Trigger and Data Acquisition for collider experiments
Present and future

By P. Le Dû

- Basic parameters and constraints
- Event selection strategy and evolution of architectures
- Implementation of algorithms: hardware and software
- Boundaries (Trigger levels/DAQ/on-line/off-line)
- Tools, techniques and standards

Menu
Brief overview of the T/DAQ issues and directions for the future illustrated by few typical exemples taken from general purpose detectors

- present (2004)  → Tevatron (CDF, D0)
- tomorrow (2007) → LHC (ATLAS, CMS)
- tomorrow (2010) → FNAL (Btev)
- future (> 2015) → FLC → NLC, GLC, Tesla
Credit

- Trigger/DaQ sessions
  - Gordon Watts
    - University of Washington, Seattle
    - NSS 2003

- M.Mur & S. Anvar (engineers CEA DAPNIA)
Constraints → a multiparameters problem

- Detector Environment
  - Pileup
  - Multiple Interactions
  - Beam-gas Interactions

- Bunch Spacing
- Bunch Train Length
- Sustained Duty Cycle

- Trigger Complexity
  - What is required to make the decision?
  - Simple Jet Finding
  - Complex Track Finding
  - Inter-detector Correlations

- Upgradability

- Trigger Requirements
  - Rejection
  - % Deadtime
  - Input/Output Rate

- Offline
  - Accept Rate
  - Archive Ability
  - Reconstruction Capability

- Physics

- Accelerator

- Implementation
  - Multi Level
  - Low level feature extraction
    - ASIC, PLA, etc.
  - Farms
    - PC Platforms
    - Transfer Technology
    - ATM, Gigabit, etc.
  - Technology/Capability Trade off.

- Cost

Trigger and DAQ go hand in hand
<table>
<thead>
<tr>
<th>Exp. Year</th>
<th>Collision rate</th>
<th>Channel count</th>
<th>L1A rate</th>
<th>Event building</th>
<th>Processing Power</th>
<th>Sociology</th>
</tr>
</thead>
<tbody>
<tr>
<td>UA’s 1980</td>
<td>3 µsec</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5-10 MIPS</td>
<td>150-200</td>
</tr>
<tr>
<td>LEP 1989</td>
<td>10-20 µsec</td>
<td>250 - 500K</td>
<td>-</td>
<td>10 Mbit/sec</td>
<td>100 MIPS</td>
<td>300-500</td>
</tr>
<tr>
<td>BaBar 1999</td>
<td>4 ns</td>
<td>150K</td>
<td>2 KHz</td>
<td>400 Mbit/s</td>
<td>1000 MIPS</td>
<td>400</td>
</tr>
<tr>
<td>Tevatron 2002</td>
<td>396-132 ns</td>
<td>~ 800 K</td>
<td>10 - 50 KHz</td>
<td>4-10 Gbit/sec</td>
<td>5.10^4 MIPS</td>
<td>500</td>
</tr>
<tr>
<td>LHC 2007</td>
<td>25 ns</td>
<td>200 M*</td>
<td>100 KHz</td>
<td>20-500 Gbit/s</td>
<td>&gt;10^6 MIPS</td>
<td>2000</td>
</tr>
<tr>
<td>NLC 2015</td>
<td>2.5 - 330 ns</td>
<td>800 M*</td>
<td>-</td>
<td>1 Gbit/s</td>
<td>~10^5 MIPS</td>
<td>5000</td>
</tr>
</tbody>
</table>

* including pixels

<table>
<thead>
<tr>
<th>Sub-Detector</th>
<th>Tevatron</th>
<th>LHC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pixel</td>
<td>-</td>
<td>150 M</td>
</tr>
<tr>
<td>Microstrip</td>
<td>~500 K</td>
<td>~10 M</td>
</tr>
<tr>
<td>Fine grain trackers</td>
<td>~100 K</td>
<td>400 K</td>
</tr>
<tr>
<td>Calorimeters</td>
<td>50 K</td>
<td>200 K</td>
</tr>
<tr>
<td>Muon</td>
<td>50 K</td>
<td>~1 M</td>
</tr>
</tbody>
</table>
Rate and Data Volume → the age of network

Fast Networking Moves Data from a local environment to a Global One

Cheap Commodity Processors allow for powerful algorithms with large rejection

Cheap DRAM allows for large buffering, and latency

L1 Rate (Hz)

10^6

10^5

10^4

10^3

10^2

10^4

10^5

10^6

Event Size (bytes)

High No. Channels
High Bandwidth
(500 Gbit/s)

