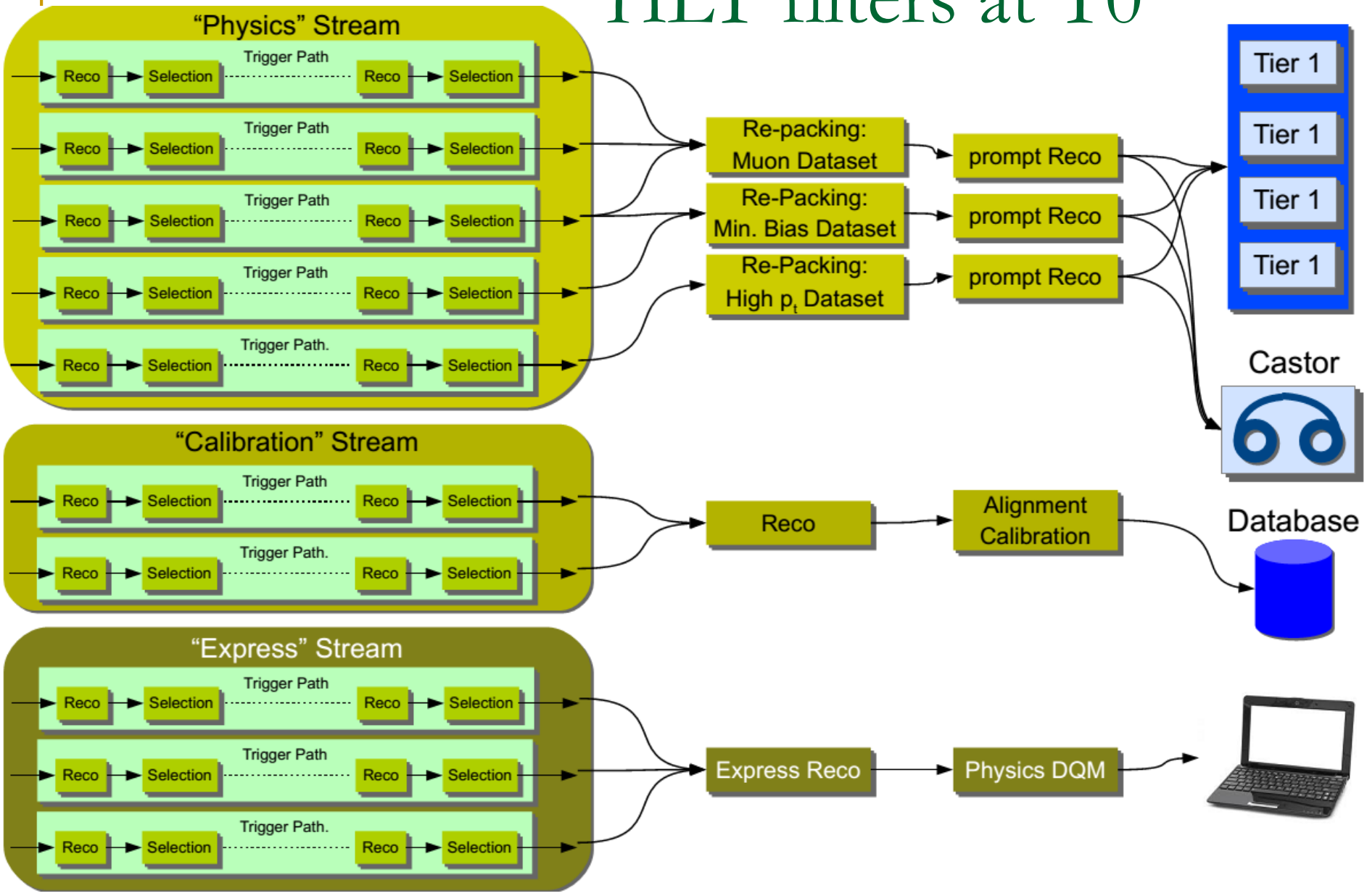
Simulation, Reconstruction and Analysis programs process "one event at a time"

Events are fairly independent
→ Trivial parallel processing

Event processing programs are composed of a number of Algorithms selecting and transforming "raw" event data into "processed" (reconstructed) event data and statistics

