Data-analysis at colliders

Prof. J. D'Hondt
Vrije Universiteit Brussel
jodhondt@vub.ac.be
Tel: 02-6293483
Office: Campus Oefenplein – 0G117

Vrije Universiteit Brussel

Content

Overview of a contemporary collider experiment

I
- General timeline of a collision experiment (eg. LHC)
- Overview of the structure of all systems in a detector
- First layer of triggers to reduce the data stream to tape
- Data Acquisition system & High Level triggers

II
- Reconstruction aspects of tracks, calorimeter deposits, muons, electrons, taus, hadronic jets of particles
- Identification of specific objects (eg. B-tagging)

III
- General analysis techniques to select the signal processes
- Advanced statistical tools for measurements
- An example analysis of how it all comes together…

It should take about 3 sessions of about 4-5 hours on average to cover the course. This will be complemented with a reading exercise and a presentation by the students.

The first part

General concepts of triggering and DAQ

- Introduction to high energy physics (HEP)
  - o Theory evolution & experiments evolution (accelerators)
  - o Evolution of particle detectors (rate of particles produced in collisions)
- General timeline of a collision experiment (eg. LHC)
  - o From the design concepts to the final paper
  - o Introduction to the Large Hadron Collider at CERN
- Overview of the structure of all systems in a detector
  - o Complexity of a particle detector (eg. Compact Muon Solenoid)
  - o General layout of a Trigger & Data Acquisition (DAQ) system
- First trigger layer to reduce the data stream to tape
  - o Hardware based trigger setups & performance
- Data Acquisition system & High Level Triggers
  - o Read-out of the detector & event building
  - o Final trigger level algorithms

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 3

Introduction to HEP

- The basic goal of the field of High Energy Physics is to reveal the fundamental particles in Nature and the interactions among them.

- It aims to construct a formal theory which allows to predict the outcome of interactions via pertubative calculations (matching general relativity and quantum mechanics in a quantum field theoretical approach).

- The domain covers a wide range of experiments of which some have a close connection to the fields of astronomy and cosmology. Due to the complexity of most particle detectors and accelerators the domain is a driving force for technology.

Evolution of the theory

- Today we know the Standard Model as a solid theory of particle physics putting the Electro-Weak and Strong force within one framework in a Quantum Field Theory.

- This model (which is basically a theory) has been the result of many experimental observations over the last ~50 years at for example fixed target experiments and collider experiments.

- To validate the Standard Model we aim to observe phenomena at always higher energy densities and we aim to measure its parameters as precise as possible.

- To achieve this we need accelerators which are as powerful as possible and detectors which can measure as precise as possible, hence driven the technology on different sides.

- One of the open questions we have today is the apparent dominating Dark Matter in the Universe for which the Standard Model particles cannot give an answer.

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 5

Timeline of a HEP experiment

The idea of a new laboratory is motivated by the need to investigate yet unexplored domains either in energy, in rate, in flavour, in exposure time, etc.
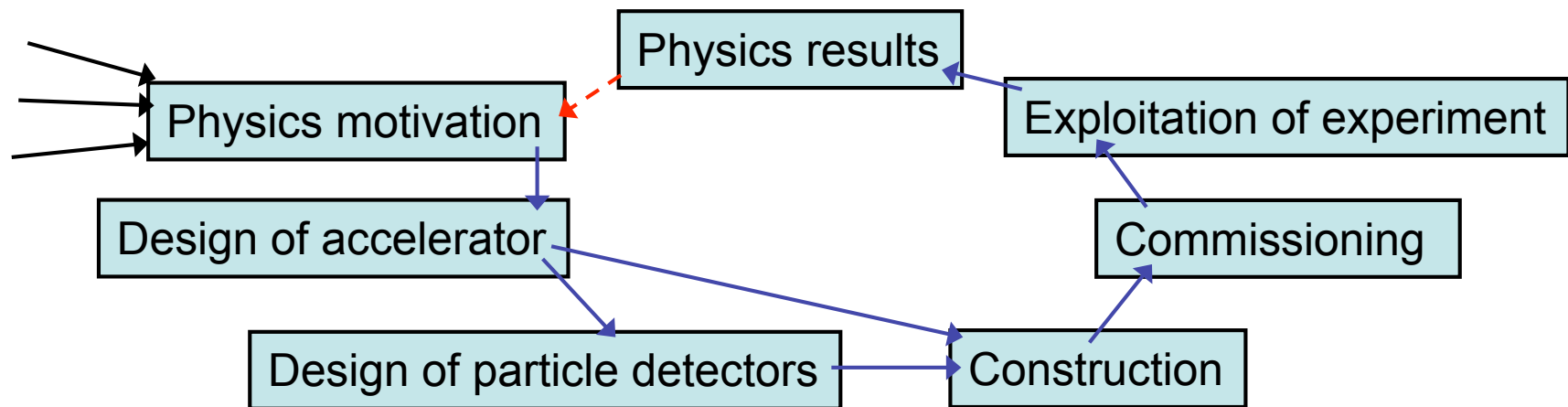
Usually theory motivates that new phenomena are to be discovered in this new domain or that the measurement of specific phenomena will give us much more insight in the current theories

Evolution of the experiments

In general new particles and phenomena are being discovered or measured at particle accelerators.

In general several types of collision experiments exist:

- Fix target experiments (eg. discovery of the quarks – proposed by J.Björken – within the proton/neutron at SLAC in 1969)
- Synchrotron experiments with colliding lepton beams (eg. LEP where electrons and positrons collided 1989-2000)
- Synchrotron experiments with colliding hadron-lepton beams (eg. HERA where electrons and protons collided 1993-2007)
- Synchrotron experiments with colliding hadron-hadron beams (eg. LHC where protons will be collided 2008-?)
- Linear colliders where leptons are colliding (eg. SLC where electrons and positrons collided 1989-1998)

Usually it takes a sequence of several experiments to reach the goal of testing the Standard Model or other theories profoundly.

Data-analysis at colliders - Prof. J. D'Hondt (VUB)

Master Level Course 2007-2008   Pag. 7

Generalities of accelerators

The main parameters of an accelerator are (relevant for this course):

- the maximum beam energy & the dipole field for circular accelerators
- the luminosity

  Rate = σ L

  σ - cross-section (units of cm^2 or barn=10^{24}cm^2), probability that an interaction will occur. If this were a game of darts, the larger the area of the dart board the more likely you will get the dart on the board.

  L - luminosity (units of cm^{-2}s^{-1} or barn^{-1} sec^{-1}), cross sectional density of the beams. The more particles per beam or the more compact (transverse) the higher the luminosity. For colliding beam, goes as the product of the two beam currents.

- the time between collisions
- the bunch length
- particles per bunch & number of bunches per ring

The luminosity (hence the event rate) at colliders has increased by large factors.

Evolution of these parameters

	LEP2 (ee)	HERA (ep)	Tevatron (p-antip)	LHC (pp)
Beam energy (GeV)	~100	~30 (e) ~920 (p)	~980	~7000
Dipole field (Tesla)	0.135	0.274 (e) 5 (p)	4.4	8.3
Luminosity (10^{30}cm^{-2}s^{-1})	100	75	171	10000
Time between collisions (ns)	22000	96	396	25
Bunch length (cm)	1.0	0.83 (e) 8.5 (p)	50	7.55
Particles per bunch (10^{10})	45	3 (e) 7 (p)	24 (p) 6 (antip)	11.5
# bunches	4	189 (e) 180 (p)	36	2808

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 9

Evolution of accelerator types

Hadron colliders can go higher in energy with the same technology compared to lepton colliders. This is due to the energy loss in synchrotron radiation when charged particles curve (hence electromagnetic radiation generated by the acceleration of ultra-relativistic charged particles through magnetic fields)

- energy loss per turn: $\Delta E \sim E^4/(r \, m^4)$
- E is the energy of a particle with mass m which described a circle with radius r
- the mass of a proton is 2000 times larger than that of an electron.

Hadron colliders can reach higher energies with the same electric RF fields compared to leptons, hadron colliders are therefore used to explore a new energy regime but lepton colliders are used to perform precise measurements.

SPS (discovery W/Z bosons) → LEP (measurement W/Z bosons)
LHC (discovery of SUSY ?) → ILC (measurement of SUSY ?)

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 10

Evolution of accelerator types

- Hadron colliders reach higher beam energies compared to lepton colliders, but the partons within the hadrons which collide have only a fraction of the beam energy. The leptons carry the full beam energy.

- Strong progress over the last decades in accelerator performance, but slowed down recently due to the scale of the projects and no real revolution in the technology which is being applied.
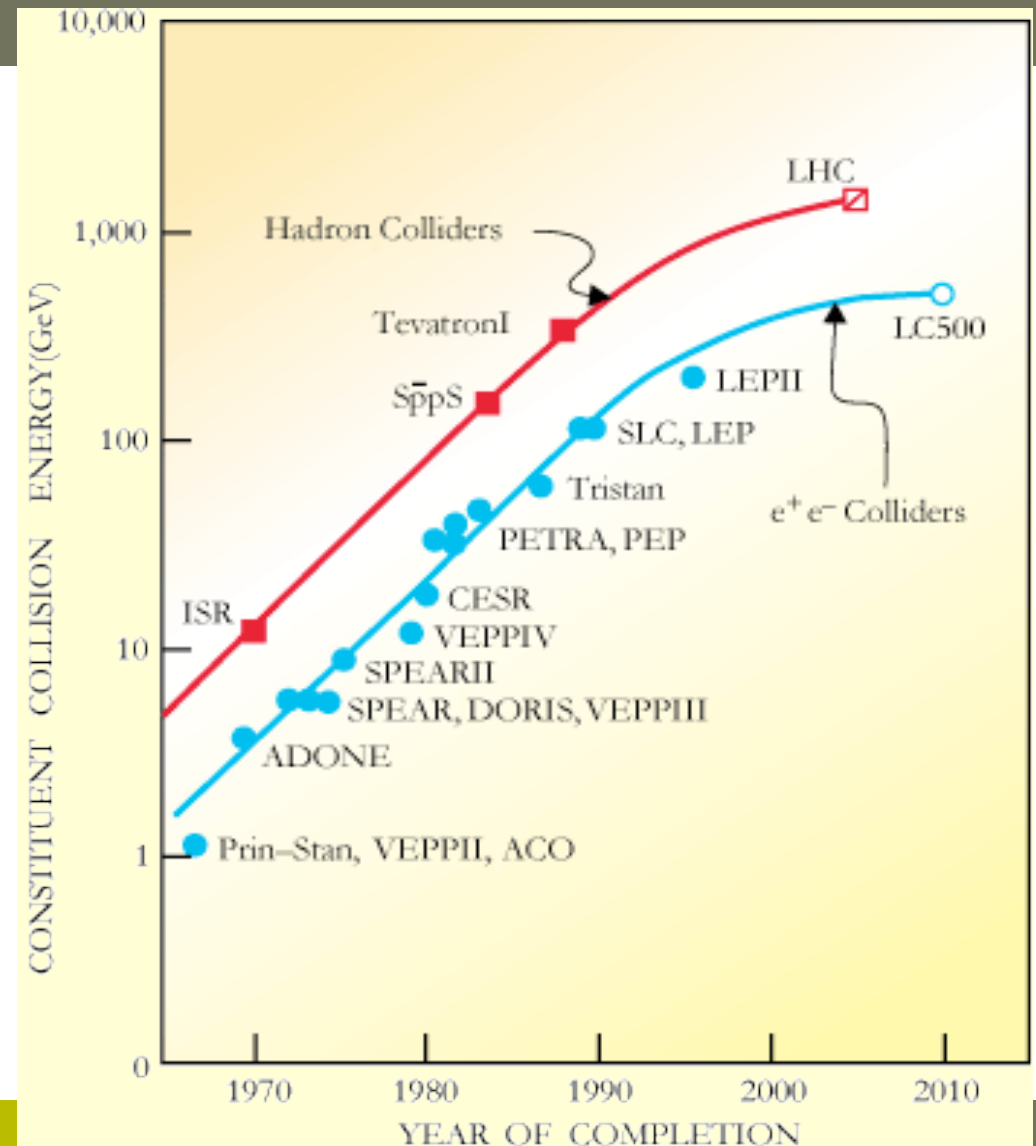
Evolution of the detectors

Particle detectors have been evolving from simple devices with few channels to very complex multi-layer systems with several millions of read-out channels.

- the data acquisition (DAQ) systems have evolved accordingly
- the main reason for this evolution is the ever increasing particle rate in the final state of the collisions we provoke in the laboratory
- together with the usually custom detector electronics
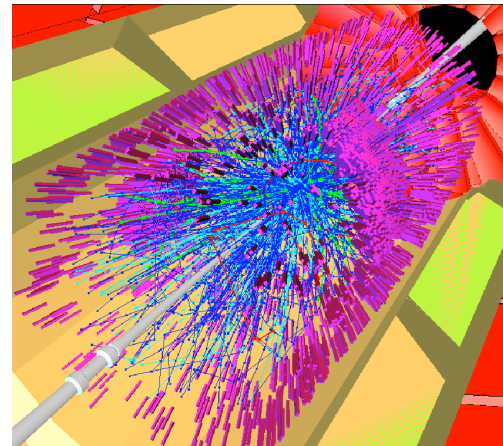
More events to be looked at in less time and also more complex events, give stronger requirements on the read-out and trigger systems of our detectors.

