# Experimental techniques in high-energy nuclear and particle physics

"Dottorato di Ricerca in Ingegneria dell'Informazione"

LECTURE 8.

Prof. Rino Castaldi INFN-Pisa

rino.castaldi@pi.infn.it

# The CERN Large Hadron Collider







 Higgs : Clarify the origin of the spontaneous symmetrybreaking Standard Model (-> Higgs, SUSY)

- New forces (symmetries)
- New particles
- Super symmetries
- Substructures

9300 Superconductor magnets 1232 Dipoles (15m,1.9<sup>o</sup>K) 8.4Tesla 11700 A 448 Main Quads, 6618 Correctors. Circonference 26.7 km





### Signal and background $\rightarrow L=10^{34} \text{ cm}^{-2}\text{s}^{-1}$

Event rate = LoBrCross sections for various physics processesvary over many orders of magnitudeHiggs (600 GeV/c<sup>2</sup>): 1pb @10<sup>34</sup>  $\rightarrow$ 10<sup>-2</sup> HzHiggs (100 GeV/c<sup>2</sup>): 10pb @10<sup>34</sup> $\rightarrow$ 0.1 Hzt t production: $\rightarrow$ 10 Hz $\forall \rightarrow \ell \vee$ :Inelastic: $\rightarrow$ 10° Hz

Selection needed: 1:10<sup>10-11</sup> Before branchina fractions...





 $\Rightarrow$  Needle in a Hay Stack



### **Data detection and data filtering**





#### LHC: a very hostile environment for DAQ (high event rate, high multiplicity, high radiation flux)



#### **Collisions at LHC**



Proton-Proton	2835 bunch/beam
Protons/bunch	10 <sup>11</sup>
Beam energy	7 TeV (7x10 <sup>12</sup> eV)
Luminosity	10 <sup>34</sup> cm <sup>-2</sup> s <sup>-1</sup>

Crossing rate 40 MHz

Collisions rate ≈ 10<sup>7</sup> - 10<sup>8</sup>Hz

New physics rate ≈ .00001 Hz

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

### **Detectors at LHC**



# Each layer identifies and enables the measurement of the momentum or energy of the particles produced in a collision







### **Bunch Crossing Times: LEP, Tevatron & LHC**





- LHC has ~3600 bunches (2835 filled with protons)
- And same length as LEP (27 km)
- Distance between bunches: 27km / 3600 = 7.5m
- Distance between bunches in time: 7.5m / c = 25ns

**LEP:** e<sup>+</sup>e<sup>-</sup> Crossing rate 30 kHz



### pp cross-sections and minimum bias





# Pile-up

In-time" pile-up: particles from the same crossing but from a different pp interaction

- Long detector response/pulse shapes:
  - "Out-of-time" pile-up: left-over signals from interactions in previous crossings
  - Need "bunch-crossing identification"









#### Detectors



Detector	Channels	Control	Ev. Data
Pixel	6000000	1 GB	50 (kB)
Tracker	1000000	1 GB	650
Preshower	145000	10 MB	50
ECAL	85000	10 MB	100
HCAL	14000	100 kB	50
Muon DT	200000	10 MB	10
Muon RPC	200000	10 MB	5
Muon CSC	400000	10 MB	90
Trigger		1 GB	16

Event size Max LV1 Trigger Online rejection System dead time 1 Mbyte 100 kHz 99.999% ~%

### **Crossing and Event Rates**



# Event selection: The trigger system

Mandate:

"Look at all bunch crossings, select most interesting ones, collect all detector information and store it for off-line analysis"

P.S. For a reasonable amount of CHF



Since the detector data are not all promptly available and the function is highly complex, T(...) is evaluated by successive approximations called :

### **TRIGGER LEVELS**

(possibly with zero dead time)

# ON/OFF-line data flow and computing model

CERN

 $\sigma$  LHC  $\sqrt{s}$ =14TeV L=10<sup>34</sup>cm<sup>2</sup>s<sup>-1</sup> Event Rate

CMS



S. CITTOIIN. CERN/CMS/RI2009







### Collision rate 10<sup>9</sup> Hz

Channel data sampling at 40 MHz

#### Level-1 selected events 10<sup>5</sup> Hz

**Particle identification** (High  $p_T e, \mu$ , jets, missing  $E_T$ )

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

#### Level-2 selected events 10<sup>3</sup> 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 10..100 Hz

#### Physics process identification

• Event reconstruction and analysis

### **DAQ structure: constraints & requirements**











LV1





### Reduce event rate (to be readout) from 40 MHz to 10<sup>5</sup> Hz

### Level-1. Calorimeters and Muons



# Level-1 Muon: CMS

- Level-1  $\mu$ -trigger info from:
  - Dedicated trigger detectors: RPCs (Resistive plate chambers)
    - Excellent time resolution
  - Muon chambers with accurate position resolution
    - Drift Tubes (DT) in barrel
    - Cathode Strip Chambers (CSC) in end-caps
  - Bending in magnetic field $⇒ determine p_T$