High Level-1 Trigger
(1 MHz)

High Data Archive
(PetaByte)

The Challenge!
Today

- Multilevel trigger
- Typical designs
- Examples: CDF & D0

We already see the evolution!
Multilevels trigger

- Required rejection is orders of magnitude
  - Algorithms to attain required rejection are too sophisticated.
  - Accelerator backgrounds can also contribute to the problem
    \[ \text{e}^+\text{e}^- \text{ vs pp} \]
- Level 1 is hardware based
  - Crude signatures (hits, local energy deposit over threshold...)
  - Operates on reduced or coarse detector data
- Level 2 is often a composite
  - Hardware to preprocess data
    - Some Muon processors, Silicon Triggers
  - Software to combine
    - Matches, Jet finders, etc.
- Level 3 is a farm \( \rightarrow \) General Purpose CPUs

Almost every one uses this scheme
- Vary number and function of levels.
L1 $\Rightarrow$ Algorithms in Hardware

- **Cut out simple, high rate backgrounds**
  - beam-gas (HERA)
    - $z$ vertex
  - QCD (TeV)
    - jet energy, track matching
- **Capabilities are Limited**
  - Feature Extraction from single detectors
    - Hadronic or EM energy, tracks, muon stubs
  - Combined at Global Trigger
    - EM Object + Track

- **Characteristics:**
  - High speed & Dead-timeless
  - Limited ability to modify algorithm
    - But thresholds typically can be modified
  - Algorithms Frequently Matched to the detector (and readout) geometry
    - Vertical Design

- **Built from Modern Components**
  - Custom (ASICs)
  - Programmable Logic Arrays (FPGAs, etc.)
The Field Programmable Gate Array

- **Board on a chip**
  - Revolutionized the way board design is done.
  - Logic design is not frozen when the board is laid out.
    - But much faster than running a program in microcode or a microprocessor.
  - Can add features at a later time by adding new logic equations.
  - Basically a chip of unwired logical gates (flip-flops, counters, registers, etc.)

Several Types

- Simple Programmable Logic Devices (SPLD)
- Complex Programmable Logic Devices (CPLD)
- Field Programmable Gate Arrays (FPGA)
- Field Programmable InterConnect (FPIC)

The FPGA allows for the most logic per chip.
**Algorithms in Software**

**Sophisticated Algorithms**
- Frequently separating background physics processes from the interesting processes.
- Some experiments run their offline reconstruction online (Hera-B, CDF)
  - Perhaps 2D tracking instead of 3D tracking.
- Data from more than one detector often used.

**Two common implementations:**
- **DSPs tend** to live at lower levels of the trigger.
  - Stub finders in muon DØ’s muon system.
  - Transputers in ZEUS’ L2, DSPs on HERAB L2 & L3
- **CPU Farms** tend to be at the upper end

**Sometimes difficult to classify a system**
- Blurring of line between software and hardware.
Tevatron architectures (RUN 2)

CDF

7.6 MHz Xing rate

Detector → L1 Buffer Pipeline) → L1 Trigger → L2 Buffer → L2 Trigger → L2 Farm → Mass Storage

- 50 KHz
- 4 µs latency
- < 1KHz
- Asynchronous 20 µs latency
- 64 PCs
- 20-50 Hz

D0

7.6 MHz Xing rate

Detector → L1 Buffer Pipeline) → L1 Trigger → L2 Buffer & digitization → L2 Global → Preprocessing Farm → L2 Global → L2 Farm → Mass Storage

- 10 KHz
- 4.2 µs
- 100 µs latency
- 48 Input nodes
- 1 s latency
- Use a combination of both Hardware and General Purpose Processors

- **CDF & DØ’s Level 2 Trigger**
  - Each box is a VME crate
  - Contain specialized cards to land and reformat data from detector front-ends
  - Contain ALPHA processor to run algorithms.
Both CDF & DØ will trigger on displaced tracks at Level 2
- DØ will use another specialized set of L2 crates.

Cluster/hit finding
- CDF: Associative memories @50 KHz
- DØ: FPGA's @ 10 KHz

Input to Level 2 Global
- Select Events with more than 2 displaced tracks
- Resolution is almost as good as offline.
- Refined track parameters

Track Fitters
- DSP's Farm.
STT can provide precise online impact parameter measurement of which CTT is not capable.