Data-analysis at colliders - Prof. J. D'Hondt (VUB)

Master Level Course 2007-2008

Pag. 12

Event rate at early experiments

- ## Bubble Chambers, Cloud Chambers, etc. (4π)
  - DAQ was a stereo photograph!
  - Effectively no Trigger:
    - Each expansion was photographed based on the accelerator cycle
    - The High level trigger was *human* (scanners).
  - Slow repetition rate.
    - Only most common processes were observed.
  - Some of the high repetition experiments (>40 Hz) had some attempt at triggering.

- ## Emulsions still used in some neutrino experiments (eg CHORUS, DONUT).
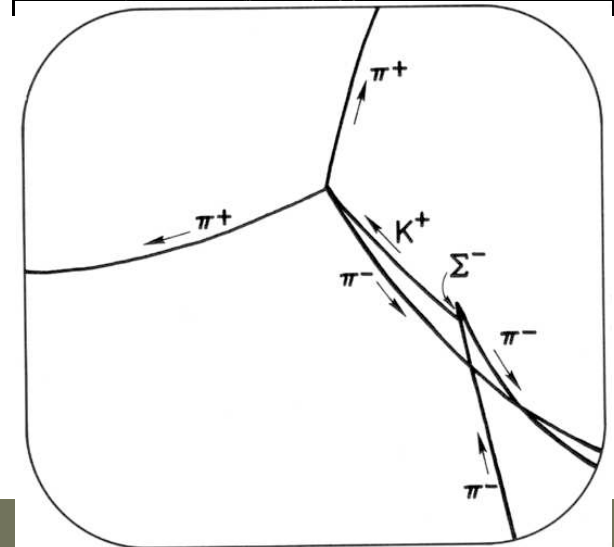
Data Acquisition

Data acquisition systems in general obtain information of a physical property, usually multi-dimensional, from the signals of detectors. Hence the system contains electronic components which can convert specific measured parameters into an electrical signal. The acquired data is hereafter stored on a disk for further analysis by a computer. The DAQ hardware are electronic devices which provide an interface between the signal and the computer.

In the language of HEP experiments, the DAQ system collects the signals of the millions of detector channels and allows the trigger system to analyze the data stored.
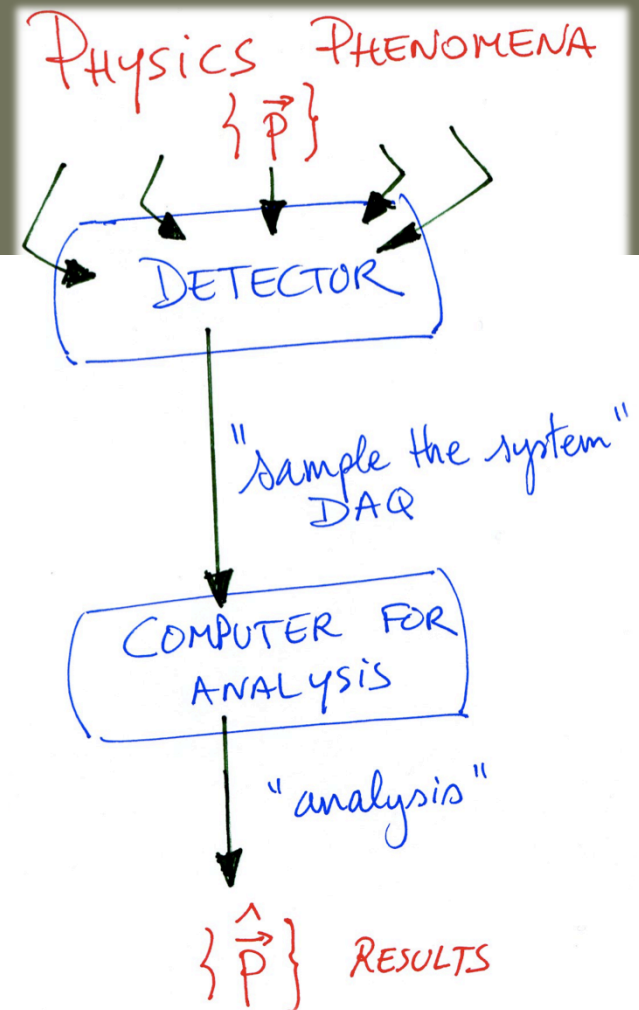
Electronic signals (reminder)

The signal is in most cases a short voltage pulse... how does this looks like?

First some basics:

$V = I \cdot R$
(Ohm's law)

$I \equiv \dfrac{dq(t)}{dt}$

"ground"

$I_1 = I_2 + I_3$
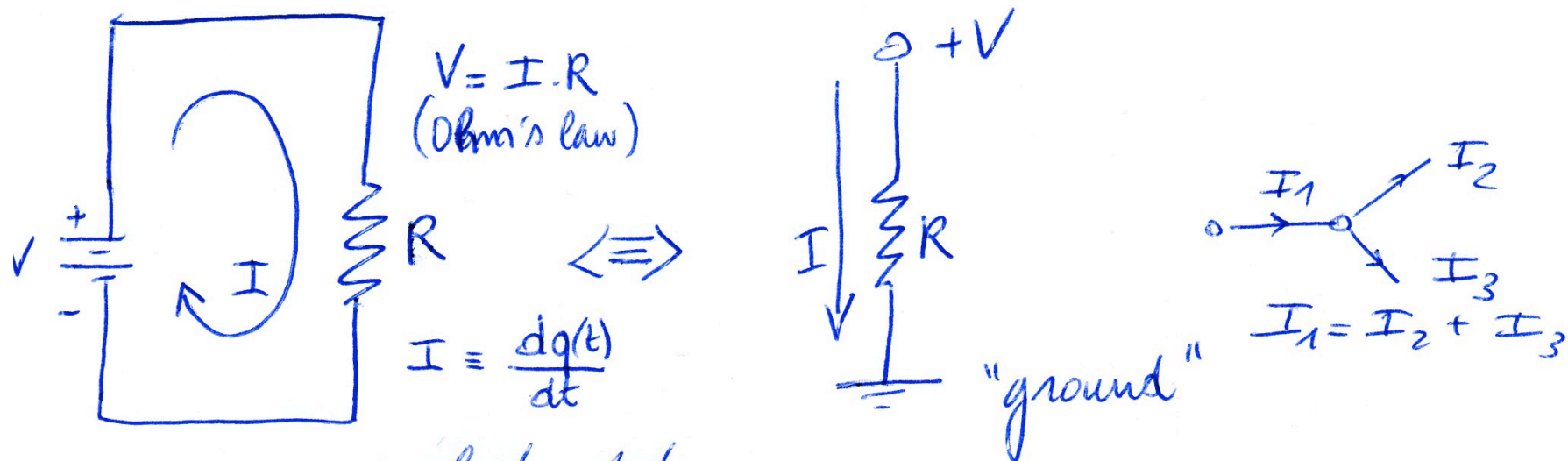
→ equivalent notation

Only the relative voltage between points matters

⇒ all relative to the "ground"

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 15

Electronic signals (reminder)

Resistors:

serie: $R = R_1 + R_2$
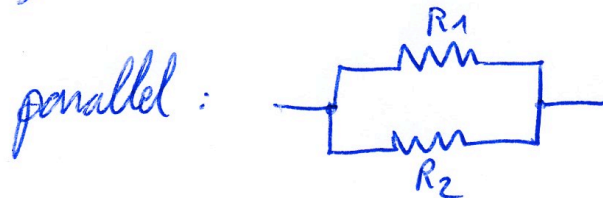
parallel: $\dfrac{1}{R} = \dfrac{1}{R_1} + \dfrac{1}{R_2}$

from $\begin{cases} IR = I_1 R_1 = I_2 R_2 \\ I = I_1 + I_2 \end{cases}$

Capacitors (stores charges, but no charges go through it)

$C \equiv \dfrac{q}{V} \quad \longrightarrow \quad \dfrac{\text{"charged stored"}}{\text{"potential difference"}}$

serie: $\dfrac{1}{C} = \dfrac{1}{C_1} + \dfrac{1}{C_2}$

parallel: $C = C_1 + C_2$

inverse behaviour of R due to Ohm's law

Electronic signals (reminder)

Assume voltage change with time (AC circuit)
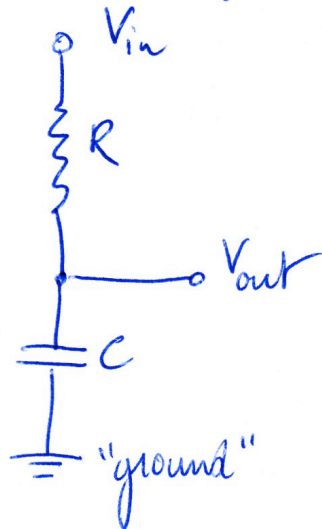
$$V_{in}(t) = 0 \qquad t \leq 0$$
$$\qquad\quad = V \qquad t > 0$$

(no change on capacitor at $t=0$)

$$V_{out}(t) = \frac{q(t)}{C} \qquad \text{"voltage drop capacitor"}$$

$$IR = \frac{dq(t)}{dt} \cdot R \qquad \text{"voltage drop resistor"}$$

$$\Rightarrow V = V_{out}(t) + \frac{dq(t)}{dt} \cdot R$$

$$\Rightarrow V_{out}(t) + R \cdot C \cdot \frac{dV_{out}(t)}{dt} = V$$

Solution of this differential equation: $V_{out}(t) = V \cdot \left(1 - e^{-t/RC}\right)$

In real life the voltage at the input has some distribution and will appear only for a short period.

→ Fourier analysis (cfr. other courses)

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 17

Electronics

The signal appears as a peak in the time-dependent voltage evolution:

$$V_{out}(t) = \int_{-\infty}^{t} V_{in}(t') \ \frac{\exp\left(-\frac{t-t'}{RC}\right)}{RC} \ dt$$

with $V_{in}(t')$ a Gaussian function
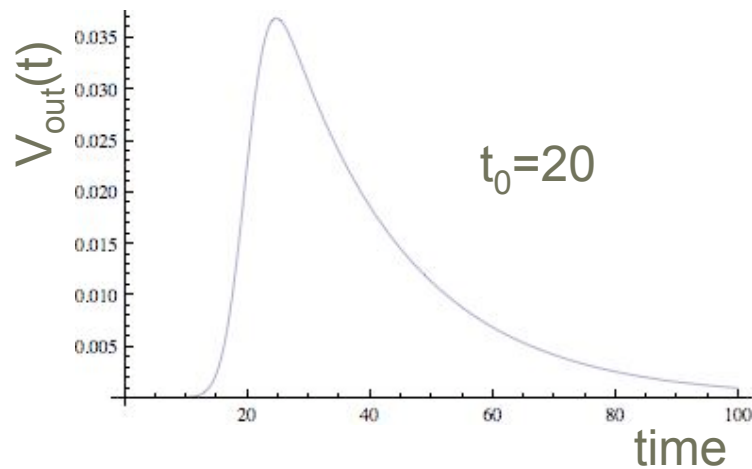
```
In[69]:= RC = 15;
t0 = 20;
Plot3D[Integrate[1/ (Sqrt[2*3.1415*sigma^2]) *Exp[-0.5*((x-t0)/sigma)^2]*
    Exp[-(t-x)/(RC)]/(RC), {x, -999 999, t}], {t, 0, 100}, {sigma, 0.1, 15}]
```

Out[71]=

$V_{out}(t)$

sigma

time
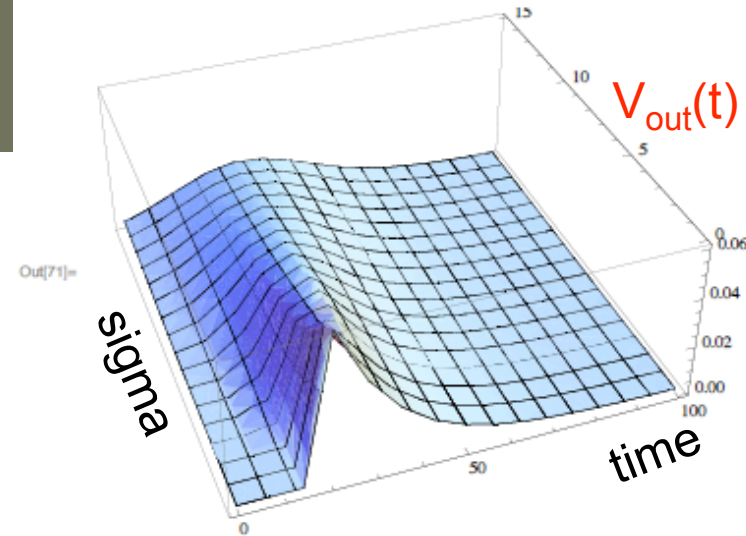
```
In[75]:= sigma = 3;
t0 = 20;
Plot3D[Integrate[1/ (Sqrt[2*3.1415*sigma^2]) *Exp[-0.5*((x-t0)/sigma)^2]*
    Exp[-(t-x)/(RC)]/(RC), {x, -999 999, t}], {t, 0, 100}, {RC, 5, 50}]
```

Out[77]=

$V_{out}(t)$

R*C

time

$t_0 = 20$

$V_{out}(t)$

time

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 18
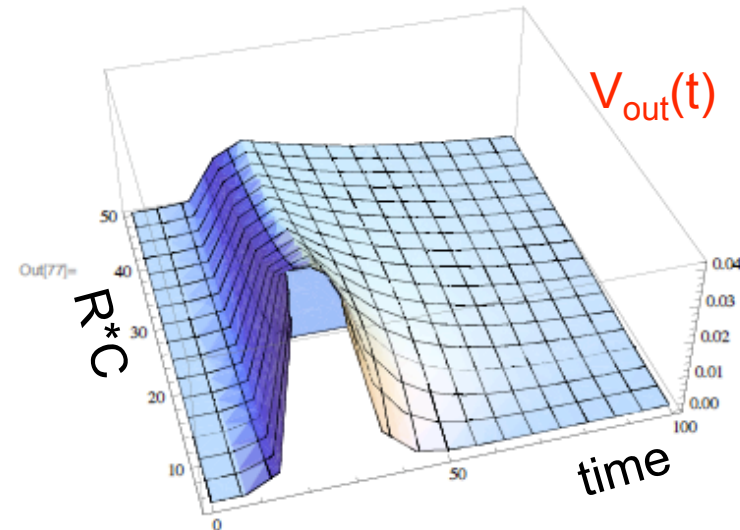
Cross section & event rate

- The cross section of specific processes is changing when going to higher energies.

