CERN SCHOOL OF COMPUTING

Sopron, Hungary 9 September 1994

Short History & New Frontiers

п

n, Hungary & September 1994	$\lambda - H/p$		/ p	$T \simeq t^{-1/2}$	1900	
		10 ⁻¹⁰ m	< KeV	>300.000 Y	Quantum Mechanics Atomic Physics	
	γ e ⁺ γ			≈ 300 sec	1940-50 Quantum Electro Dynamics	
of Data Acquisition for LHC S. Cittolin/CERN-ECP	e S	10 ^{.15} m	∝ GeV		1950-65 Nuclei, Hadrons Symmetries Field theories 1965-75	
 Introduction LHC experiments Data acquisition structures Frontend Level-1 trigger Fast controls Readout Event Builder 		10 ⁻¹⁶ m	> GeV	$\approx 10^{-6}$ sec	Quarks Gauge theories	
	ever ever ever ever ever ever ever ever	10 ^{.18} m	100 GeV	$\approx 10^{-10}$ sec	SPS, pp 1970-83 ElectroWeak Unification, QCD	
	Charge 6-3 Quarks 6 L 26.3 U 17.3 C 3 "Colours" C	eptons Charge ν _e 0 e -1 ν _μ 2 "Flavours"			LEP 1990 3 families	
- High trigger levels - Conclusion		μ in each v _τ doublet τ			Tevatron 1994 Top quark	
	Higgs ? Supersyn	nmetry ?			LHC	
		10 ^{.19} m	≈ TeV	$\approx 10^{\cdot 12} \; sec$	Origin of masses The next step	
	GRAND Unified T	heories ? 10 ^{.31} m	10¹³ TeV	$\approx 10^{.38}$ sec	Underground Labs Proton Decay ?	
	Quantum Gravity? Superstrings ?					
		10 ⁻³⁴ m	~ 10¹ ⁶ TeV	≃ 10 ^{.42} sec	?? The Origin of the Universe	

1

Overview of Data Acqui

• HIGGS

Electroweak symmetry breaking Origin of masses



of massive objects. The hadron colliders can provide the exploratory physics with high constituent \sqrt{s} and with high luminosity, but at the expense of clean experimental conditions

ſ

LHC project

Two superconducting magnet rings in the LEP tunnel.





LHC detectors and experiments



ATLAS A Toroidal LHC ApparatuS



CMS Compact Muon Solenoid



Selection rate 1/10000000000000

I .

LHC parameters

	LEP	SppS	HERA	LHC
BX period	22 µs	3.8 µs	96 ns	25 ns
Interactions/BX	<< 1	~1	<< 1	~ 20
% Intn to select	100 %	0.1 %	100 %	0.001%
# calo towers	~ 10⁴	~ 10 ³	~ 104	~ 10⁵
# tracking chan.	~104	~104	~10⁴	~107

302



t i

Compact Muon Solenoid. CMS



CMS central tracking

Central tracking at L = 1.8 10^{34} (50 ns integration, \approx 2000 tracks)



ĩ

Data acquisition structure



- Trigger levels
- Trigger steps in present colliders
- Multilevel data acquisition

- Second level trigger
- CMS data acquisition

Rates from background of physics triggers.

Тор		
l + jets	p, ^µ > 40 GeV	200 Hz
	p,e > 40 GeV	5 •10 ⁴ Hz
e+μ	p, ^μ > 50 GeV	100 Hz
	$p_t^e > 50 \text{ GeV}$	400 Hz
Higgs		
41	p,²μ > 20 GeV	25 Hz
	p ^{2e} > 20 GeV	10 ³ Hz
e v j j	$p_{t}^{e} > 100 \text{GeV}$	15 Hz
SUSY		
p, m+jets	p _t 3i> 200 GeV	500 Hz
4 I +jets	p ₁ ^{2µ} > 30 GeV	10 Hz
	p _t ^{2e} > 30 GeV	10 ² Hz

LVL-1 rate 104 - 105 Hz

ст<u>я</u> — с

Event selection

The trigger is a function of :



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)

The final aim is to look for \approx 1/1000000000000

п



10..100 Hz

]]



· · ·



Multi-level readout and analysis system

i .