STT proposed 1998 as addendum to DØ baseline. Received approval and funding in 1999
Layout of Run II A STT Crate

6 Identical Crates with
1 Fiber Road Card
9 Silicon Trigger Cards
2 Track Fit Cards

6 Identical Crates
1 Fiber Road Card
9 Silicon Trigger Cards
2 Track Fit Cards
STT decision time requirement: \(~50\mu s\).

TFC fitting performance is acceptable. Still tuning for the final configuration.

<table>
<thead>
<tr>
<th>Fit_Times</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Entries</td>
<td>257</td>
</tr>
<tr>
<td>Mean</td>
<td>15.32</td>
</tr>
<tr>
<td>RMS</td>
<td>11.73</td>
</tr>
</tbody>
</table>
**CDF L2 legacy decision Crate**

Technical requirement: need a **FAST** way to collect/process many inputs…

→ With the technology available back then (1990s), had to design custom (alpha) processor & backplane (magicbus) …

→ had to deal with the fact that each data input was implemented in a different way …

→ 6 different types of custom interface boards + custom processor & backplane …

---

Input: up to 50KHz  
Output: 300Hz
Can we design an universal board to replace ALL?

→ Pulsar project ...

- Convert/pre-process/merge different data inputs into a standard link (CERN S-LINK)
- Interface with commodity PC(s) via high bandwidth SLINK → PCI
- Much enhanced performance ...
- Much easier maintenance

Pulsar is designed to be able to do more than these ...

But will only discuss how we address the issues.
Pulsar is designed to be: Modular, universal & flexible, fully self-testable (board & system level)

Has ALL interfaces L2 decision crate has

Personality cards
- Hotlink IO
- Taxi IO
- SVT/XTRP
- Level 1
- TS IO
- S-LINK IO

User defined interfaces

PULSAR
- Pulser And Recorder
- SRAMs
- RF clock

VME
CDFctrl

AUX
- S-LINK
- S-LINK

S-LINK

S-LINK to PCI

Gigabit Ethernet

Spare lines

Standard link
- another Pulsar

All interfaces are bi-directional (Tx & Rx)
Lego-style, open design ...

LXWS042 - P. Le Dû
Results: No blue wire on all prototypes → All became production boards

MOAB (Mother of All Boards)

- Design & Verifications: ~3 people in ~9 months
- Initial checkout of ALL interfaces: ~2 people in ~1 month

4 types of custom mezzanine cards (Tx/Rx)

・ Pulsar

Works up to 100 MHz

Is this luck?
Modular, universal & flexible

One possible Pulsar formation

⇒ General purpose DAQ
⇒ CDF L2 upgrade

L1 muon
L1 track

SLINK

muon

L1 trigger

SLINK

merger

L2 SVT

cluster

Pulsar pre-processors

PC 0
PCI→SLINK

PC 3
SLINK/GBE→PCI

PC 1

PC 2

Trigger Supervisor

ShowerMax

Pulsar mergers or GBE switch

L2toTS

LXWS042 - P. Le Dû
α vs. PC L2 algorithm timing

<table>
<thead>
<tr>
<th>PC timing</th>
<th>α timing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entries</td>
<td>Entries</td>
</tr>
<tr>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>RMS</td>
<td>RMS</td>
</tr>
<tr>
<td>Underflow</td>
<td>Underflow</td>
</tr>
<tr>
<td>Overflow</td>
<td>Overflow</td>
</tr>
</tbody>
</table>

Entries: 63057, Mean: 3.017, RMS: 2.87, Underflow: 0, Overflow: 198
Entries: 63058, Mean: 7.23, RMS: 5.079, Underflow: 0, Overflow: 1.524e+04

1.8GHz Linux
500MHz No OS

Tested with real L2 trigger data
Summary: how we addressed the issues

• **Uniformity & modularity & flexibility**
  Lego-style, general purpose design, backward & forward compatible. Many applications within & outside CDF:
  Plan to replace/upgrade > 10 different types of CDF trigger board
  Compatible with S-LINK standard → commodity processors …
  Knowledge gained transferable to and from LHC community…

• **Design & verification methodology**
  simulation & simulation: single/multi-board/trace & cross talk analysis …
  → no single design or layout error (blue wire) on all prototypes

• **Testability & commissioning strategy**
  Board & system level self-testability fully integrated in the design,
  Suitable to develop and tune an upgrade system in stand-alone mode
  Minimize impact on running experiment during commissioning phase:

• **Maintenance & firmware version control**
  Built-in maintenance capability; Common firmware library (VHDL), CVS …

• **Documenting the work**
  [http://hep.uchicago.edu/~thliu/projects/Pulsar/](http://hep.uchicago.edu/~thliu/projects/Pulsar/)
Farm Triggers

- Industry has revolutionized how we do farm processing
  - Wide scale adoption by HEP.
  - Online and Offline.
  - One CPU, one event
    - Not massively parallel

- Data must be fed to farm
  - As many ways as experiments (push/pull)
  - Flow Management & Event Building

- Farm Control: an issue!
  - Distributed system can have 100’s of nodes that must be kept in sync.
CDF

- Data Flows when Manager has Destination
  - Event Builder Machines
- One of first to use large switch in DAQ.