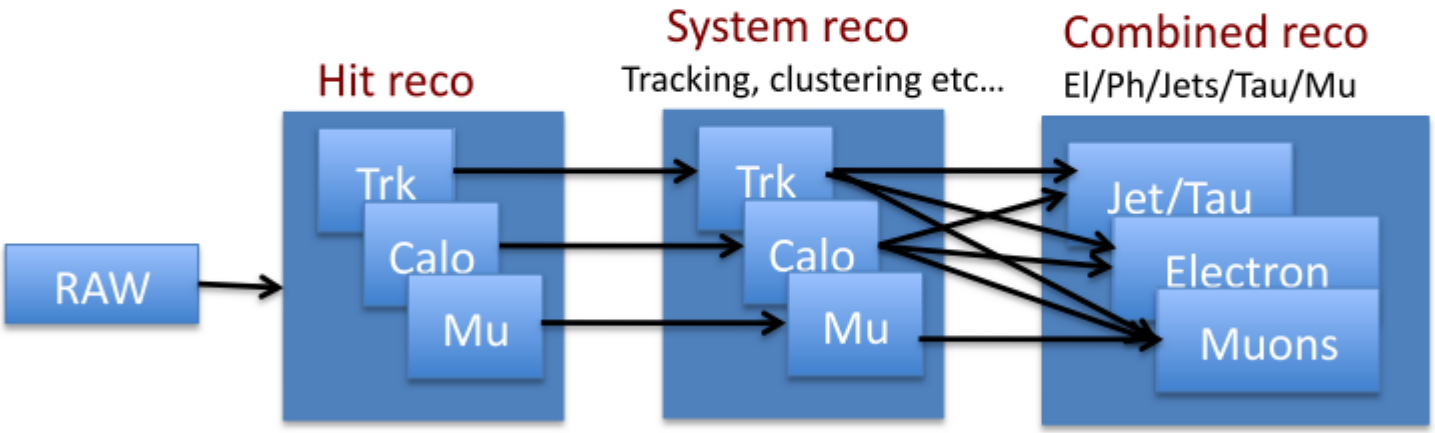


HLT filters at T0



Data processing - reconstruction

- The RAW data of triggered events are written to disk/tape
- This data is processed to produce outputs for physics analysis
  - The processing 'reconstructs' the data
    - RAW ADC counts -> detector 'hits'
    - Track and cluster finding
    - Physics object reconstruction (combining information from different detectors)
    - Applying calibrations and alignment in many of these steps



- Often the data is processed promptly at the Tier-0 and then reprocessed at a later time (with improved software and/or calibrations)

Calibration Workflows

- Provide most up-to-date conditions @ all stages of the data processing
- Different workflows depending on the time scale of updates:
  - quasi-online calibrations for HLT and express:
    - e.g. beam-spot → quick determination online
  - prompt calibrations: monitor/update conditions expected to vary run-by-run (or even more frequently):
    - updated conditions must be ready before prompt-reconstruction
  - offline re-reco workflows:
    - more stable conditions
    - workflows which need higher statistics:
      run on AlCa streams produced during prompt-reco or offline re-reco

Alignment & Calibration Streams

- All workflows fed using dedicated skims or datasets:
  - event selection tuned on the needs of the workflow
  - event content reduced to optimize bandwidth/disk space usage
- 2 kind of calibration streams:
  - produced directly @ HLT level
    - workflows statistically limited or requiring dedicated selection:
      - e.g. ECAL π^0 stream and ϕ-symmetry....
    - profit from High Level Trigger flexibility → software based
  - produced offline during express and prompt reconstruction (and offline re-processing)
    - just skimming events dedicated to calibrations

Data Processing Summary

ATLAS/CMS have similar data processing model

Prompt reconstruction using a calibration loop

- Processing of the physics data at Tier0 delayed by ~36/48hours to allow use of calibrations in the processing
- Means the output of prompt reconstruction is of high quality and can be used for physics analyses
  - Many physics papers published promptly processed data
  - in long term when the luminosity is stable and when we have more sophisticated calibrations may want to only publish papers based on reprocessed data
- Promptly processed data available for physics a few days after the data is taken

Data reprocessed with improved software and calibrations 1-2 times a year

In order to have consistent data and MC samples the MC needs to be reconstructed with the same release as the data