### **Local Track finding**





RPC: pattern recognition using pre-calculated pattern table



DT: Meantimers form track vectors Correlation engine combines tracks into segments

CSC: Pattern comparator with half-strip resolution. Hits combined to form track segment

# Level-1 Calo: Isolated electron at CMS

- Electromagnetic trigger based on 3 × 3 trigger towers
  - Each tower is 5 × 5 crystals in ECAL (barrel; varies in end-cap)
  - Each tower is single readout tower in HCAL



- Isolation in both FCAL and HCAL sections

two towers

# **CMS Level-1**

- Information from calorimeters and muon detectors
- Custom-built electronics for trigger processors (ASICs, FPGAs)
- Synchronous, pipelined
  - Processing logic: 25 ns pipelined system
  - Must work dead time free
  - Latency: < 3.2 μs (128 bx)</li>
  - readout + processing: <  $1\mu$ s
  - signal collection + distribution:  $\sim 2\mu s$
- Max. output rate: 100 kHz
- Organized in 3 subsystems:
  - Muon Trigger, Calorimeter Trigger, Global Trigger
- Backgrounds are huge
  - Large rejection factor: 40 MHz (×20 events/crossing)  $\rightarrow$  100 kHz
  - Rates: steep functions of thresholds







40 MHz digitizers and 3.2µs x 25ns step **pipeline readout buffers** 40 MHz Level-1 trigger (**massive parallel pipelined processors**) High precision (~ 100ps) **timing, trigger and control distribution** 

# **ATLAS Level-1**



# Level 1 Muon: ATLAS



The Level-1 trigger logic is almost fully programmable; this flexibility will allow to optimize carefully the signal trigger efficiency vs. the background rejection.

- RPC in barrel regions
  - 3 stations
  - 430,000 channels
- TGC (Thin Gap Chambers) in end-cap regions
  - 3 stations
  - 800,000 channels
- Coincidence logic ( $\eta$  and  $\phi$ )
- Two p<sub>T</sub> threshold ranges
- Low p<sub>T</sub> (6 10 GeV):
  - Require hits in 3 out of 4 layers in inner two stations
- High p<sub>T</sub> (8 35 GeV):
  - Require hits in 3 out of 4 layers in inner two stations
  - Require hits in 1 out of 2 layers of the outer station (2 out of 3 in the end-caps)

### **Detector front-end structure**



# E.g. CMS silicon strip front-end





# **Front-end synchronization**





- LHC has ~3600 bunches (2835 filled with protons)
- And same length as LEP (27 km)
- Distance between bunches: 27km / 3600 = 7.5m
- Distance between bunches in time: bx =7.5m / c = 25ns
- Apparatus dimensions 30-70 m -> 5...7 bx



### Timing and Trigger and Event kinematics





### And things are not always simple





Revolution time Revolution frequency RF frequency Bunch crossing rate No of bunches/beam Filling factor Bunch train length SPS injection kicker gap LHC injection kicker gap LHC extraction kicker gap RMS bunch length RMS collision length Interbunch spacing 88.924 μs 11.246 kHz 400.8 MHz (2 x SPS) 40.08 MHz 2835 0.795 81 220 ns 950 ns 3.17 μs 0.075 m 0.053 m, 177 ps 7.5 m, 25 ns



- The massive data rate after LVL1 poses problems even for network-based event building — different solutions are being adopted to address this, for example:
  - In CMS, the event building is factorized into a number of slices each of which sees only a fraction of the rate
    - Requires large total network bandwidth (⇒ cost), but avoids the need for a very large single network switch
  - In ATLAS, the Region-of-Interest (Rol) mechanism is used with sequential selection to access the data only as required – only move data needed for LVL2 processing
    - Reduces by a substantial factor the amount of data that need to be moved from the Readout Systems to the Processors
    - Implies relatively complicated mechanisms to serve the data selectively to the LVL2 trigger processors ⇒ more complex software

# **The Main Challenge: Event Building**



PC motherboards for data Source/Destination nodes

Three Major issues: 1) Maximize link utilization (CHF); 2) Avoid bottleneck at outputs; 3) Assemble large number of ports







Total 99.999 % rejected 0.001 % Accepted



### **CMS DAQ baseline structure**





Collision rate		40 MHz
Level-1 Maximum trigger rate		100 kHz
Average event size	~	1 Mbyte
Flow control&monitor	≈	10 <sup>6</sup> Msg/s

Readout concentrators/links Event Builder bandwidth max. Event filter computing power Event Builder GBE ports Data production Processing nodes