Standard DAQ layout

П

n

CMS data acquisition

CMS two physical levels

- Reduces the number of building blocks simpler design, easier maintenance and upgrades
- · Simplifies the data flow

Line and

• Exploits the commercial components 'state of the art' memory, switch, CPU

• Upgrades and scales with the machine performances flexibility in logical redistribution of resources

• Makes full use of the computing power anyway needed for the off-line analysis

Technology ansatz

• The CPU processing power increases by a factor 10 every 5 years (at constant cost)

• The memory density increases by a factor 4 every two years (at constant cost)

• The 90's are the data communication decade

CMS data acquisition parameters

Number of channels and data volumes (at 10³⁴ luminosity)

			• ·
Detector	No. Channels	Occupancy%	Event size (kB)
Pixel	80000000	.005	100
InnerTracker	16000000	3	700
Preshower	512000	10	50
Calorimeters	250000	10	50
Muons	1000000	.1	10
Trigger			10
Level-1 trigger rate			100 kHz
Level-1 trigger rate			100 kHz
No. of Readout units (200-5000 Byte/event)			1000
Event builder (10	00•1000 switch) bar	ndwidth	500 Gb/s (*)
Event filter computing power			5•106 MIPS
Data production			Tbyte/day
No. readout crates			300
No. electronics boards			10000

(*) In order to achieve the data acquisition II jure of 100 kHz event rate after the level 1 trigger, this tracking data must not be moved into the readout network until the associated event has passed the test of the high trigger levels based on the information from the other detectors. This operation (called until level-2) is expected to reduce the event rate (for the tracker data) by at least one order of migratude.

CMS data acquisition subsystems

• Detector frontend

• First level trigger

Readout network

Fast controls

Event builder

• Event filter

 Subsystems and functions:

 RC
 Readout Crate.

 Event Builder.
 Multiport switch network

 EFC
 Event Flow Control.
 Event scheduling and filter task control

 FFI
 Filter Fram Interface.
 Event data assembly into processor memory. The system includes the farm status communication with the EFC and the control. FFI functions may be part of the future farm computer architecture.

 CPU
 Event filter processing unit. It may be a single workstation board or a unit in a multiprocessor server.

 Fast Control.
 Trigger and EFC signals broadcasting to readout crates.

 Control.
 System test,initialization, monitoring etc.

 Services.
 Control room analysis, display and monitoring consoles, WAN connections, mass storage and data archives

Π

. -

Frontend functional model

11

ĩ

ſĨ

312

i ,

313

Large dynamic range (15bits) Digital filtering Radiation hard High power consumption

Silicon strip detector. RD-20

Slow amplifier

Analog memory

• Bunch crossing recovery

Pileup removal

[]

. 1

Analog to digital conversion

Digitizing means measuring something (charge, amplitude, time..) that is compare it with a reference unit.

1128-1-1 (No) (L.B.)

Entity to be measured Ruler unit

Compare entity with a series of rulers in sequence (standard ADC, counting) or in parallel (FADC, decoding)

Compare entity with a ruler, subtract ruler from entity (Sigma Delta ADC)

Compare entity with a ruler then subtract ruler from entity and halve the result (HARP pipeline ADC)

١

Compare entity with a series of rulers in sequence (standard ADC, counting) or in parallel (FADC, decoding)

Flash ADC

Dynamic range Resolution " Power consumption Speed critical for Calorimeter (15-16 bits) Calorimeter (9-10 bits) "Inner detectors "All (n• 67MHz n≥1)

Ш

314

-1

Sigma delta

Compare entity with a ruler, subtract ruler from entity (Sigma Delta ADC)

Sigma Delta ADC

Pipeline conversion

Compare entity with a ruler then subtract ruler from entity and halve the result (HARP pipeline ADC)

ADC pipeline (CERN-LAA HARP) 12 Bits 1MHz

Time Memory Cell (KEK/Hiroshima/NTT)

Industry

In the last 20 years, the requirements of telecommunication and in particular to television have strongly contributed to the development of standard technology (CMOS, BiCMOS) and mass production by industry.