Stability of software very important to facilitate physics analysis

25

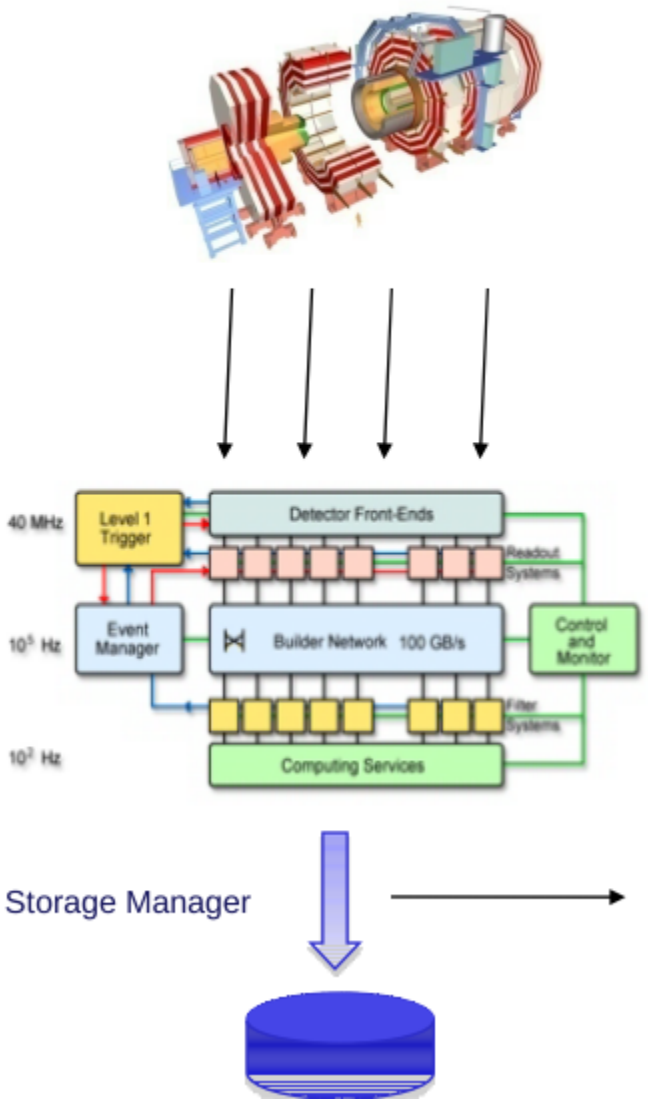
Data Certification

- The complexity of the offline workflows requires robust validation
- Several stages of Data Quality Monitoring (DQM):
  - online DQM → monitor detector performance during data-taking
    - dedicate event stream (sampling)
  - offline DQM → monitor performance of physics objects
    - runs on full statistics available for analysis:
      - express reco → fast feedback
      - prompt-reco → continuous monitor
      - offline re-reco → validation of software and condition updates
- Physics Validation Team → coordinates the validation activity.
  Feedback from:
  - groups responsible for physics objects
  - detector performance groups
  - analysis group

Data Quality - Introduction

- Data Quality (DQ) is the system for telling people **what data to use for physics analysis**
  - DQ also maintains a 'known problems' database
- Data can have bad quality because
  - Detector problem (dead channels, noise, data corruption)
  - Trigger / DAQ problem
  - Bad calibration / Reconstruction problem
- Data time granularity
  - ATLAS data is divided into 1 minute luminosity blocks (LB)
  - CMS use 23s lumi sections
- This is the time unit used for DQ and luminosity measurement
  - Eg. If a detector has a problem for 5 mins the corresponding LBs will be marked bad for physics for that detector
- DQ recorded for different systems separately
  - Can have a LB good for muons but not good for calorimeter
- DQ includes offline reconstruction and calibrations
  - Can recover some DQ efficiency in future data reprocessings

Online DQM



Online DQM: suite of CMSSW applications that run either on all events in the Filter Farm or on a selection of events served by the Storage manager