- The total cross section in hadron collisions is increasing according to the measurements made (extrapolation to higher energies).

- Together with the luminosity increase this results in a higher event rate R=$\mathcal{L}\,\sigma$, hence important technical challenges for the design of the experiments.

proton-proton

proton-antiproton
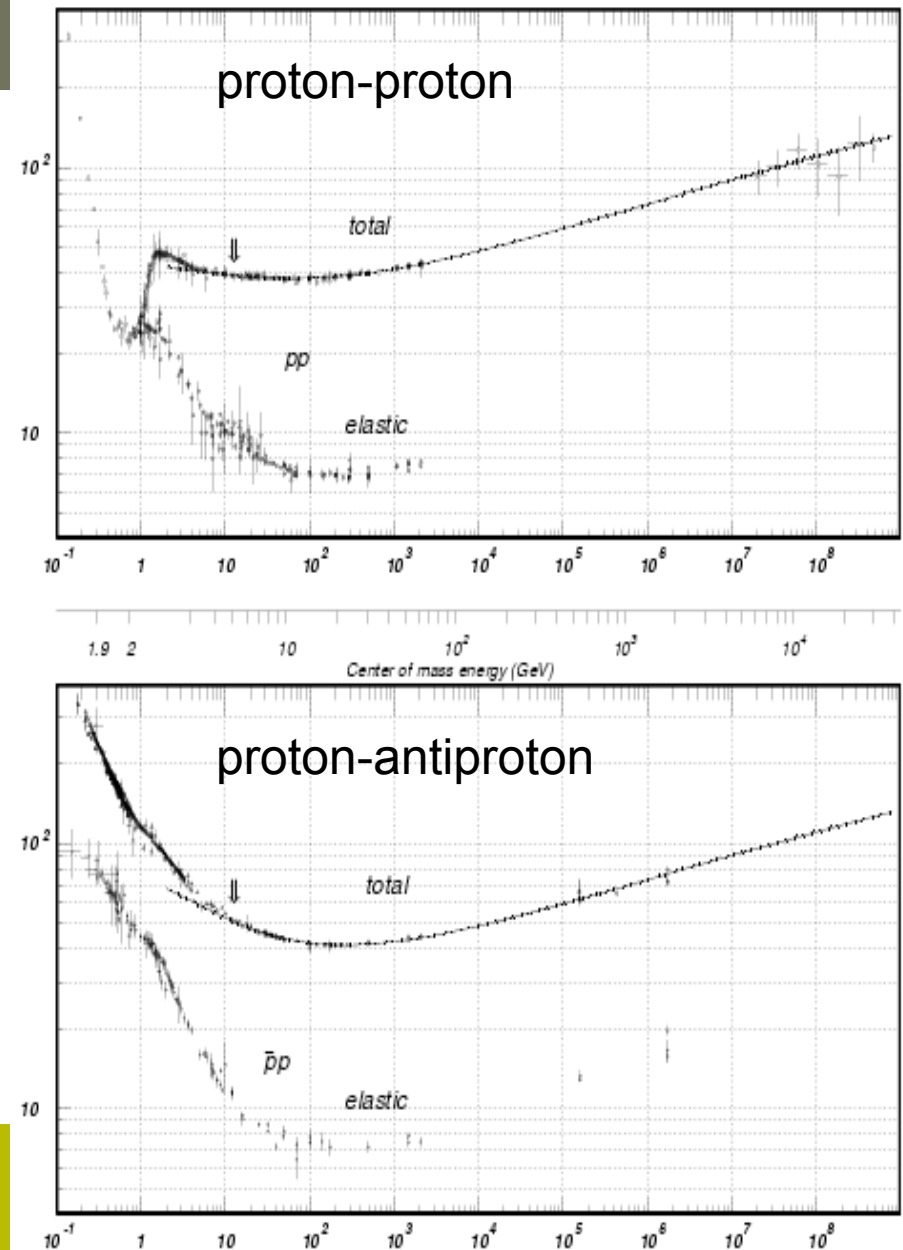
Cross Section (mb)

Cross section & event rate

- The total cross section at the LHC (14 TeV) ranges between 90 and 130 mb

JINST 3 S08007 (2008)

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag.
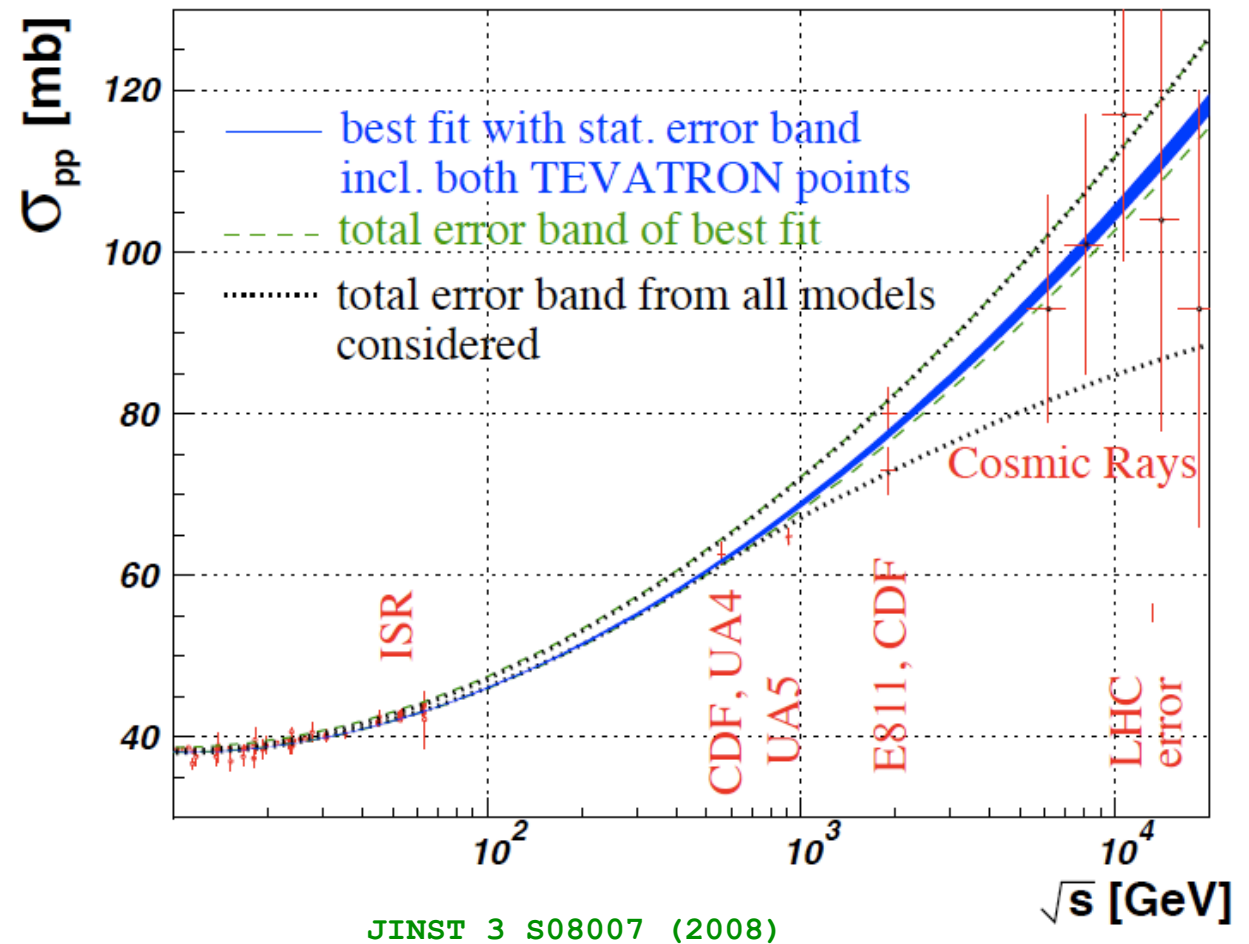
Cross section & event rate

- In general the cross sections of Standard Model processes is increasing when going to higher centre of mass energies.

- The processes of interest to be discovered (Higgs bosons, SUSY, etc.) have much lower cross sections compared to the Standard model processes, but as they require the creation of particles at higher mass scales they usually increase more rapidly compared to Standard Model processes with the centre of mass energy.

References for LHC & experiments:

http://www.iop.org/EJ/journal/-page=extra.lhc/jinst
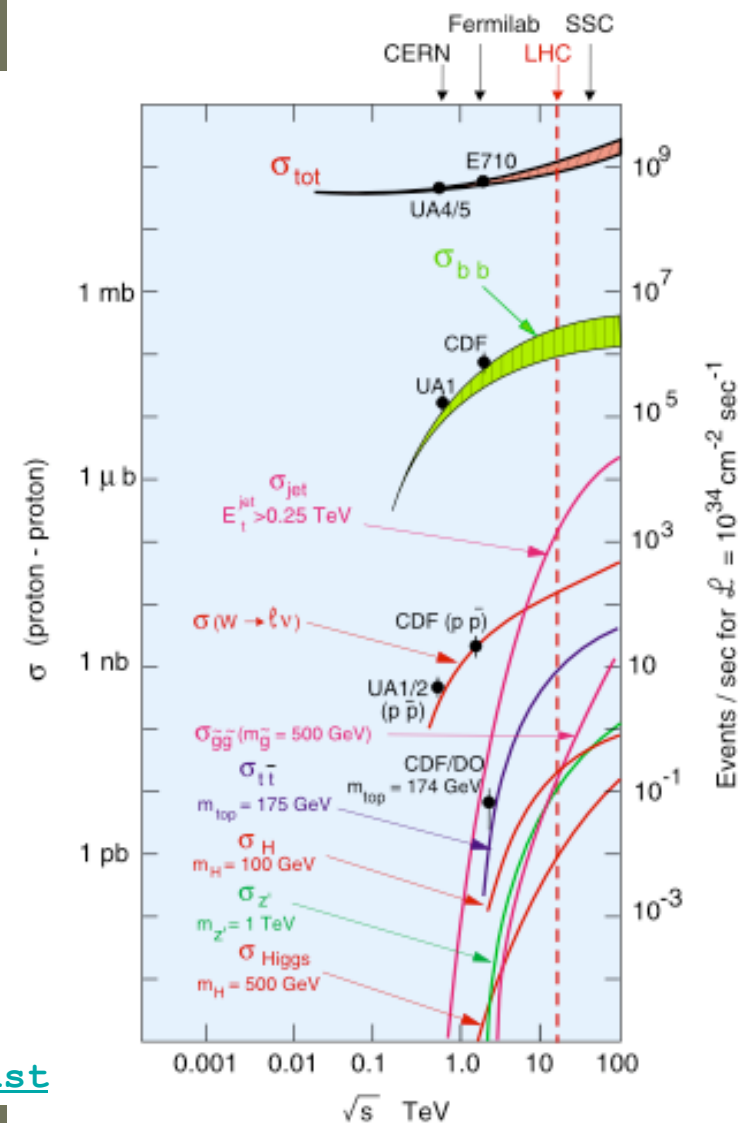
Goal of the trigger setup

- Goal of Trigger and DAQ systems is to store the maximize amount of data for the desired process with minimal cost (=time)
- Relevant efficiency is for events that will be useful for later analysis
- Low rate process (eg. Higgs production at Tevatron or LHC), try to accept them all in the trigger, hence maximize their selection efficiency.
- Dead-time induced do to fluctuations when rate into a stage of trigger (or read-out) approaches the rate it can handle.

Efficiency to be maximized

$$\mathcal{E} = \mathcal{E}_{operation} \cdot \mathcal{E}_{trigger}^{Higgs} \cdot (1 - dead\,time)$$

$$\mathcal{E}_{t_n} = \frac{\#\ accepted\ Higgs}{\#\ produced\ Higgs}$$

Simple case (no buffer)

Dead-time = "Input Rate" * "Execution Time"

Execution time needed for the "averaged" process to be recorded to tape/disk.

Buffering incoming data reduces dead-time, more buffering less dead time
- If <Incoming Rate> is larger than 1/<Execution Time>, dead no matter what!

Minimizing dead-time helps all processes
- 1% of machine time * 1 year = exploitation time lost (=money lost)

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 22

Trigger on what?