512 x 4 Gb/s 2 Tb/s

- ≈ 10 TeraFlop
- > 4000
- ≈ Tbyte/day
- ≈ x Thousands



### **DAQ: scaling x 8 units**





#### Data to surface:

1 Mbyte Average event size No. FED S-link64 ports 700 DAQ links (2.5 Gb/s) Event fragment size 2 kBFED builders (8x8 dual) 80 Technology(2004) Myrinet

# 512+512

#### **Readout Builder (x8):**

Lv-1 max. trigger rate 12.5 kHz GBE RU Builder 80x80x2 .125 Tb/s Event fragment size 16 kB **RU/BU** systems 80 Event filter power ~ TeraFlop **EVB** technology Ethernet

# **ATLAS DAQ Baseline Structure**



# **The Main Challenge: Level 2**

### Driven by Level-1 information

- Crucial parameters: data routing and CPU time (latency)
- ROIBuilder: custom hardware to combine ROI pointers
- Supervisor farm: collect info, allocate event to processor, distribute result to ROBs
- Processor farm: collect data from ROBs (upon request) execute algorithm; decision to Supervisor farm



# CMS trigger levels and DAQ architecture



# **CMS High Level Trigger**

### • Rejection:

1:1000 selection

### • Algorithms:

- Algorithms can almost be as sophisticated as in offline analysis
- Avoid unnecessary calculations; reject as soon as possible
- Hence, internal "logical" trigger levels:
  - Level-2: use calorimeter and muon detectors
  - Level-2.5: also use tracker pixel detectors
  - Level-3: use of full information, including tracker
- In principle continuum of steps possible
- Regional reconstruction: e.g. tracks in a given road or region

### Resources/CPU time:

- 100 kHz  $\rightarrow$  10  $\mu s/event$
- If  $T_j$  is the time taken by the Level-J decision (J=2,3,...) and the rejection factors are  $R_J$

$$T_{tot} = T_2 + T_3 / R_2 + T_4 / (R_3 R_2) + \dots$$

• A 50 kHz system at startup will need ~2000 CPU's





### LHC trigger and DAQ summary





S. Cittolin. CERN/CMS/RT2009



### **Evolution of DAQ structures with technologies**





#### PS/ISR/SPS:1970-80: Minicomputers

Readout custom design First standard: CAMAC Software: noOS, Assembler • kByte/s

Detector Readout		
Event building		
On-line processing		
Off-line data store		



#### p-p/LEP:1980-90: Microprocessors

HEP proprietary (Fastbus), Industry standards (VME Embedded CPU, servers Software: RTOS, Assembler, Fortran • **MByte/s** 





#### LHC: 200X: Networks/Clusters/Grids

PC, PCI, Clusters, point to point switches Software: Linux, C,C++,Java,Web services Protocols: TCP/IP, I2O, SOAP,

• TByte/s



S. Cittolin. CERN/CMS/RT2009



### **Trigger and data acquisition trends**

CERN





#### Cessy: Master&Command control room



#### Fermilab: Remote Operations Center



#### **Meyrin: CMS DQM Center**



#### CR: Any Internet access.....







#### O(10) Luminosity increase (2014-16) will require:

•New front-end electronics and readout links

- •Higher level-1 selection power (to maintain 100 kHz max. output)
- •Event builder (>10 Tb/s) with an order of magnitude higher



#### 2014 SLHC. A possible DAQ design:

A multi-Terabit/s network congestion free and scalable (as expected from communication industry). In addition to the Level-1 Accept, LV1 Trigger transmits to the front-end readout units additional information such as the event type and the event destination IP address that is the processing system (CPU, Cluster, TIER..) where the event has to be eventually built and analyzed.

The event fragment delivery and therefore the event building will be warranted by the network protocols and (commercial) network internal resources (buffers, multipath, network processors, etc.)

**Real time buffers of Pbytes** temporary storage disks will cover a real-time interval of days, allowing to the event selection tasks a better exploitation of the available distributed processing power. More Slides