Since Dec2009 Online DQM consume DCS information in addition to Event data

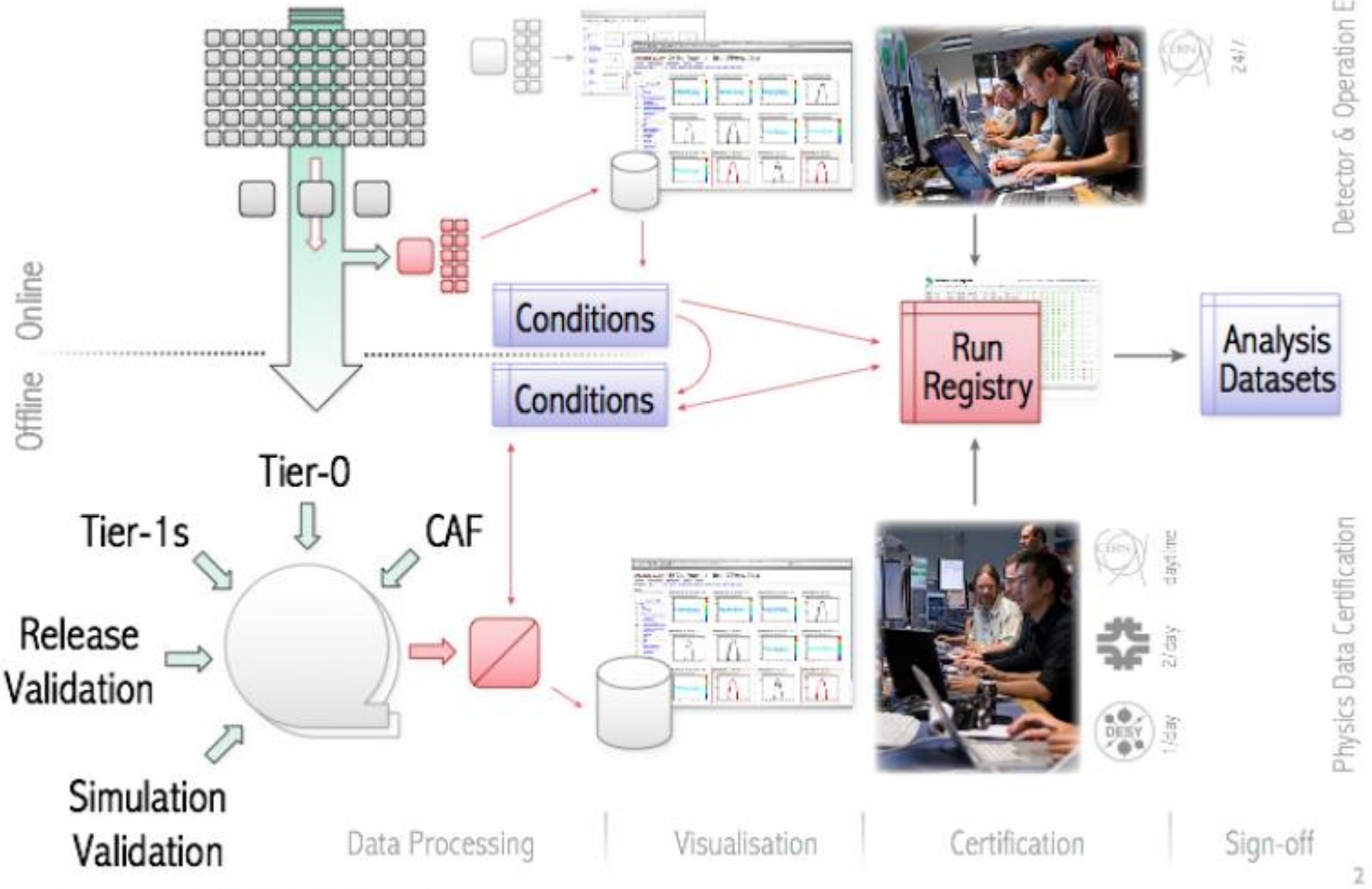
Online DQM Infrastructure is completed by the DQM data transfer system up to the DQM servers where the histograms and other DQM data are uploaded and visible to the shifters and the CMS community

Scope of Online DQM Shifts:

- Identify problems with detector performance or data integrity during the run

Online

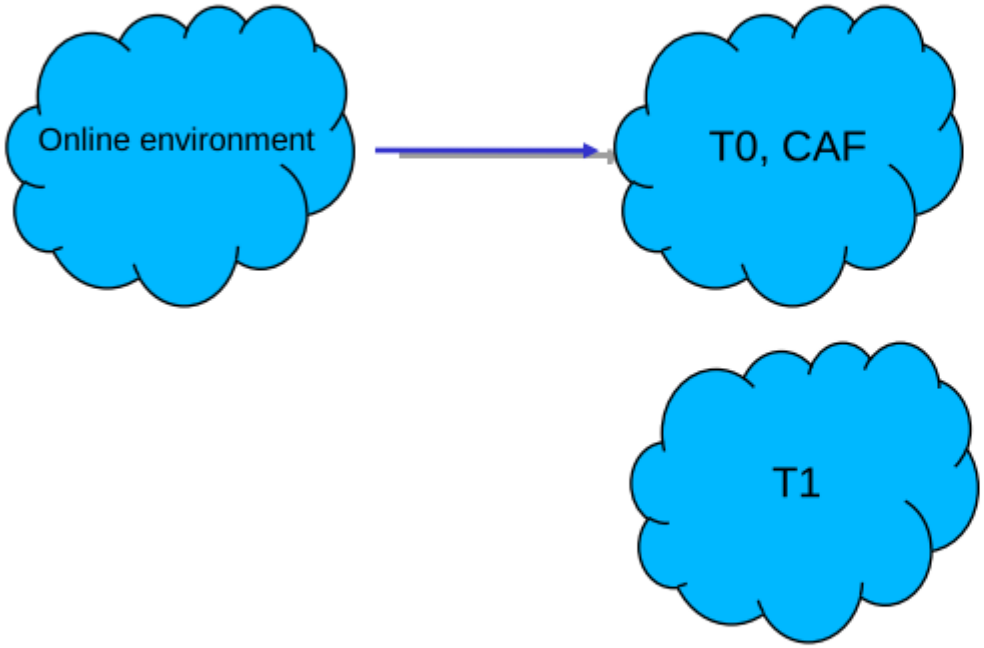
DQM end to end



Scope of Online DQM Shifts:
Identify problems with detector performance or data integrity during the run
➔ SPOT PROBLEMS QUICKLY FOR OPTIMAL OPERATION EFFICIENCY