- Dataflow over ATM switch
  - Traffic Shaping
- Backpressure provided by Manager.
D0 3 Level Trigger Design

Detector component raw data

L1 Trigger

L2 Trigger

L3 DAQ and Trigger Farm

- Average event size 250 kB
- 1 kHz design L2 accept rate

2.5 MHz

5 kHz

1 kHz

~ 50 Hz

Tape storage
L3 DAQ Requirements

1 kHz 250 kB/event

Total of 63 VME Readout Crates
1-10 VME modules/crate
1-20 kB total/crate

L3 Trigger Processor Farm
(currently 82 dual-CPU nodes)
Commodity Ethernet L3 DAQ Design

- VME Single Board Computers (SBC), Linux
- Ethernet TCP/IP
- Linux farm nodes

![Diagram of Ethernet L3 DAQ Design]

- ROC SBC
- Ethernet concentrator
- Supervisor CPU
- Cisco 6509 Ethernet Switch
- L1/L2 Trigger frame work
- Routing Master CPU
- Farm CPU
- Offline Storage

Run control

Nuclear Science Symposium, Portland, October 2003
Cisco 6509 Ethernet Switch

- Backplane easily handles 250 MB/s design rate
- Using:
  - One 16 x 1 Gb/s module
  - Two 48 x 100 Mb/s modules with 1 MB output buffer per port
- Room for expansion
The L3 Trigger Farm

dual Ethernet 100 Mb/s  
dual CPU 1 GB memory

34 dual AMD Athlon 2000  
48 dual PIII 1 GHz

(Farm typically ~50% busy, <100 ms/event on average)
Error Recovery

- Potential source of inefficiency: occasional human errors by DAQ shifters
The Human Factor

- System are complex
  - 1000’s of interacting computers
  - Complex Software with 100’s of thousands of lines of code
- We get the steady state behavior right.
- What about the shifter who does a DAQ system reset 3 times in a row in a panic of confusion?
- The expert playing from home?

Anonymous messing up DAQ system from home

The Microsoft Problem?

Self Healing Systems?
(too smart for their own good)
Conclusions

D0 L3 DAQ built from commodity hardware

- 63 VME sources to 82 node processor farm
- 250kB events at 1kHz, without incurring dead time
- Ethernet – TCP/IP communication

One year to design and build

- Hardware availability
- Open software tools
- On-hand expertise VME, Linux, TCP/IP
- Prior DAQ experience

Stable performance

- Stable, reliable, low maintenance and smooth operation
- Meets current and future needs of D0
NA60 readout system → ALICE prototype

- PCI based system:
  - Good performances / low cost
  - Readout of several different detectors
    - General purpose PCI readout card
    - Detector specific mezzanine

- NA60 readout is spill-buffered
  - DAQ handshake with readout system
Tomorrow

- LHC architecture
- Evolution 2005 → 2010

Conceptual models & options
# LHC Multilevels Selection Scheme

## Production Rate

<table>
<thead>
<tr>
<th>Event Type</th>
<th>Production Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>QCD</td>
<td>$10^8$</td>
</tr>
<tr>
<td>W, Z</td>
<td>$10^6$</td>
</tr>
<tr>
<td>Top</td>
<td>$10^4$</td>
</tr>
<tr>
<td>Z'</td>
<td>$10^2$</td>
</tr>
<tr>
<td>Higgs</td>
<td>$10^2$</td>
</tr>
</tbody>
</table>

## Available Time

- 25 ns
- Few µsec
- ~ms
- > Sec

## Hardwired Processors
- (ASIC, FPGA)

## Standard CPUs Farms & Networks

## Recorded Events

LXWS042 - P. Le Dû
Strategy for event selection @ LHC

40 MHz

- "Local" identification of high Pt objects → use coarse dedicated data
  - Electrons /Photons, Hadrons & Jets → Energy clusters
  - Muons → Track segments
  - Neutrinos → Missing Et

100 KHz

- Particle signature (e/g,h,Jet,μ ...) → use final digitized data
  - Refine Pt cuts → fine granularity & track reconstruction
  - Combine detectors → Converted electron,“Punchthrough”, decays

1 KHz

- Global topology → multiplicity & thresholds

100 Hz

- Identification & classification of physics process → trigger menu
  - Partial event reconstruction → Vertices, Masses, Missing Et...