Accept specific decays modes

– high P_T leptons from W, Z, top, W/Z +Higgs QCD: High Et jets

– $B_s/\psi \rightarrow \mu\mu$, medium p_T leptons for B physics

Reject:

– lower P_T objects (QCD)

Select on object/event kinematics:

– E_T of Calor Tower (cluster) or missing E_T

– muon P_T (+ track P_T)

– track P_T (+ impact parameter/ detached vertex)

Et = 306.26 GeV

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

23
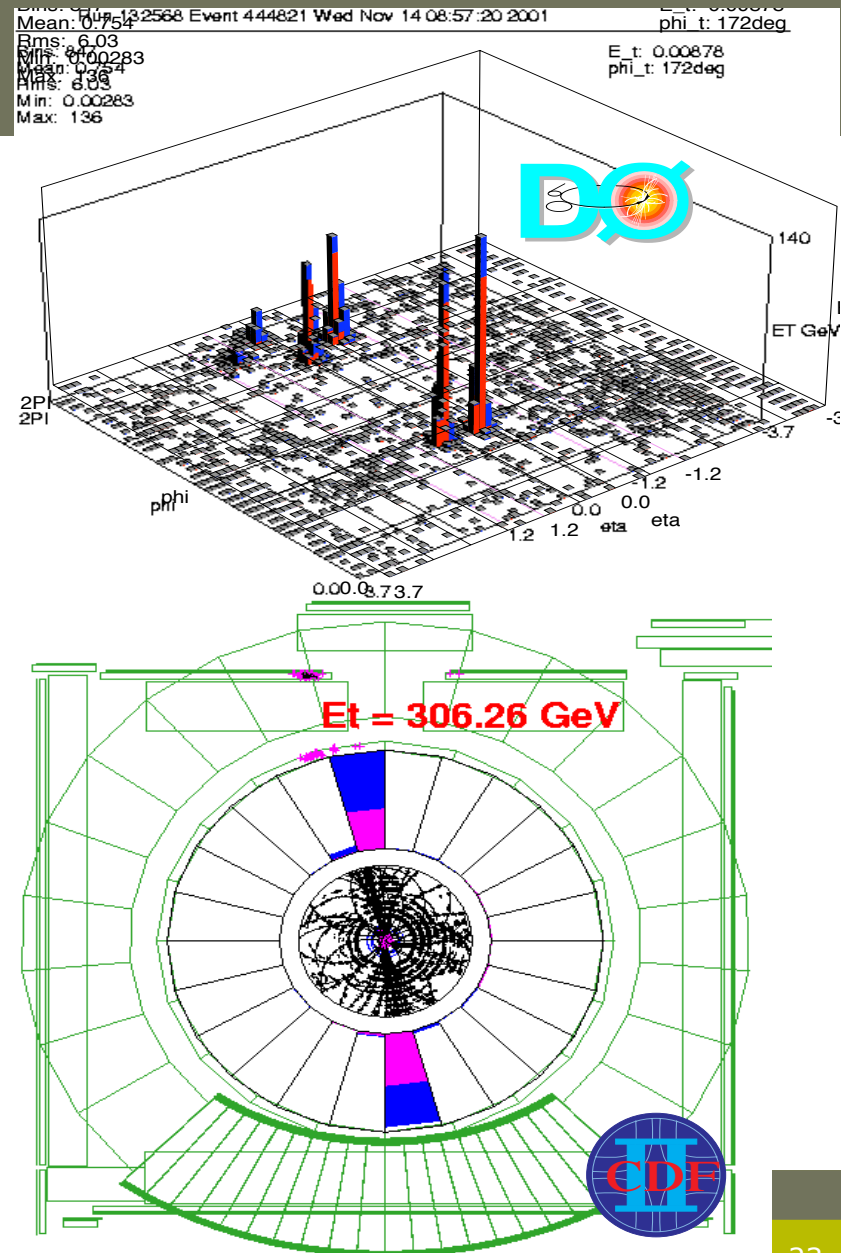
Hadron collisions

What do you expect to see in the detector?

Warning: simplified schematic view which is not to scale

p

p/$\overline{p}$

The incoming particles are contained into bunches of for example 10^{11} hadrons (eg. protons). The proton is a bag full of partons which have probability densities depending on Björken-x (longitudinal momentum fraction) and the Q^2 (momentum transfer), hence the partons collide with each other, not the full hadron.

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 24

Hadron collisions

What do you expect to see in the detector?
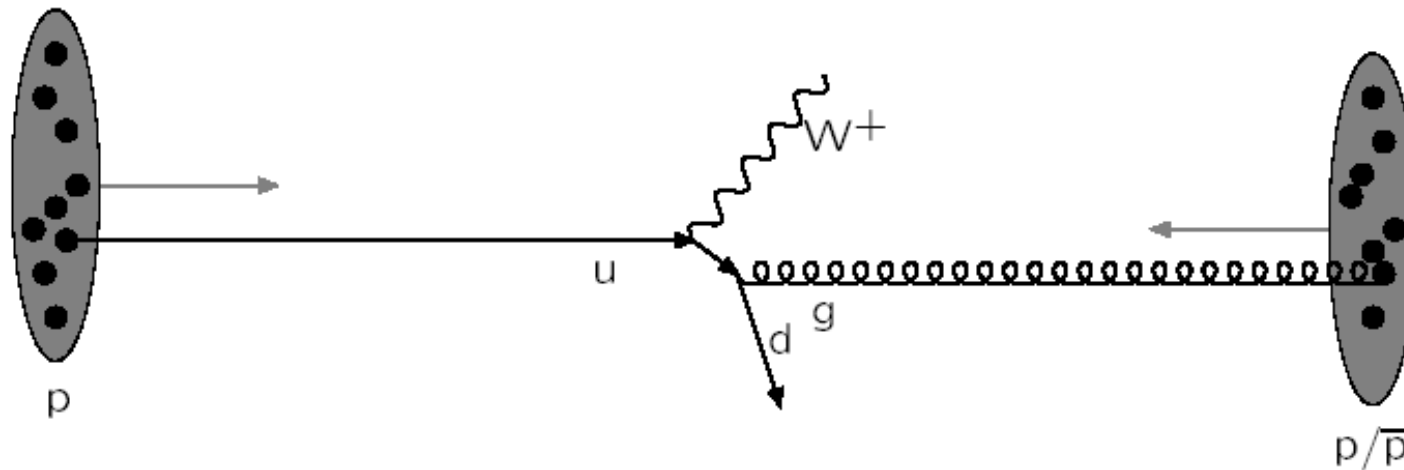
The partons (being quarks or gluons) interact reflecting the hard subprocess described by matrix elements obtained from quantum field theory calculations (Feynman diagrams).

Hadron collisions

What do you expect to see in the detector?
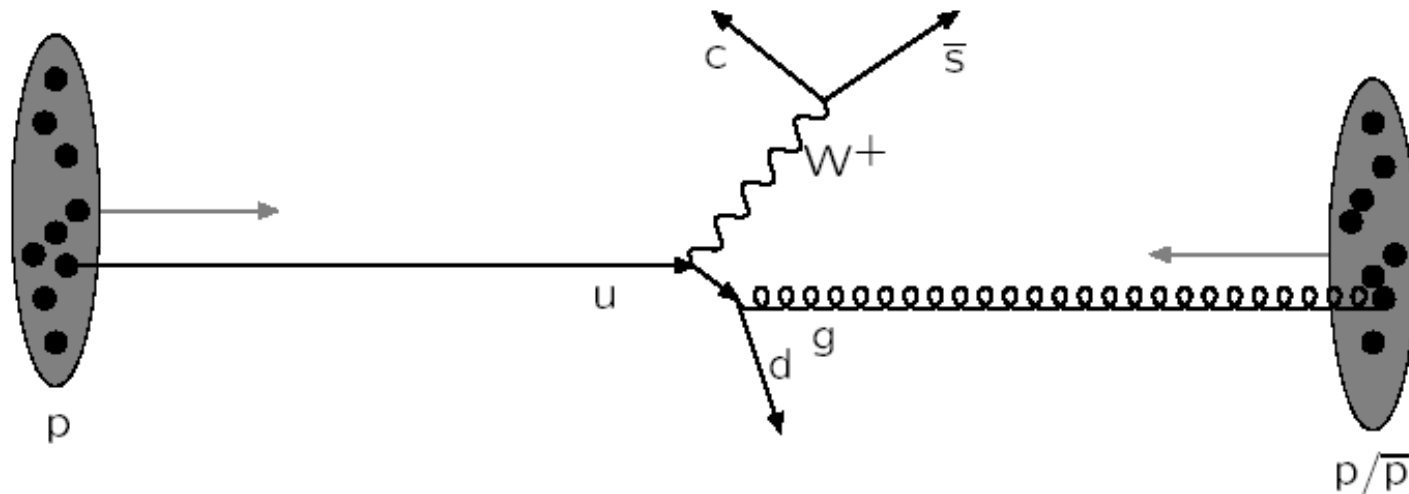
The heavy resonances decay, for example the W boson with a mass around 80 GeV. This is still correlated with the hard subprocess and is usually calculated with the same tools.

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 26

Hadron collisions

What do you expect to see in the detector?
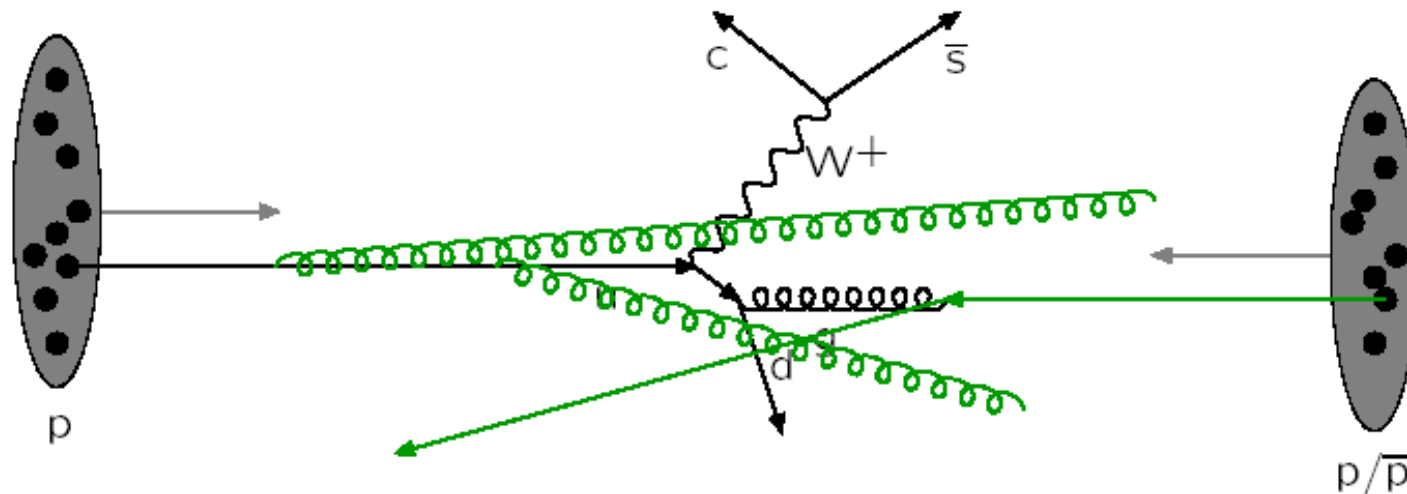
The partons could radiate coloured particles before interacting with each other, this is called initial state radiation (=space-like parton showers).

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 27

Hadron collisions

The partons could radiate coloured particles after interacting with each other, this is called final state radiation (=time-like parton showers).

Hadron collisions
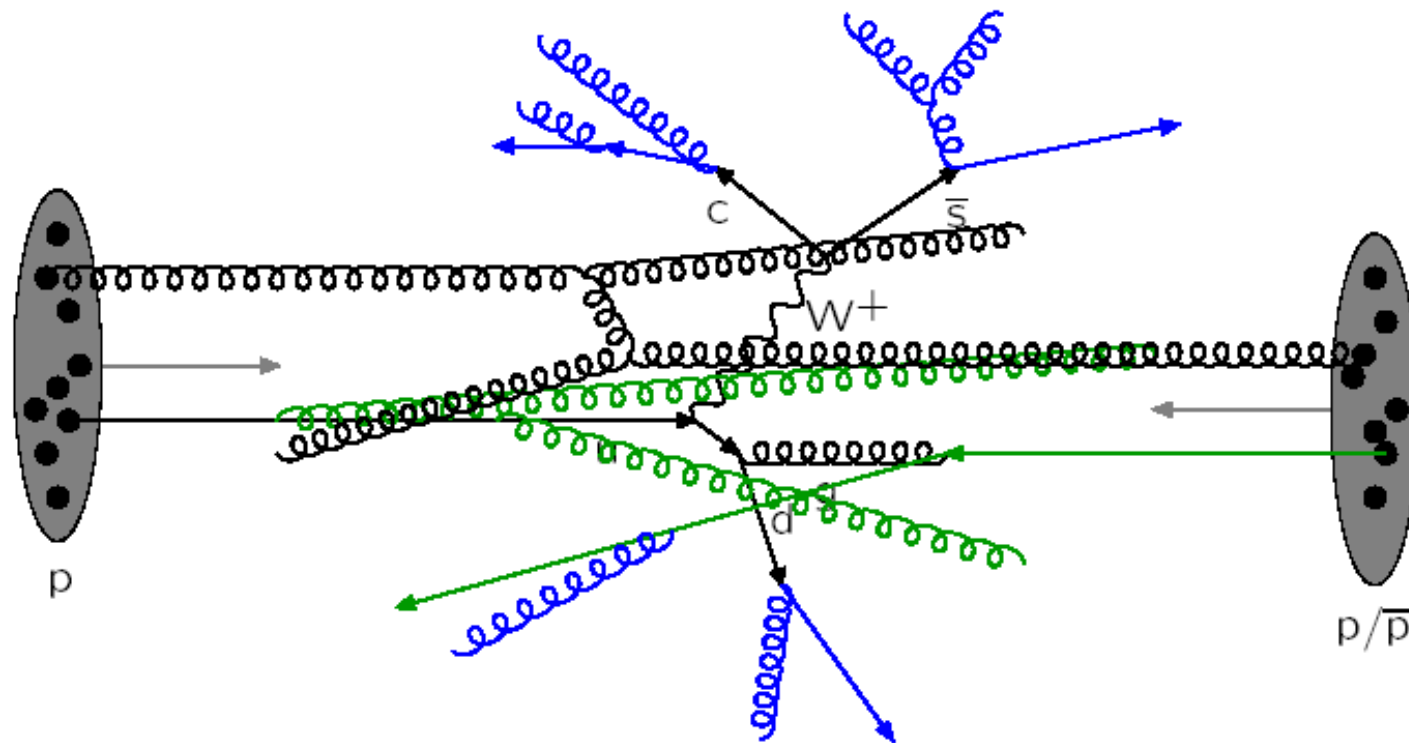
The hadrons consist out of several partons, hence there can be more than one parton-parton interaction. We speak about multi-parton interactions.