GUI – Summary Workspace

Offline Data Processing and Offline DQM



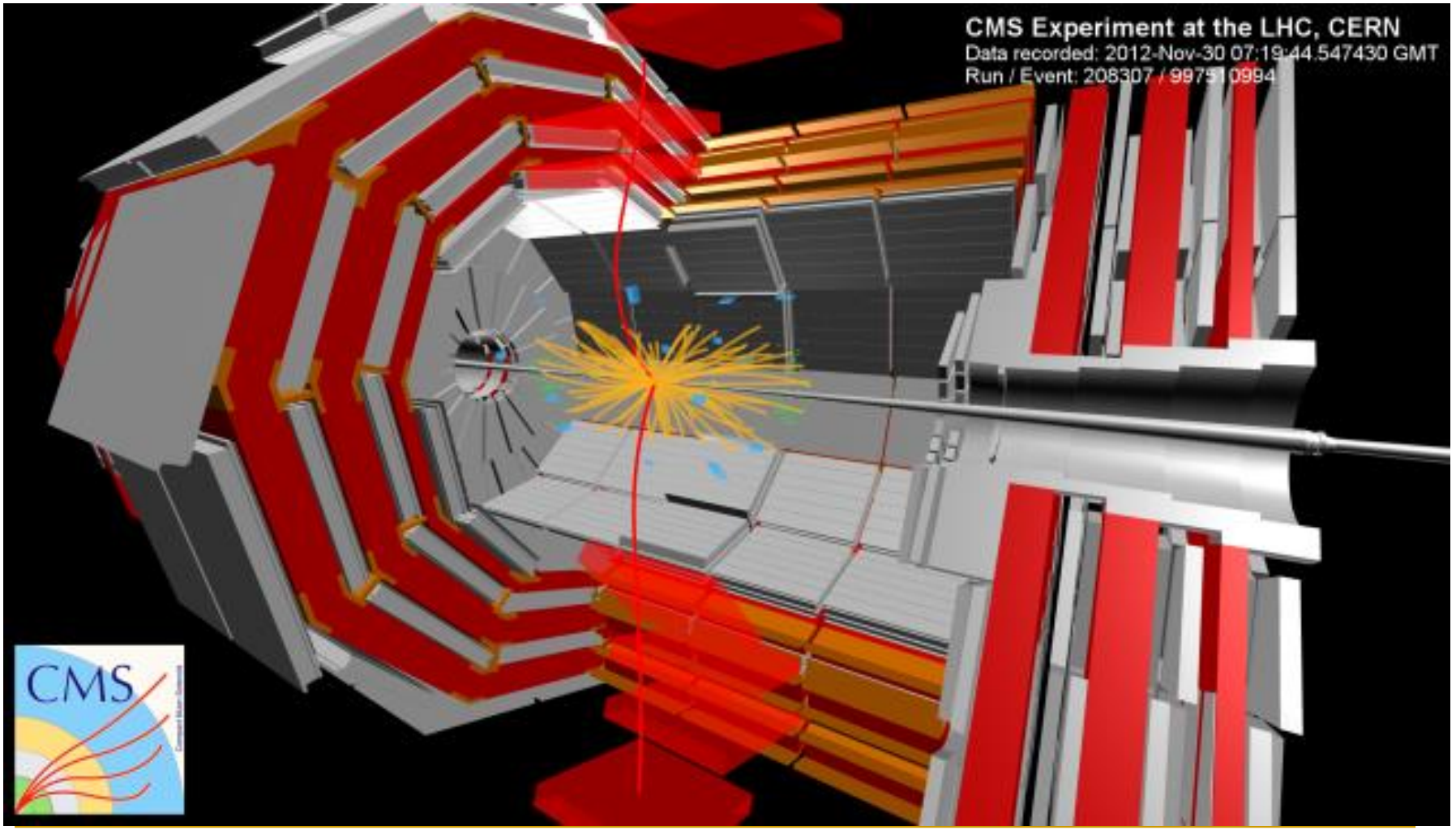
Prompt Reconstruction at T0 and CAF is performed from within one hour up to 48 hours after data is transferred from P5 to T0 and CAF (CERN)

Subsequent iterations of re-reconstruction at the T1's follow periodically the Prompt Reco with improved Alignment and Calibration constants, bug fixes.

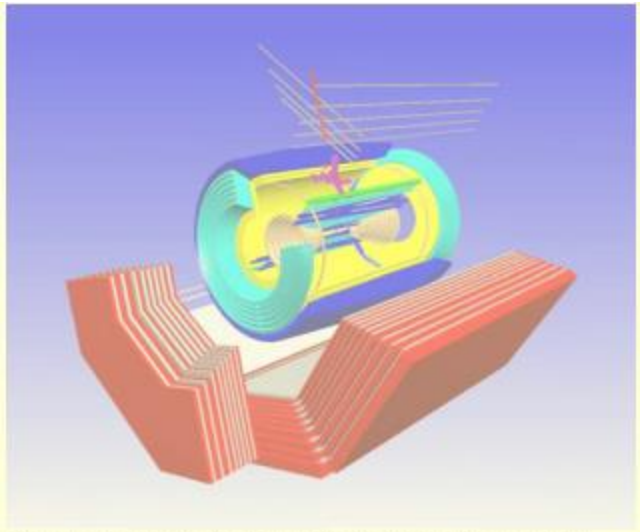
Offline DQM is part of the Offline data processing that, in addition to detector data analyses, includes higher level reconstruction objects, aka Physics Objects (POG's)

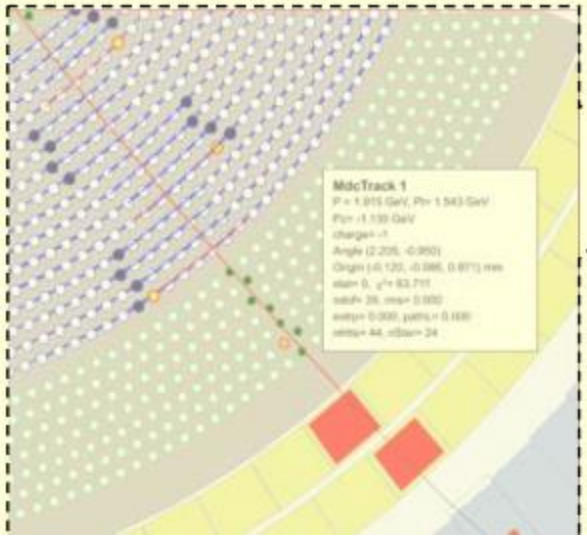
Scope of Offline DQM Shifts: Produce the data certification for various reconstruction iterations ➔ USED FOR CMS OFFICIAL GOOD RUN LISTS!!!