- Physics analysis
  - “Off-line” type analysis of classified events
ATLAS CMS

Level 1

Front end buffers (~10^7 channels)
1000 - 1500 Gbit/s

Level 2

Read Out Buffers
1000-1500 units
100 Mbit/sec
20 Gbit/s
256 x 256
1 KHz

Level 3

Processor farm (~2 to 5 x 10^6 MIPS)
10 Gbit/s

Switching Network
(Event builder)

Permanent storage
100 MBytes/sec --> TBytes/day

Custom made Hardware pipelined rigger processor

100 KHz

Region Of Interest

Commercial DSPs, RISCs, etc.

256 x 256

100 MHz

Level 1

Sub- Detectors

ROI

100 KHz

Read Out Buffers

1000-1500 units

40 MHz

Level 2

Switching Network
(Event builder)

100 KHz

1KHz

100 Hz

Switching Network
(Event builder)

Permanent storage
100 MBytes/sec --> TBytes/day

High Levels

2.5 μs fixed

Synchronous

Asynchronous

few m s variable

1-10 s variable

100 KHz

20 Gbit/s
250-500 Gbit/s

10 KHz

1 Gbit/sec
1000 x1000

“Virtual” level 2

1000 -1500Gbit/s

100 Mbit/sec

LHC Logical structure

LXWS042 - P. Le Dû
Requirements and design parameters

<table>
<thead>
<tr>
<th>Detector</th>
<th>Channels</th>
<th>Control</th>
<th>Ev. Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pixel</td>
<td>60000000</td>
<td>1 GB</td>
<td>70 (kB)</td>
</tr>
<tr>
<td>Tracker</td>
<td>10000000</td>
<td>1 GB</td>
<td>300</td>
</tr>
<tr>
<td>Preshower</td>
<td>145000</td>
<td>10 MB</td>
<td>110</td>
</tr>
<tr>
<td>ECAL</td>
<td>85000</td>
<td>10 MB</td>
<td>100</td>
</tr>
<tr>
<td>HCAL</td>
<td>14000</td>
<td>100 kB</td>
<td>80</td>
</tr>
<tr>
<td>Muon DT</td>
<td>200000</td>
<td>10 MB</td>
<td>10</td>
</tr>
<tr>
<td>Muon RPC</td>
<td>200000</td>
<td>10 MB</td>
<td>5</td>
</tr>
<tr>
<td>Muon CSC</td>
<td>400000</td>
<td>10 MB</td>
<td>80</td>
</tr>
<tr>
<td>Trigger</td>
<td>1 GB</td>
<td></td>
<td>20 ~ 800</td>
</tr>
</tbody>
</table>

Event size ~ 1 Mbyte
Max LV1 Trigger 100 kHz
Online rejection 99.999%
System dead time ~ 0%

Detectors

- MUON BARREL
  - Drift Tube Chambers (DT)
  - Resistive Plate Chambers (RPC)

- CALORIMETERS
  - ECAL Scintillating PbWO₄ Crystals
  - HCAL Scintillator sandwich

- IRON YOKE

- SUPERCONDUCTING COIL

- TRACKERs
  - Pixels Silicon Microstrips
  - Cathode Strip Chambers (CSC)
  - Resistive Plate Chambers (RPC)

Total weight : 12,500 t
Overall diameter : 15 m
Overall length : 21.6 m
Magnetic field : 4 Tesla
Collision rate
Level-1 Maximum trigger rate
Average event size
Event Flow Control Mssg/s
† 50 kHz at startup (DAQ staging)

40 MHz
100 kHz†
≈ 1 Mbyte
≈ 10^6

No. of In-Out units
Readout network bandwidth
Event filter computing power
Data production
No. of PC motherboards
‡ 6×10^5 at startup

512
≈ 1 Terabit/s
≈ 10^6 SI95‡
≈ Tbyte/day
≈ Thousands

† 50 kHz at startup (DAQ staging)
Filter Farm

Event data partitioned into about 500 separated memory units

Two architectures

Massive parallel system
ONE event, ALL processors
- Low latency
- Complex I/O
- Parallel programming

Farm of processors
ONE event, ONE processor
- High latency (larger buffers)
- Simpler I/O
- Sequential programming

≈ 50..100 kHz

≈ 200 Hz per BU
Filter Farm

...Builder Units...