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 29
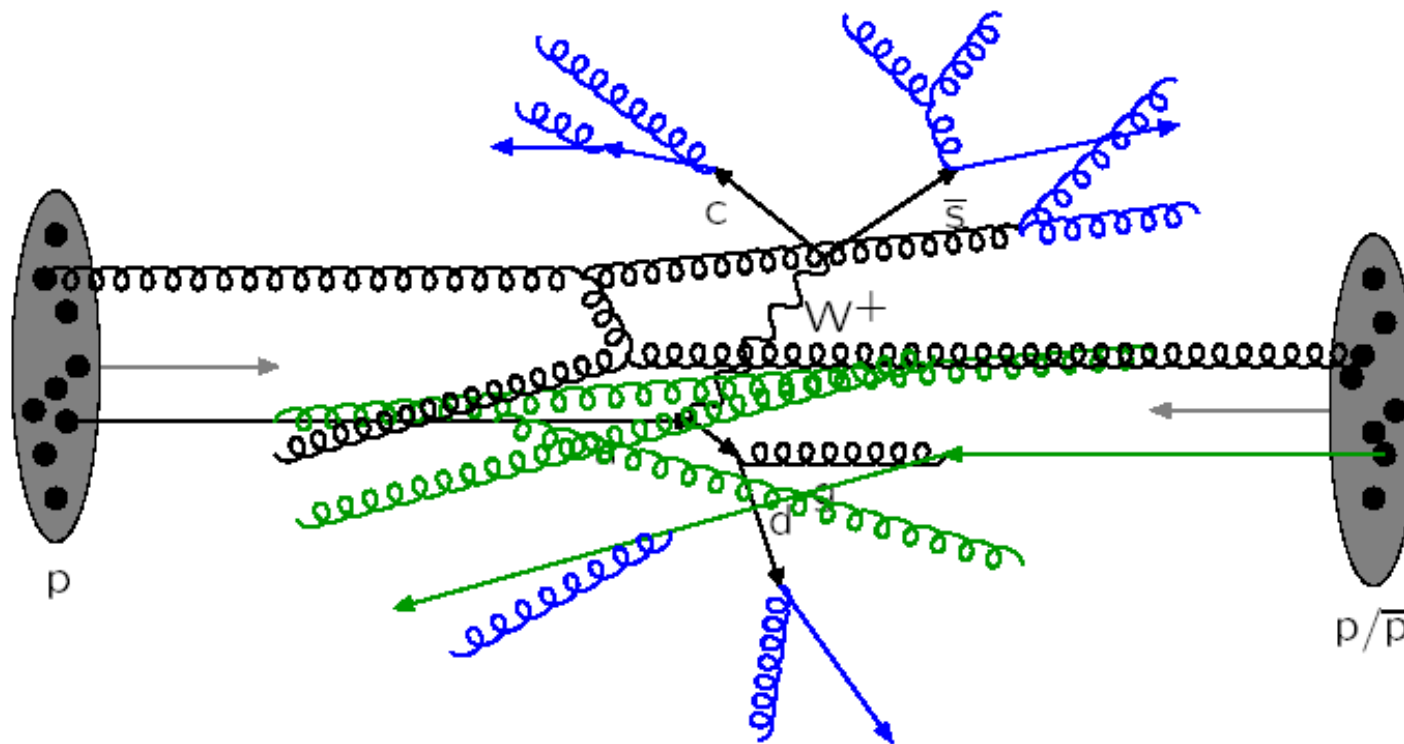
Hadron collisions

And in these extra parton-parton interactions there can also be initial and final state radiation of extra partons. More blue and green lines…

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 30
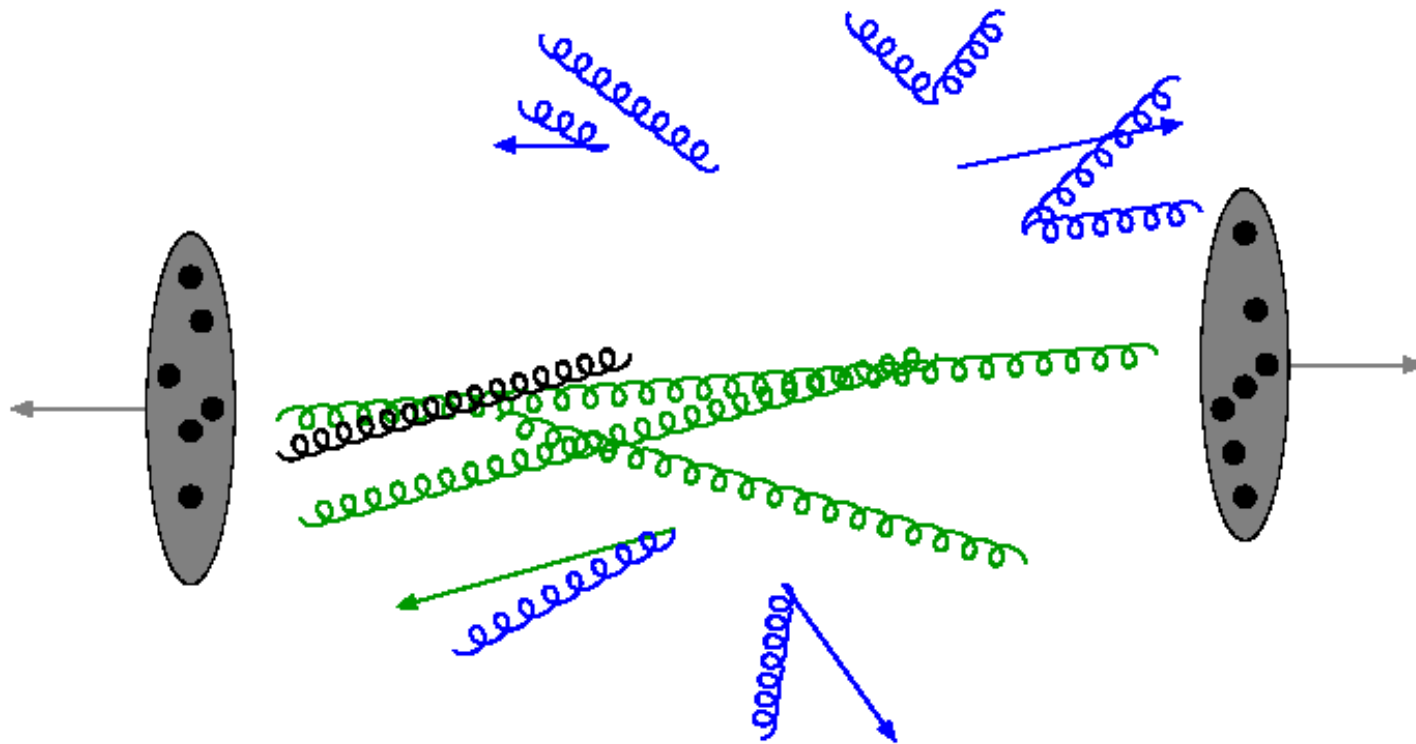
Hadron collisions

After the collision of course we do not "see" the internal lines, we only "see" the outgoing partons and the beam remnants.

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 31
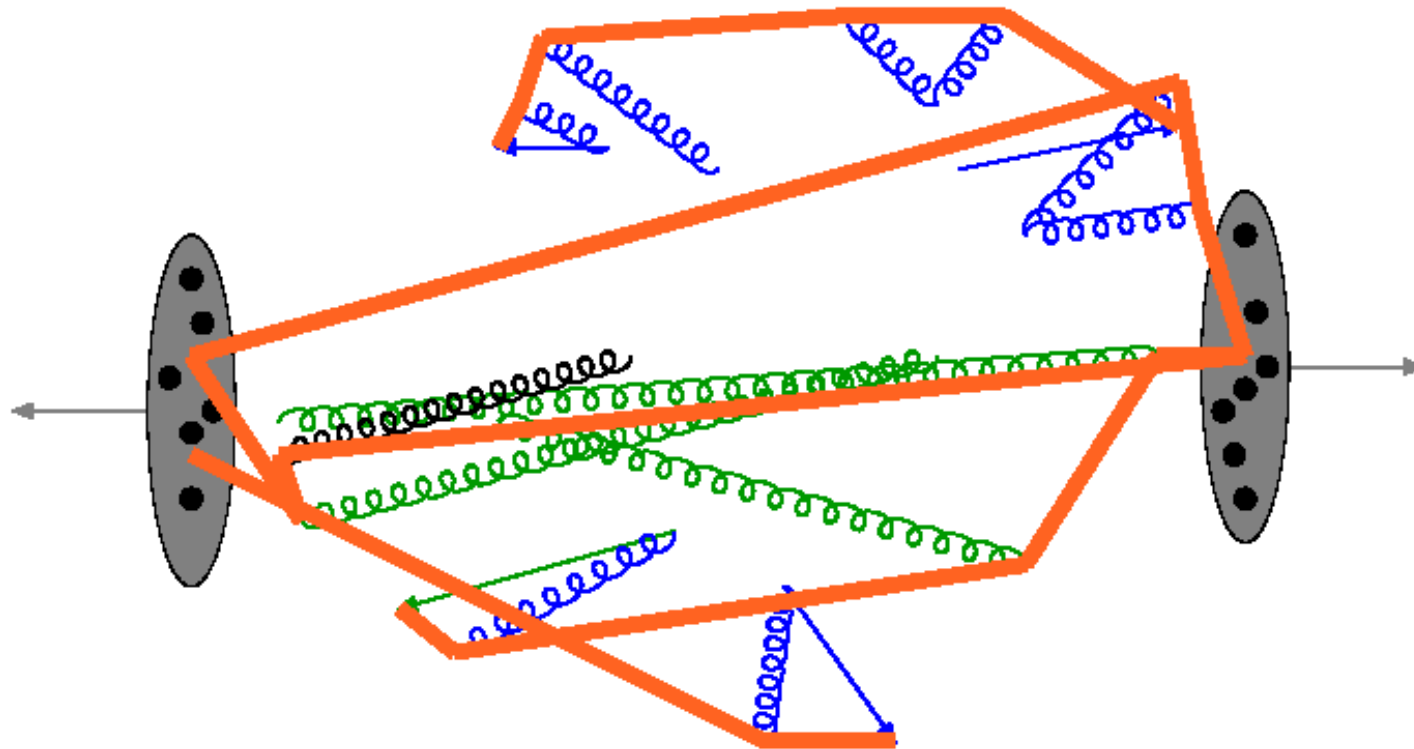
Hadron collisions

All partons are connected by colour confinement strings (colour field between partons as you have an electromagnetic field between charged particles).