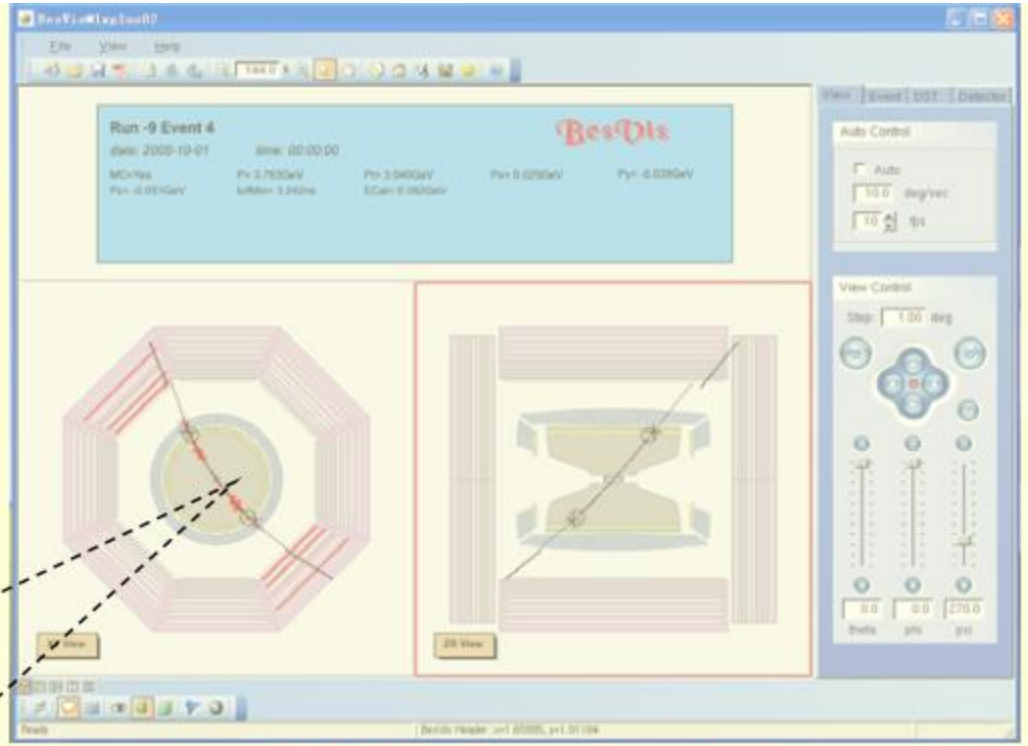
Event Display: $B_s^0 \to \mu^+\mu^-$



Event Display: online and offline







- Based on ROOT, OpenGL, X3D and XML
- Support both 2D and 3D view
- Operations and controls available through menu and toolbar items

Further References:

- Bi-annual CHEP:
  - http://chep2015.kek.jp/
  - http://www.chep2013.org/
- Annual TWEPP conference:
  - http://www.lip.pt/events/2015/TWEPP/ 2015(in Sep)
  - https://indico.cern.ch/event/299180/overview 2014
- CMS TriDAS TDR.
  - V1: CERN-LHCC-2000-038 ; CMS-TDR-6-1
  - V2: CERN-LHCC-2002-026 ; CMS-TDR-6
- ATLAS TriDAS TDR
  - V1: CERN-LHCC-1998-014, ATLAS-TDR-12
  - V2: CERN-LHCC-2003-022, ATLAS-TRD-016
- ISOTDAQ: the international school of trigger and data acquisition
  - http://isotdaq.web.cern.ch/isotdaq/isotdaq/Home.html