Filter Data Network

FS
DEMONSTRATOR results:

Myrinet: baseline technology:
- Data to surface FED and Readout Builders
- Feasible with ‘today’s’ (2003) technology

GbEthernet: backup technology:
- first results up to 31x31+BM. Raw packet and TCP/IP
- for nominal conditions (16 kB and RMS = 8 kB): 75 MB/s (60% of wire) speed
- it scales, not clear if scaling persists to larger N or higher switch load
- longer term: with good NIC, fast PC, jumbo frames etc might be possible to reach 100 MB/s per node and twice that for two-rail configuration
- option to track

Two-stage Event Builder Architecture:
- Allows scaling and staging.
- RU-builder: 16 kB fragments at 12.5 kHz easier than 2 kB at 100 kHz
- EVB switch size 8x8 and 64x64 rather than 512x512
The ATLAS Trigger

Interaction rate
~1 GHz
Bunch crossing rate 40 MHz

Level 1 Trigger
75 kHz

Level 2 Trigger
~ 2 kHz

Event Filter
~ 200 Hz

Target processing time

Rate

Level 1
Hardware trigger

High Level Triggers (HLT)
Level-2 + Event Filter
Software trigger
Regions of Interest concept (ROI)

- **Bandwidth/Physics Compromise**
  - Farm usually reads out entire detector
  - HW often look at single detector

- **ROI sits in the middle**
  - CP Farm CPU requests bits of detector
    - Uses previous trigger info to decide what regions of the detector it is interested in.
  - Once event passes ROI trigger, complete readout is requested and triggering commences.

- **Flexible, but not without problems.**
  - Pipelines in front ends must be very deep
  - Farm decisions happen out-of-event-order
    - Pipelines must be constructed appropriately.

ATLAS, HERA-B, BTeV...
Region of Interest (Rol) Concept

- **Level-1**
  - Uses only *coarse* calorimeter and muon spectrometer data
  - *Local signatures* dominate selection
  - *No matching of different detectors*

- **Rol**
  - The *Region of Interest* is the *geometrical location* of a LVL1 signature (identified high $p_T$ object)
  - Allows access to *local full granularity data* of each relevant detector
  - $<\text{Rol/Level-1 accept}> \sim 1.6$

- **Level-2**
  - Seeded with Rol
  - *Matching of full detector data within Rol*
  - Uses $\sim 2\%$ of the full event data for decision
High Level Triggers

- **Level-2** (75 kHz → 2 kHz, 10 ms target processing time)
  - Selection software runs in Level-2 Processing Unit (L2PU) in *multiple worker threads* (each thread processes one event)
  - Feature extraction in Rol with *specialized algorithms* that are optimized for speed and *combine information of all sub-detectors* sequentially

- **Event Filter** (2 kHz → 200 Hz, 1-2 s target processing time)
  - Independent *Processing Tasks (PT)* run selection software on Event Filter (EF) farm nodes
  - *Full event reconstruction* (seeded by Level-2 result) with *offline-type algorithms* which have access to full calibration data
  - Software *based on ATLAS offline environment* ATHENA/GAUDI

- **HLT Selection Software Framework requirements**
  - *Common to Level-2 and EF*
  - Possibility to *move algorithms from Level-2 to EF* for optimization
  - *Development and evaluation in offline environment*

**Need offline interfaces in Level-2:** *Steering Controller*
BTeV

B_s Mixing, Rare Decays

23 million pixels - in Level 1!
132 ns Bunch Spacing
100 kb/event
Trigger full rate. 800 GB/sec!!

**Level 1**

1. Hit clustering in Pixels - FPGAs
2. Cluster linking in inner and outer layers - FPGAs
3. Track finding in B field (p_T>0.25) - Embedded CPU Farm
4. Hard Scatter Vertex Finding - Embedded CPU Farm
5. DCA test for displaced tracks due to B decays - Embedded CPU Farm

8 Planes of L1+L2/L3 (round robin)
BTeV Trigger Design

Buffering for 300k interactions (300ms)

800 GB/sec optical links

L1 Farms

Global L1

L2/L3 Farms

8 Data Highways

200 MB/sec

3.8 kHz

78 kHz

Large # of High End Boxes (dual 4 GHz)