# Taus

no stop mixing  $m_{etop} = 1 \text{ TeV}$ 

800

900

1000





#### ALICE<sup>(Pb-Pb)</sup> trigger levels and DAQ architecture σ LHC √s=14TeV L=10<sup>34</sup>cm<sup>-2</sup>s<sup>-1</sup> Event Rate barn GHz — σ inelastic LVO mb - bb MHz LV1 Collisions rate 4 kHz LV2 LV-0 μb 4 us •First Level~ 500 Hz 🏓 kHz 100µs **2** Storage ytels Rate on tape ~50 Hz Readouters HĽ. • $Z \rightarrow \ell v$ HLI nb • tt Hz SUSY qq+qg+qg $gg \rightarrow H_{SM}$ $\tan\beta = 2.\mu = m_{\tilde{a}} = m_{\tilde{a}}/2$ anß=2,u=m<sub>2</sub>=m ALICE: aā→qāH\_M pb **Trigger Levels** $H_{SM} \rightarrow \gamma \gamma$ Z<sub>ARL</sub>→2ℓ mHz $h \rightarrow \gamma \gamma$ Level-1 rate 100 Hz **Readout bandwidth GB**/s 5 fb Storage bandwidth **GB**/s ABSORRER 1.2 μHz scalar LQ $Z_n \rightarrow 2$ ∠<sub>SM</sub>→3γ ns μs ms year hour sec 500 1000 2000 10-6 10<sup>-3</sup> 10-0 10<sup>3</sup> 10<sup>6</sup> sec 50 100 200 10-9 jet E<sub>T</sub> or particle mass (GeV) **ON-Line OFF-Line**





#### 12 yeas of R&D (too much?)

the project has lasted more or less a man generation from design to implementation...

#### **Custom/Standards attention**

Pay attention to maintenance and replacement issues. Survey new standards in the field of telecommunication, server packages, data centers, cooling etc.

#### Will we buy computing power and network bandwidth?

New kind of commodities: CPU power, memory, mass storage and bandwidth are becoming commercial products..

#### **Configuration and control of complexity is an issue**

Data taking efficiency depends on the real-time system performances but also on the prompt handling of all on-line resources. Configuration, fault tolerance, remote controls, monitor and security have to be improved.

#### Super LHC DAQ upgrade

The best R&D for the SLHC DAQ upgrade will come from the completion and operation of the current systems

#### Moore law "confirmed"

High energy physics experiments still provide "exceptional challenges" to data processing and data communication technologies, however in the majority of cases somebody else is driving the locomotive.

# Level-1 trigger table (low/high lumi)

Low Luminosity	Trigger	_Threshold _ (ε=90-95%) (GeV)	₋Indiv. ₋Rate (kHz)	_Cumul rate(kHz)
Total Rate: 50 kHz	_1e/γ, 2e/γ	<i>_</i> 29, 17	_4.3	_4.3
Eactor 3 safety	_1μ, 2μ	_14, 3	-3.6	_7.9
	$-1\tau$ , $2\tau$	-86, 59	-3.2	-10.9
allocate 16 KHZ	_1-jet	_177	-1.0	_11.4
	_3-jets, 4-jets	-86, 70	-2.0	_12.5
	₋Jet * Miss-E <sub>T</sub>	_88 * 46	-2.3	_14.3
	₋e * jet	_21 * 45	-0.8	_15.1
	-Min-bias		-0.9	_16.0

High Luminosity	_Trigger	_Threshold _(ε=90-95%) (GeV)	₋Indiv. ₋Rate (kHz)	_Cumul rate (kHz)
Total Rate: 100 kHz	_1e/γ, 2e/ γ	_34, 19	_9.4	_9.4
Factor 3 safety, allocate 33.5 kHz	_1μ, 2μ	-20, 5	_7.9	_17.3
	-1τ, 2τ	<b>_101, 67</b>	-8.9	_25.0
	_1-jet	-250	-1.0	-25.6
	_3-jets, 4-jets	_110, 95	_2.0	-26.7
	_Jet * Miss-E <sub>⊤</sub>	_113 * 70	_4.5	_30.4
	₋e * jet	_25 * 52	_1.3	_31.7
	-μ * jet	_15 * 40	_0.8	-32.5
	-Min-bias		_1.0	_33.5

# **Technologies in Level-1 trigger systems**

- ASICs (Applicatio-Specific Integrated Circuits) used in some cases
  - Highest-performance option, better radiation tolerance and lower power consumption (a plus for on-detector electronics)
- FPGAs (Field-Programmable Gate Arrays) used throughout 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

### **DAQ structures at LHC**





### 1978-1989. UA1 DAQ system (0.1 MHz, 30 Hz)





1981-84

- Remus data acquisition (≈200 CAMAC crates)
- rate on tape ≈ 1Hz (event size ≈100 Kbyte)

Rates (Hz) 10<sup>5</sup> Trigger Muon Central detector Hadron Streamer TMP BATS Level 1 30 **Data reduction** Level 2 Readout Supervisor Parallel read-out (30 units, 60 Mbytes/sec) Multi-event buffer 10 VME DMA event readout ( 8 Mbyte/sec) VME 680xx, Macintosh Event multiprocessor system Manager VME Event builder **Event filter** 3081E Hadron Muon Event 3081E monitor monitor display 3081E 3 (20) VME monitor Mass storage **DAQ** console **IBM 3840** Trigger control....

1985-1989 VME, IBM-emulators, Desktops

#### Proprietary/Standards: CAMAC, embedded µp, emulators, VME