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 32
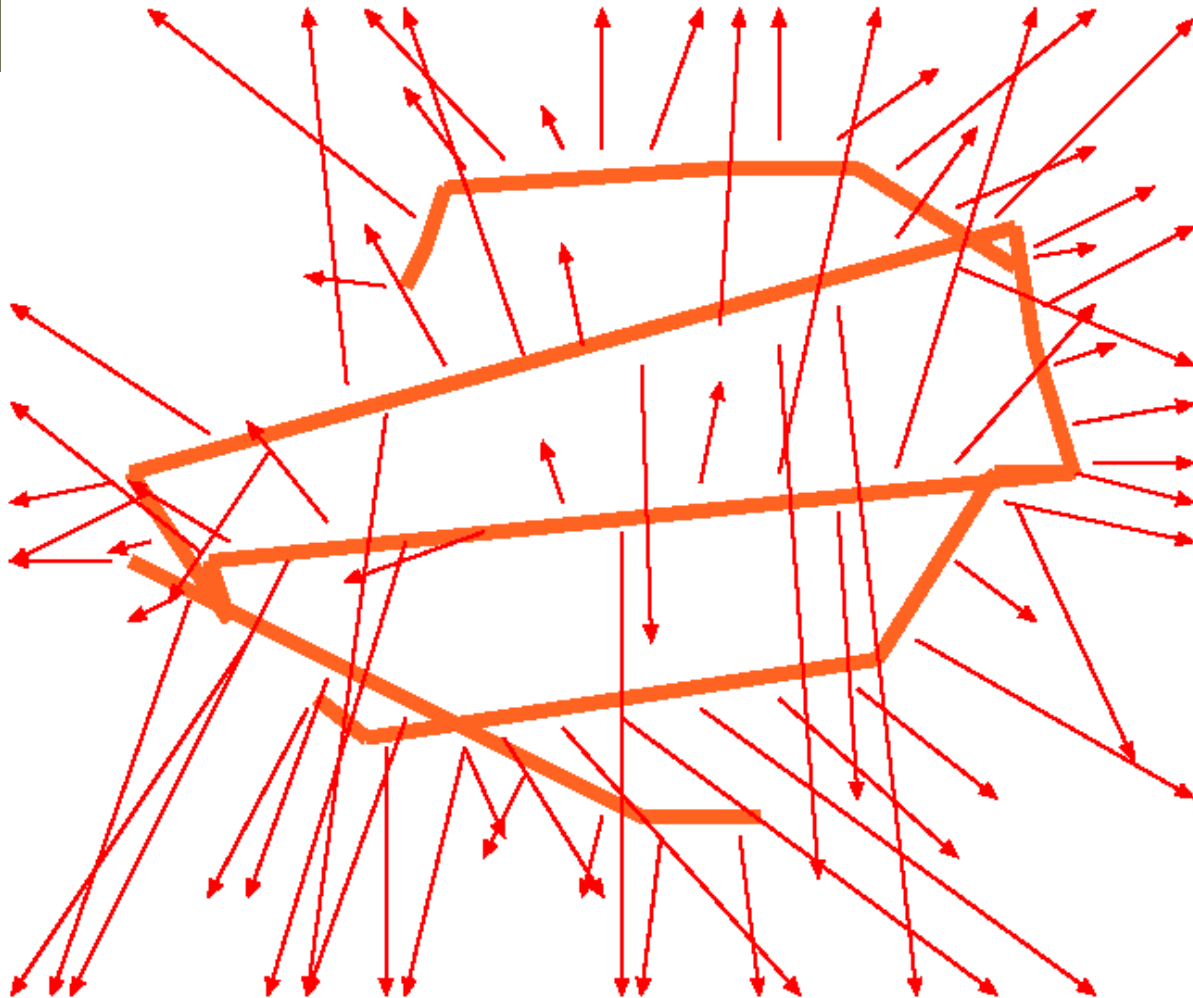
Hadron collisions

The strings will fragment to produce in general hadrons, hence colour-neutral particles.

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 33
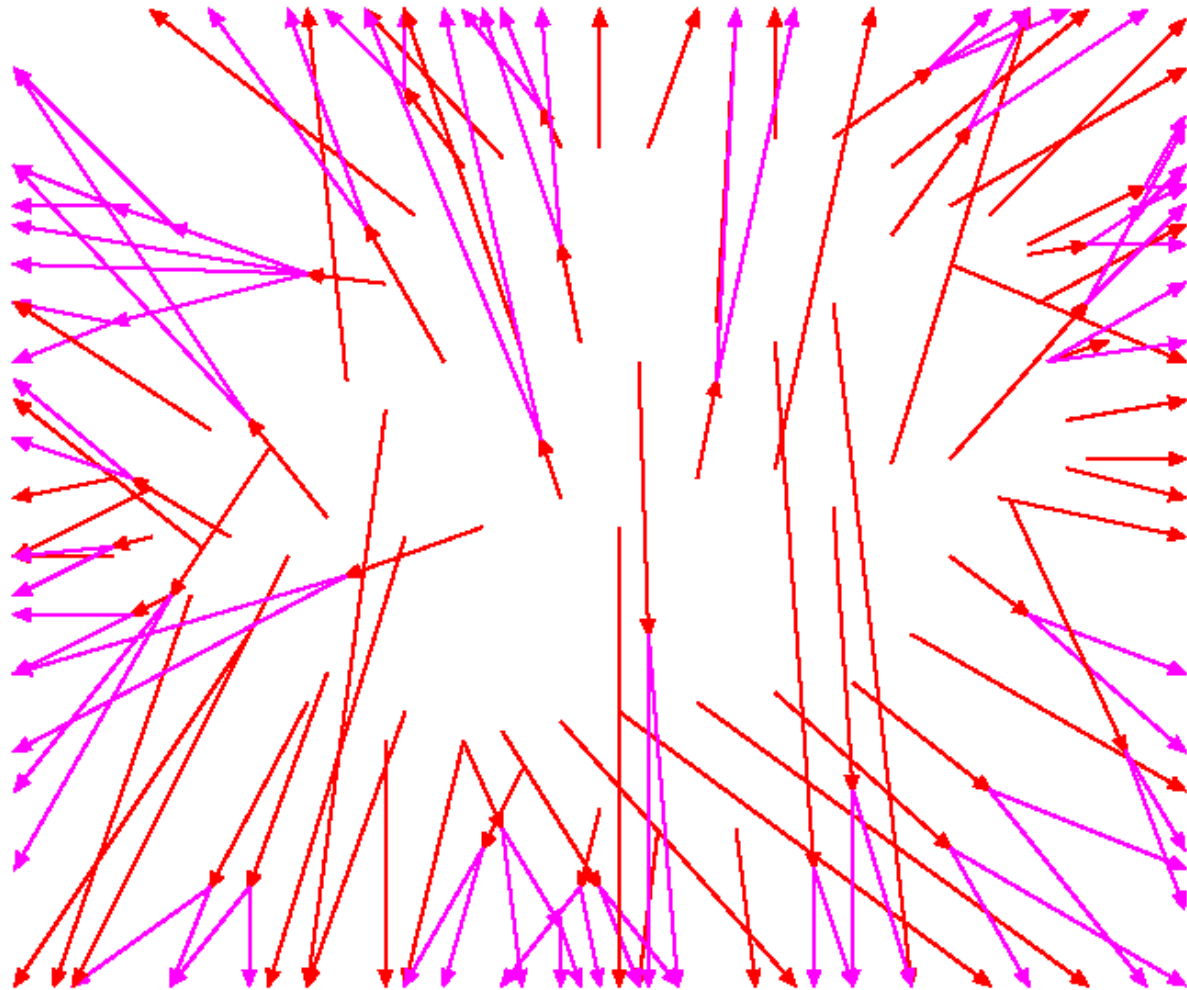
Hadron collisions

A lot of hadrons are not stable and will decay further into stable particles which we will observe in our detectors.

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 34

The Large Hadron Collider (LHC)

The next in line accelerator is the Large Hadron Collider at CERN which will collide protons up to centre of masses of 14 TeV

Several theory based motivations are driving the design of both the accelerator and the particle detectors

- Search for the Higgs boson motivated by the spontaneous symmetry breaking mechanism [Brout, Englert, Higgs] and the divergences of cross section of the WW→ZZ scattering process
- Or alternative models explaining the origin of mass
- The search for particles able to solve the Dark Matter problem in the Universe, eg. by introducing Supersymmetry in our theories
- Try to incorporate gravity into the Standard Model
- Many exotics …

The accelerator will start taking data in 2009

Data-analysis at colliders - Prof. J. D'Hondt (VUB)

Master Level Course 2007-2008

Pag. 35
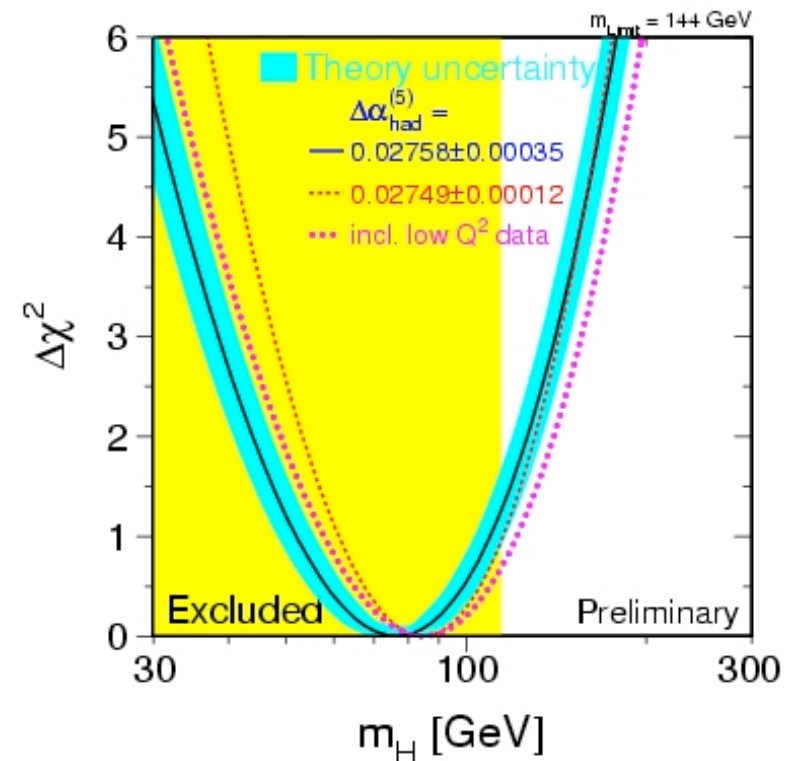
The timeline of the project

The history of this accelerator is long:

1982: first studies for the LHC project

1983: Z0 & W boson discovered at the SPS proton anti-proton collider

1989: start of LEP operation (Z boson factory)

1994: approval of the LHC by the CERN Council

1996: final decision to start the LHC construction

1996: LEP operation to produce W boson pairs (W boson factory)

2000: last year of LEP operation above 100 GeV beam energy

2002: LEP equipment removed

2003: start of the LHC installation

2005: start of LHC hardware commissioning

2008: expected LHC commissioning with beam

Large Hadron Collider basics

Colliding protons at the LHC √s = 14 TeV

foreseen start this year!

LHC - B
Point 8

CERN

ATLAS
Point 1

ALICE
Point 2

CMS
Point 5

SPS

TI 8

TI 2

LHC - B

ATLAS

ALICE

CMS

LEP/LHC

circumference ~26.7km
tunnel diameter 3.8m
depth 70-140m
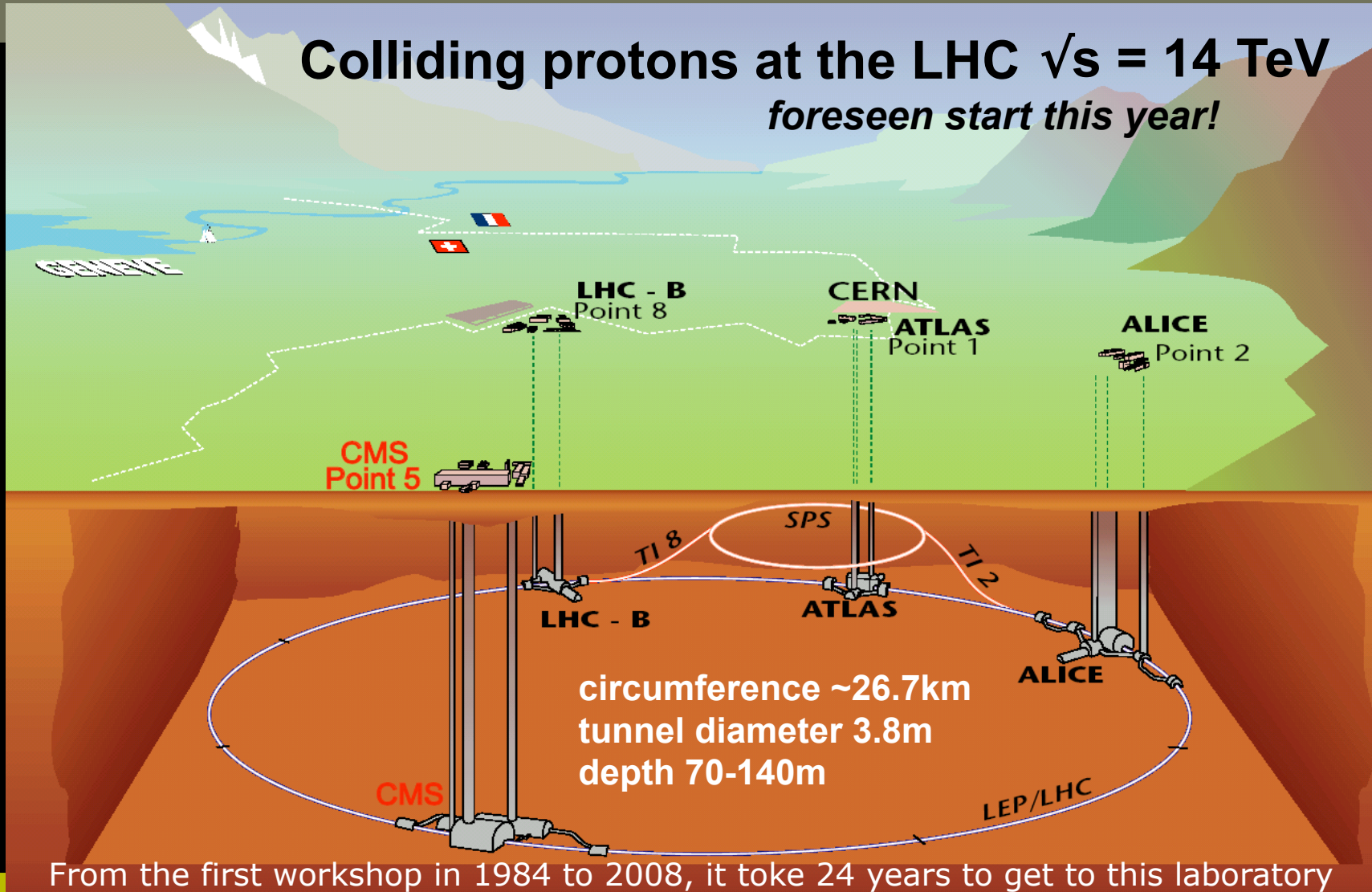
From the first workshop in 1984 to 2008, it toke 24 years to get to this laboratory

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 37

Laboratory is getting ready!