Uses ROI - Similar to ATLAS

But L2/L3 are interleaved on same box

HLT Algorithm
Evolution → 2005-2010

- **LHC (ATLAS & CMS) → Two levels trigger**
  - L1 = physics objects (e.g., jet, m ..) using dedicated data
  - L2 + L3 = High Levels « software » Triggers using « digitized data »

- **Complex algorithms like displaced vertices are moving downstream**
  - New L1 Calorimeter trigger (sign electron, jets and taus)
  - CDF/DO : L2 vertex trigger
  - LHCb/Btev : L0/L1 b trigger

- **Use as much as possible comodity products (HLT)**
  - No more « Physic » busses → VME, PCI ..
  - Off the shelf technology
    - Processor farms
    - Networks switches (ATM, GbE )
  - Commonly OS and high level languages
On-off line boundaries

- Some computing issues
- Common aspects
- Evolution

Detectors will produce a huge data volume
- Few Tbytes/year! (according to the present models)
- Is it reasonable (and useful!)
- Can we reduce it? \( \rightarrow 100 \)

A personal, critical .. non conformist view!
About On-Off line boundaries

- **Detectors are becoming more stable and less faulty**
  - High efficiency, Low failure rate
  - Powerfull “on-line” diagnostics and error recovery (expert systems)

- **On-line computing is increasing and not doing only “data collection”: More complex analysis is moving on-line**
  - “Off-line” type algorithms early in the selection chain (b tag ..)
  - Selection of “data streams” --> Important role of the “Filter”
  - Precise alignment needed for triggering
  - Detector calibration using Physics process available
  - On-line calibration and correction of data possible

- **Common aspect**
  - Algorithms, Processing farms, Databases...
  - use similar hw/sw components (PC farms..)

*Boundaries become flexible*
Trigger & Event Analysis common strategy

**HLT Algorithms Select**
« Physics tools »

**Select objects**
and compare to Menus

**Fast Analysis Stream**
Partial/Full Event Building

**Physics streams**

- **Simple signatures**
  - e/g, μ, taus, Jet
  - Refine Et and Pt cut
  - Detector matching
- **Complex signatures**
  - Missing Et, scalar Et
  - Invariant and transverse mass
  - separation …
  - vertices, primary and displaced
- **Selection**
  - Thresholding
  - Prescaling
  - “Intelligent formatting “

**Event Candidate & classification**

- Simple signatures
- Complex signatures
- Others signatures
- Topology
- Menu

**On-Line Processing**

- Calibration Constant
- Sub-Detector performance
- Event Background Infos to the LHC

**Database**

- “Analysis” farm
- Candidates Storage
- Sample Prescale Compress
- “Analysis” farm

**Temporary storage**

- S1
- S2

**Final storage**

- Sn

**Keep only “Physics process”**
(with specific data format)
Prescale and compress the rest

- **Prescale**
- **Compress**

**Simple signatures**
- e, µ, τ, Jet

**Complex signatures**
- Missing Et, scalar Et
- Invariant and transverse mass
- separation …
- vertices, primary and displaced

**Selection**
- Thresholding
- Prescaling
- “Intelligent formatting “

**LXWS042 - P. Le Dû**
Summary of T/DAQ architecture evolution

- **Today**
  - Tree structure and partitions
  - Processing farms at very highest levels
  - Trigger and DAQ dataflow are merging

- **Near future**
  - Data and control networks merged
  - Processing farm already at L2 (HLT)
  - More complex algorithms are moving on line
  - Boundaries between on-line and off-line are flexible
  - Commodity components at HLT
Technology forecast (2005-2015)

• **Processors and memories**
  - Continuous increasing of the computing power
    - More’s law still true until 2010!  \( \rightarrow \times 64 \)
    - Then double every 3 years  \( \rightarrow \)
  
  - Memory size quasi illimited!
    - Today: 64 MBytes
    - 2004 : 256 MB
    - 2010 : > 1 GB

• **Networks: Commercial telecom/computer standards**
  - Multi (10-100) GBEthernet

Systematic use of:
Off the Shelves comodity products
Evolution of technologies since 30 years
What next?

- the Next Linear Collider
- Software trigger concept
- Evolution

The ultimate Trigger?
LC beam structure

- Relatively long time between bunch trains: **199 ms**
- Rather long time between bunches: **337 ns**
- Rather long bunch trains (same order as detector read-out time): **1 ms**

- Relatively long time between bunch trains (same order as read-out time): **6.6 ms**
- Very short time between bunches: **2.8 ns/1.4 ns**
- Rather short pulses: **238 ns**
Software trigger concept → No hardware trigger

Detector Read-out of a complete bunch train
- Signal processing - digitization
- Perform Zero suppression and/or data compression
- Buffering – dead time free pipeline

Data Collection is triggered by every train crossing
- Full event building of one bunch train
- Software Selection
  - Sequential selection (equivalent to L1-L2-L3 …) with off line algorithms
  - Event classification according to physics, calibration and machine needs
  - Select bunch of interest

On-line processing → Event Filtering …..