Together with other fields high energy physics experiments have exploited these developments extensively.

316

Flash ADC Analog memory Personal computers Helical scan recording Data compression Image processing Cheap MFlops for image synthesis

In recent years the world television industry has undertaken a new challenge :

High Definition TeleVision (HDTV).

This represents a tremendous effort of research and development in the field of standard technology. In Europe it is organized under the project EUREKA 95.

- HDTV chain
- Digital
- Data compression
- Multi-standards

• Transmission

- High quality pictures
 - Digital signal processor
 Memory delays
 - Pipeline and data driven architectures
 - High speed (2 GB/s) optical links

- ADC (144, 27 MHz, 10 bits)

- Mass storage
- Parallel to serial converterHigh speed (100 MB/s) recorders

ADC 288 MB/s Diagonal Filter	
Parallel to Serial 144 MB/s	
≫ ^{<™} 1.3 Gbit/s	
Serial to Parallel 144 MI Diagonal Filter	
DAC 288 MB/s	One HDTV ≈ One LHC channel
60 MHz	Analog bandwidth \approx 100 MHz Digital throughput \approx 1 Gbit/sec

≈ 10⁵ Hz rate

317

· • · · ·

= 107 Hz rate

Digital signal processing

The signals generated by an LHC calorimeter cell will be diffused over more than one bunch crossing. A digital channel with programmable filter capability is needed to extract the physical information and to associate events within bunch crossings. In addition the digital analysis can be the first step of the

In addition the digital analysis can be the first step of the trigger process.

Apart from considerations of power consumption, electronics packaging and radiation hardness, it seems very attractive to go digital immediately after the preamplifiers, complementing the analog shaping with a **pipelined digital signal processor**.

П

Sampling

, I

Feature extraction

Baseline normalization

The purpose of the base line normalization is to remove the constant background and low-frequency components. The methods presented here are based on the subtraction of a signal mean value from the original signal. The analysis window length in base line normalization is substantially larger than the length of a typical filter.

Mean/Median. Baseline estimation

· · ·

High precision measurements (16-bit range, 10-bit resolution) Pipelined system running at 40 MHz

Timing, trigger and control distribution. TTC

Development of basic hardware and software components of a **multichannel** optical-fibre timing distribution system for LHC detector front-end electronics, and investigation of the feasibility of simultaneously exploiting the timing system for transmission of the level-1 trigger acceptance and addressable control information

Timing, trigger and control distribution

First level trigger

Signatures

Missing E_τ
Background (e/γ)

Cluster finding Segment finding

Synchronous system 25ns pipelined ≈ μs latency (processing)
 ≈ μs latency (distribution)

Muon level-1 coincidence matrix

Segment identification in low occupancy detectors.

Muon first level trigger

Muon level-1

Trigger based on tracks in external muon detectors that point to interaction region

Low-p, muon tracks don't point to vertex
Multiple scattering
Magnetic deflection
Two detector layers
Coincidence in "road"

- Content addressable memories
- Matrix logic (ASIC, Gate array...)

П

Detectors: RPC (patterr recognition), DT(track segment)

Level-1 structure

- Readout crate

- Readout unit

- Dual port memory

П

- Fast data link (~1Gb/s)
- Pipelined processing elements (40,80,160 MHz)

I .

• High speed interconnection network

Readout crate

Readout unit

П

RCS Readout Crate Supervisor. RCS is a commercial VME processor. ROC performs the data acquisition control and monitoring and the local detector fes, and maintenance. RFC Readout Flow Controller. RFC Unit gen vates and processes the signals used to monitor and control the data liw. The unit provides also multi-crate inte connection facilities (e.g. SCI) for local detector test and maintenance.