... new era for physics !

Higgs boson production at LHC

- How does the Higgs boson shows itself: several production and decay possibilities

- Would like to have 20-30 Higgs boson events per year at the highest masses (lowest cross section)

- Hence need 10^{34} cm^{-2}s^{-1} of luminosity and ~2800 bunches in the 27km long accelerator

decay branching ratios

production cross-section
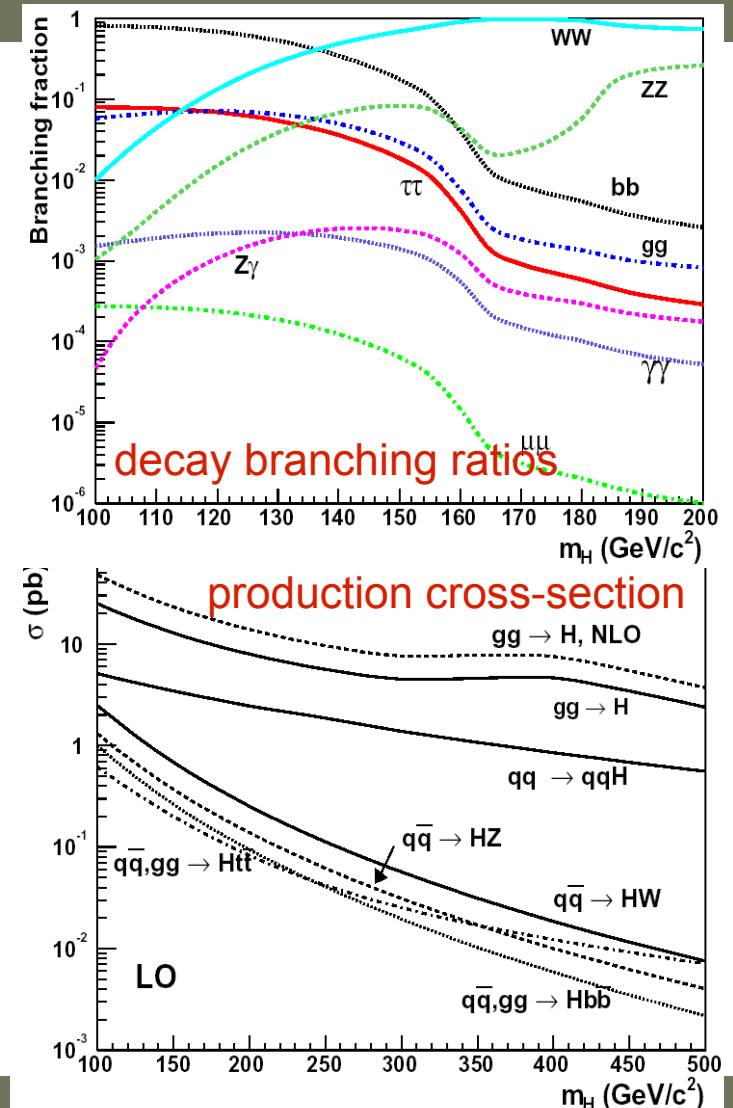
Beam crossings evolution

- Also empty bunch spaces are needed in order to be able to dump the proton beam properly
- Therefore need ~3600 bunch spaces (= 25 ns)
- 25 ns = 27 km / (c . 3600)

LEP: e⁺e⁻ Crossing rate 30 kHz

22μs

Tevatron Run I

3.5μs

Tevatron Run II

396ns

LHC: pp Crossing rate 40 MHz

25ns

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 40
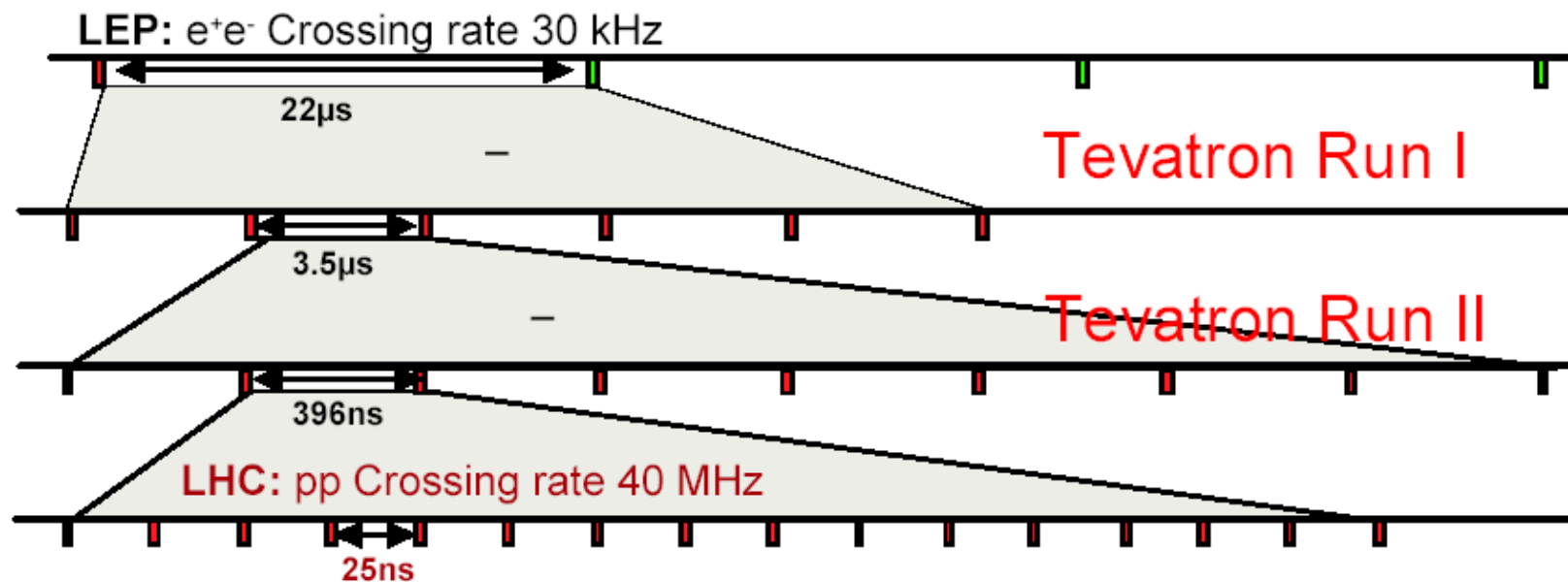
The Large Hadron Collider (LHC)
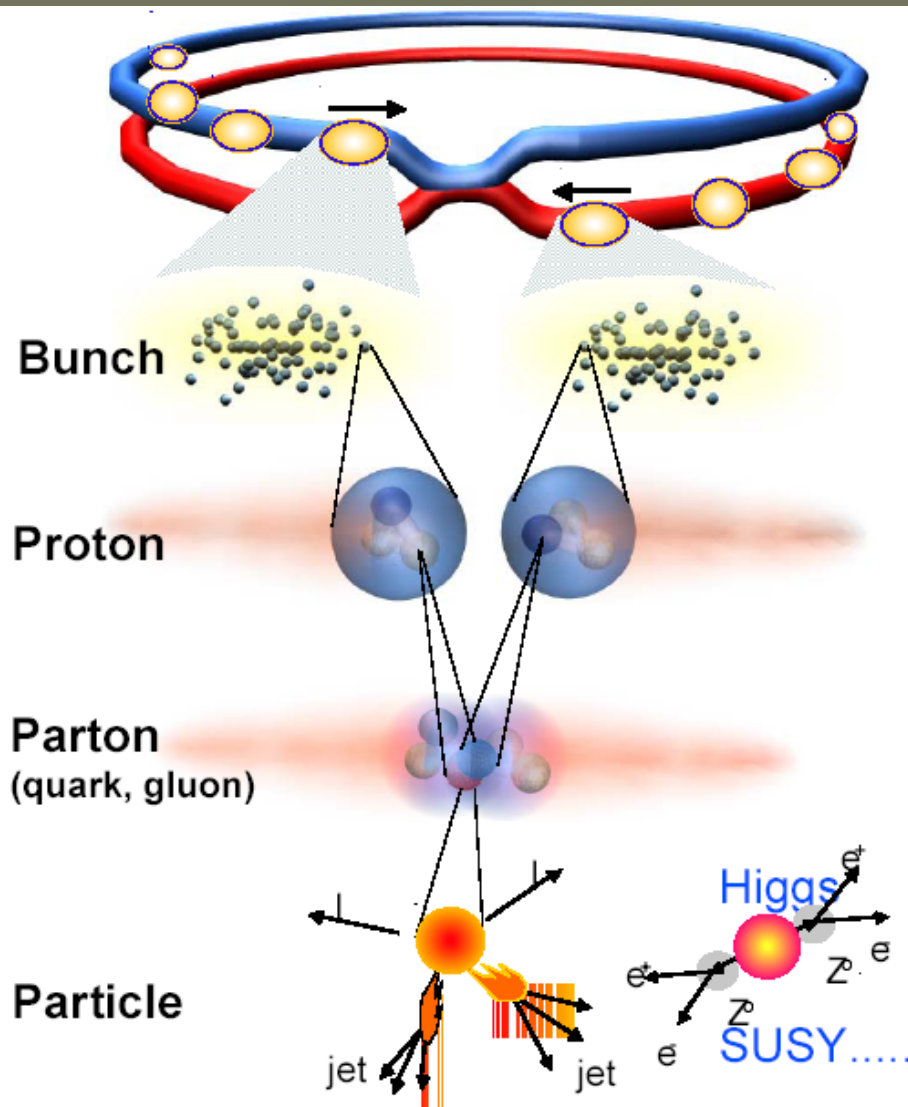
Proton - Proton	**2804 bunch/beam**
Protons/bunch	10^{11}
Beam energy	7 TeV (7×10^{12} eV)
Luminosity	10^{34}cm^{-2}s^{-1}
Crossing rate	40 MHz
Collision rate $\approx$	10^7-10^9

New physics rate $\approx$ .00001 Hz

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

Cross section and minimum bias: in-time pile-up

- How many interactions per beam crossing ?
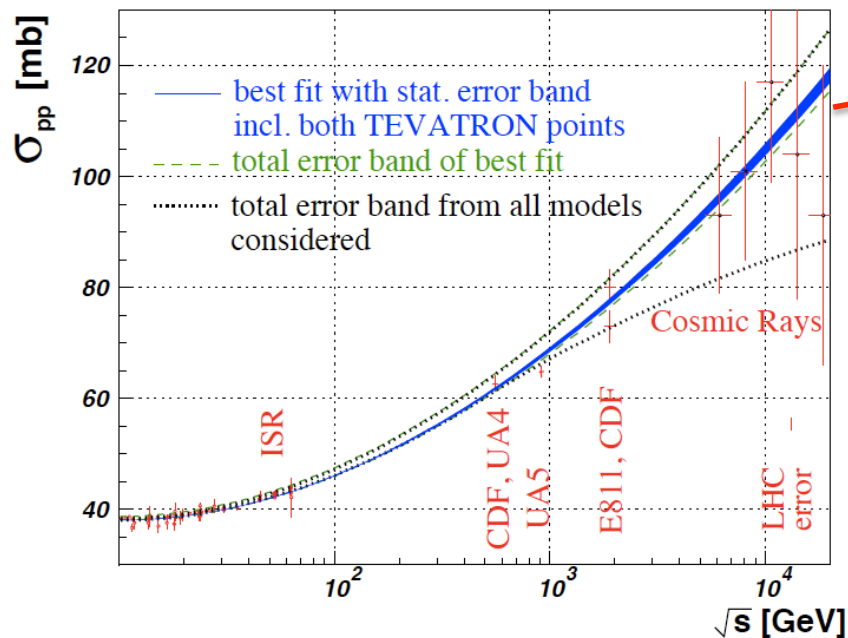
- Therefore a Higgs boson process (pp->HX) comes together with ~36 minimum bias events

- Luminosity $= 10^{34}\ cm^{-2}\ s^{-1}$
  $= 10^{7}\ mb^{-1}\ Hz$

  $\sigma(pp) \sim 115\ mb$
  $\Rightarrow rate = L * \sigma = 1.15\ 10^{9}\ Hz$

- time for one beam crossing
  $\Delta t = 25\ ns = 2.5\ 10^{-8}\ s$
  $\Rightarrow interaction/crossing = rate \cdot \Delta t$
  $= 28,75$

- not all bunches are full
  (only 2835 out of 3564)
  $\Rightarrow interactions/"active"\ crossing$
  $= 28,75 \left(\frac{2835}{3564}\right)^{-1} = \underline{36}$

Time-of-flight of signal: out-of-time pile-up

c=30cm/ns; in 25ns, s=7.5m

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
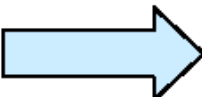Master Level Course 2007-2008