Data Flow

Detector Front End

Read-Out Buffer

Switching Network

Processor Farm

Storage & Analysis

Data streams: S1, S2, S3, S4, Sn

1 ms
3000 Hz

1 MBytes
Average event size

200 ms
30 Hz

Few sec

LXWS042 - P. Le Dû
Advantages → all

- Flexible
  - fully programmable
  - unforeseen backgrounds and physics rates easily accommodated
  - Machine people can adjust the beam using background events

- Easy maintenance and cost effective
  - Commodity products: Off the shelf technology (memory, switches, processors)
  - Commonly OS and high level languages
  - on-line computing resources usable for « off-line »

- Scalable:
  - modular system

Looks like the ‘ultimate trigger’ → satisfy everybody: no loss and fully programmable
Consequences on detector concept

- **Constraints on detector read-out technology**
  - **TESLA:** Read 1ms continuously
    - VTX: digitizing during pulse to keep VTX occupancy small
    - TPC: no active gating
  - **JLC/NLC:**
    - 7 ms pulse separation
    - detector read out in 5 ms
    - 3 ns bunch separation
    - off line bunch tagging

- **Efficient/cheap read-out of million of front end channels should be developed**
  - silicon detectors (VTX and SiW calorimeters)
**Tesla Architecture**

- **Detector Channels**
  - VTX: 799 M
  - SIT: 300 K
  - FDT: 40 M
  - TPC: 1.5 M
  - FCH: 20 K
  - ECAL: 32 M
  - HCAL: 200 K
  - MUON: 75 K
  - LAT: 40 K
  - LCAL: 20 K
  - Detector Buffering (per bunch train in Mbytes/sec)
  - Processor farm (one bunch train per processor)
  - Computing resources (Storage & analysis farm)
  - Event building network
  - Event manager & Control

- **Network**
  - 10 Gbit/sec (LHC CMS 500 Gb/s)

- **Storage & Analysis Farm**
  - 30 Mbytes/sec → 300 TBytes/year
FLC 'today' Network model

On-Detector Front End RO (Silicon On Chip)

Interface: Intelligent PCI 'mezzanines'

Local/global Remote Control

Run Control

Monitoring Histograms

Event Display

DCS

Databases

...
Future Trends

- **Evolution of standards → no more HEP!**
  - HEP: NIM (60s), CAMAC (70s), FASTBUS (80s),
  - Commercial OTS: VME (90s), PCI (2000) → CPCI?
    - High Speed Networking
    - Programmable Logic Arrays (FPGAs)
    - Commodity CPUs

- **Looking ahead → where is the market moving now?**
  - No wide parallel data buses in crates
  - Backplanes used for power distribution, serial I/O, special functions
  - Small networked devices & embedded systems
  - High speeGb/s fiber & copper serial data links
  - Wireless devices and very-local area networks
Summary

- Higher level trigger decisions are migrating to the lower levels → Software Migration is following functional migration
  - Correlations that used to be done at Level 2 or Level 3 in are now done at Level 1.
  - More complex trigger (impact parameter!) decisions at earlier times (HLT) → Less bandwidth out of detector??

- Boundaries
  - L2 and L3 are merging into High Levels Triggers
  - DAQ and trigger data flow are merging
  - On-line and off-line boundaries are flexible

- Recent Developments in Electronics
  - Line between software and hardware is blurring
  - Complex Algorithms in hardware (FPGAs)
  - Possible to have logic designs change after board layout
  - Fully commercial components for high levels.

- “Software trigger” possible for the future LC
FLC issues for the next steps

- **NLC/GLC scheme?**
  - Particularities .... need initiate discussions... Soon!
- **Interface with machine**
  - Which infos are needed?
- **Gaine experience with Model Driven Technologies**
  - Follow XML implementation experience (CMS)
- **Modelling using ‘simple tools’**
  - Quantitative evaluation of the size of the system
- **What about GRID ???**
- **Define ‘boundaries’ → functional block diagram**
- **Integration of GDN & DAQ?**
- **Trigger criteria & algorithms**