FED The FrontEnd Driver is a detector dependent module performing the data collection from the detector frontend electronics. The FED cost is included in the detector electronics estimations

trontend electronics. The FED cost is included in the detector electronics estimations RDPM. The Readout Dual Pon Memory is the data acquisition module used to buffer multi-exent (up to 10) data and to communicate with the event builder switch. It may be implemented by a custom design, by an embedded processor system or by a standard desktop con puter (the solution depends on the attainable performances such as 100 MBS throughput and 100 MB memory) RDL The digital data link (transmitter receiver and fiber) between the RDPM and the event funder

RDL The digital data link (transmitter receive: and tiber) between the RDPM and the event functions. The composition (number of FED boards and the number of RUs in a crate) is determined by the average data volume produced in each partition and can vary with the detector type and the detector region. Moreover the total number of RUs (event builder ports) is related to the data link speed and the exploitable bandwidth (- 500 Gb/s) of the switch able to comply with the CMS requirements (100 kHz rate of TMB events with high level trigger scheme based on particil event data building).

Dual port memory functions

- Autonomous Input/Output
- 100 MB/s throughput
- Large memory (up to 100MB)

SeqIN allocates buffers and reads data from F/E modules

CPU monitors the status and handles faulty conditions. CPU(VME) may spy event data for monitoring

SeqOUT formats the data according to the data link standards and controls the transmission

A DPM module operates as a du I access disk controller. At input, files (event buffers) have to be created, written and closed. At output, files (event buffers) have to be open, read and deleted.

Given the high throughput (≈ 100 MB/s) all I/O and control operations may have to be done by hardware units. A standard processor is used to recover from faulty situations and to manage errors and conflicts

DPM structure

Event builder

Readout Network & Event Builder

- $\bullet\approx$ 1000 \bullet 266 Mb/s data links
- \approx 100 (10 to 1) multiplexers
- \approx 100 100 2.4 Gb/s Switching network

• \approx 500 Gb/s aggregate bandwidth

Event building

ATM

Asynchronous Transfer Mode network system. ATM telecommunication standard 1992. 155 Mb/s ... 620 Mb/s ... 2.4 Gb/s 199X

- Alcatel ATM. 155 Mb/s
- IBM prizma. 400 Mb/s
- AT&T Phoenix. 320 Mb/s
- More on the market in the future

Fiber Channel

Fiber Channel can be a very efficient way to event assembly when dealing with large data records. It can be a valuable candidate for event building in a scheme where many (\approx 100) event data blocks are transferred at the time

[]

Scalable coherent interface. SCI

SCI is IEEE standard for processor intercommunication. It has to demonstrate its impact in the computer industry. Node interfaces are almost available and switches are under study

A SCI based event builder allows an event builder architecture where the processor accesses only the data needed at a given trigger level

Event filter

The event filter consists of a set of high performance commercial processors organized into many farms convenient for on-line and off-line applications. The farm architecture is such that a single CPU processes one event.

• \approx 100 farms of 50 processors

- $\bullet \approx 1000 \; MIPS/processor$
- Trigger levels 2, 3 .. & Data analysis

1 event 1 sec	Event Filter	Analysis	Total Online&Offline
UA/LEP	4 MIPS	400 MIPS	50 MIPS (30 Hz)
LHC	400 MIPS	4•10⁴ MIPS	4•10 ⁶ MIPS (10 KHz)
			Latency \approx ms to sec

11

Massive parallel system

Farm of processors

I .

CMS virtual level-2

1) The level-2 selection uses the calorimeter, muon and preshower data. The sub-events are built using a fraction of the switch bandwidth (e.g. 30%).

The two operations take place in parallel

The sharing between the level-2 data sources and the rest is such to match the event builder bandwidth with the level-2 rejection rate.

Up to 100 kHz with virtual level-2 mode
 Up to 50 kHz reading the full event data

2) The rest of the event data is sent after the level-2 decision if the event is accepted

Event filter today

• RD-24 SCI interconnect link.

- Hardware node (Dolphin design GaAs and CMOS).
- Switches and bridges under study as well

[]

Technologies

The CMS data acquisition system exploits standard industry components as far as possible, and a considerable part of the system is expected to be developed and installed by industry.

11

Switching network and farm power extensions can be made to track LHC machine improvements by adding the most performant and cost effective processors available at the time.