Pag. 43
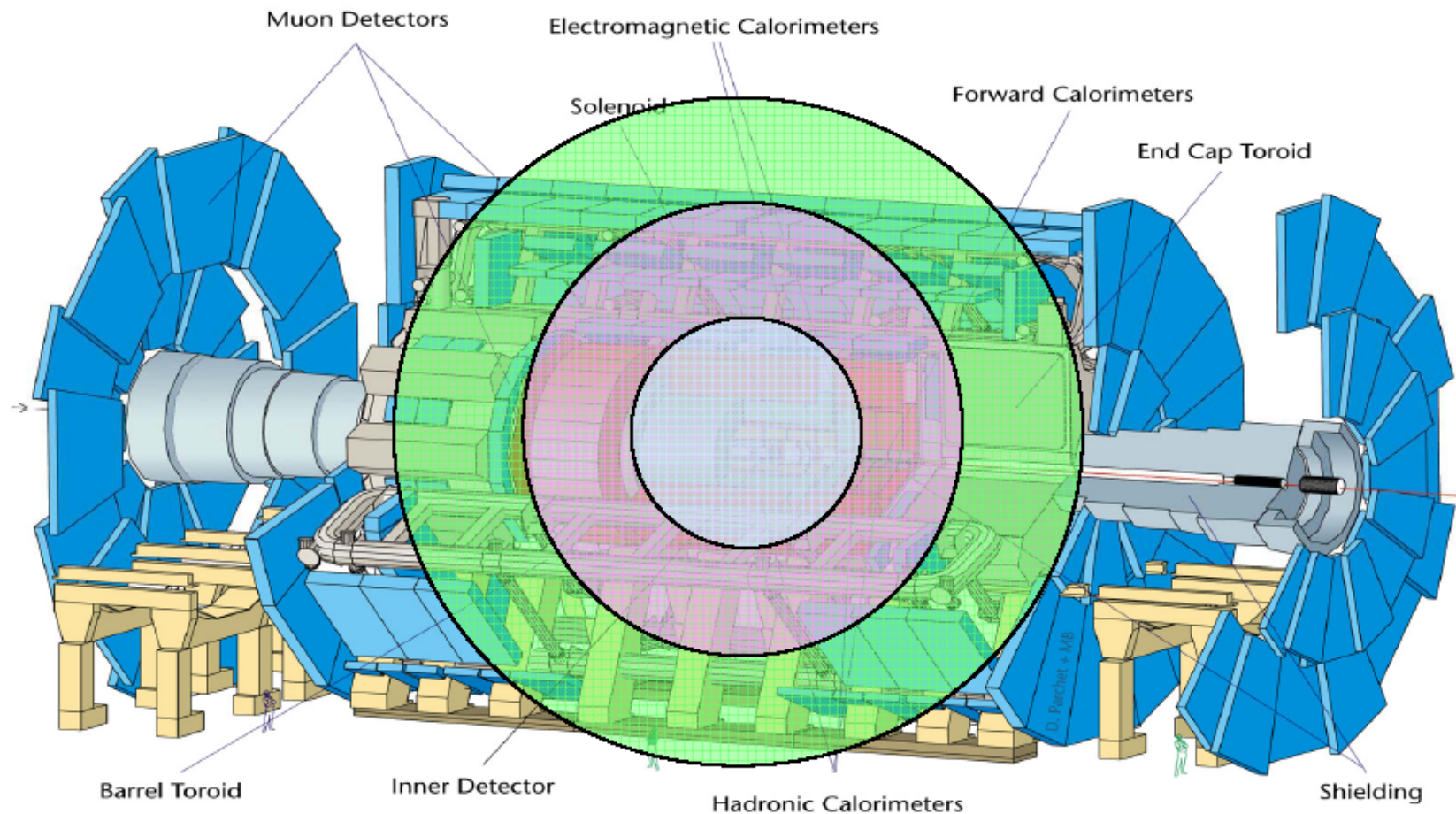
Superimposed pile-up collisions

- **"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"**
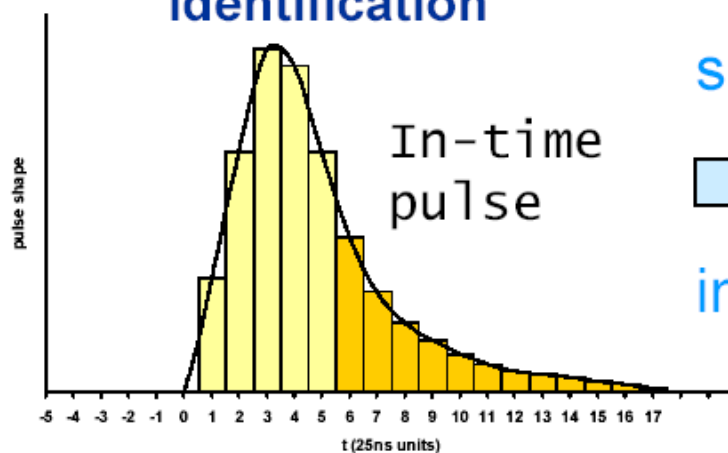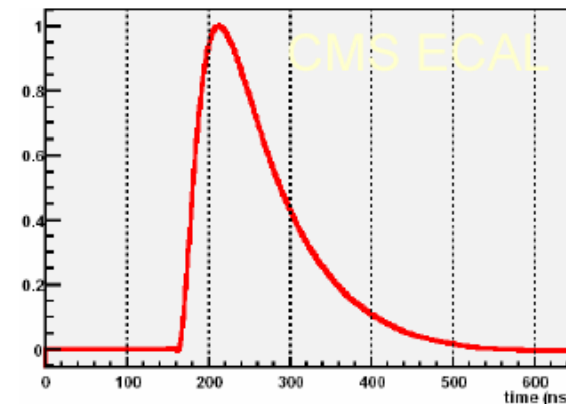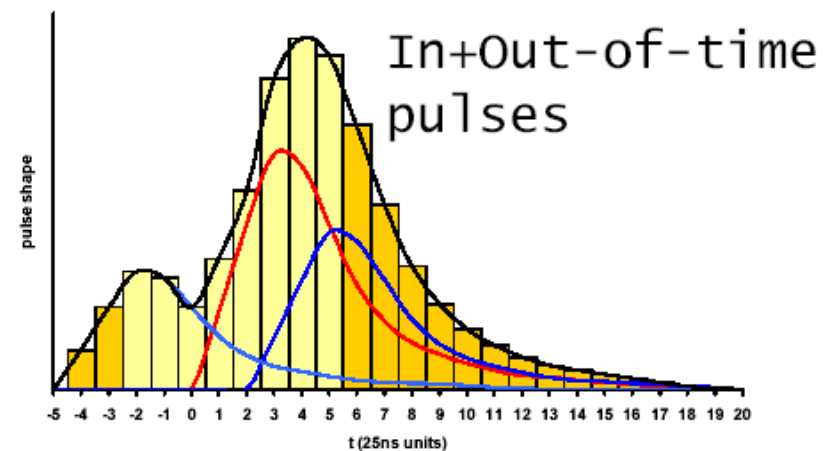
Impact on detector design

- **LHC detectors must have fast response**
  - ◆ Avoid integrating over many bunch crossings ("pile-up")
  - ◆ Typical response time : 20-50 ns
    - → integrate over 1-2 bunch crossings → pile-up of 25-50 min-bias events → very challenging readout electronics

- **LHC detectors must be highly granular**
  - ◆ Minimize probability that pile-up particles be in the same detector element as interesting object (e.g. γ from $H \rightarrow \gamma\gamma$ decays)
    - → large number of electronic channels

- **LHC detectors must be radiation resistant:**
  - ◆ high flux of particles from pp collisions → high radiation environment e.g. in forward calorimeters:
    - • up to 10^{17} n/cm^2 in 10 years of LHC operation
    - • up to 10^7 Gy (1 Gy = unit of absorbed energy = 1 Joule/Kg)

Data-analysis at colliders - Prof. J. D'Hondt (VUB)

Master Level Course 2007-2008

Pag. 45

The Compact Muon Solenoid

One of the main particle detectors which will collect the outcome of the proton collisions at the LHC is the CMS detector.
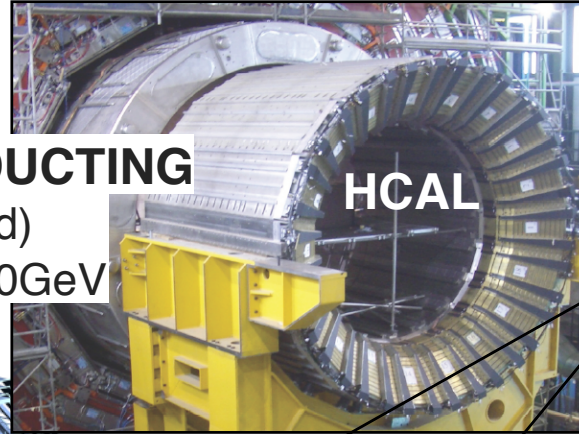
A multi-purpose detector with millions of read-out channels and electronics at sampling frequencies of 40MHz similar to the collision rate.
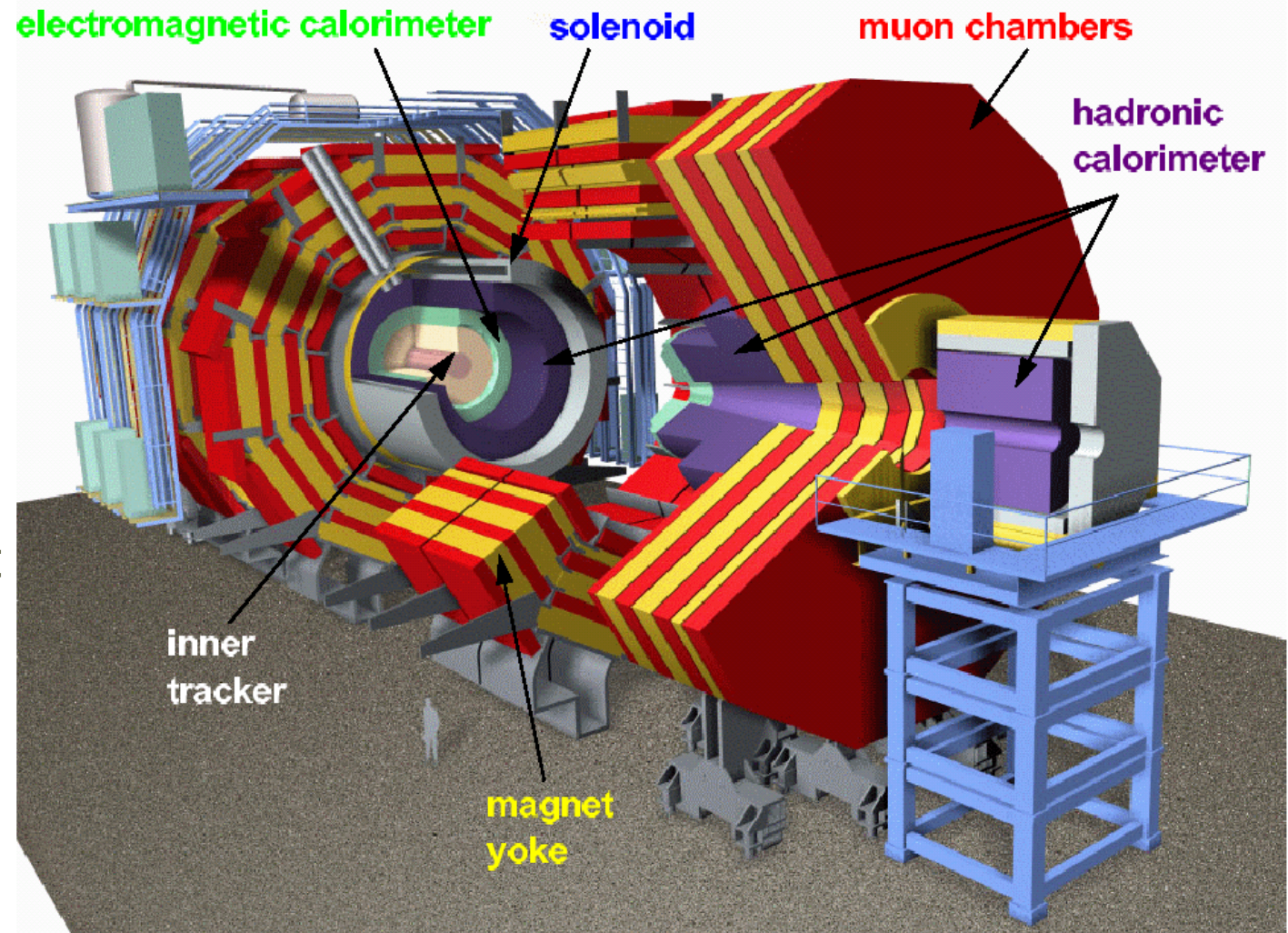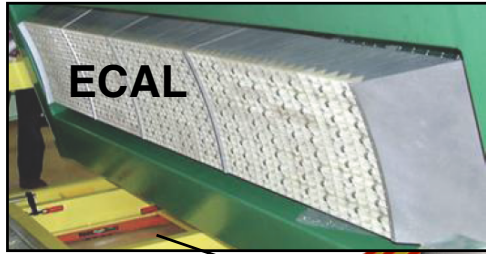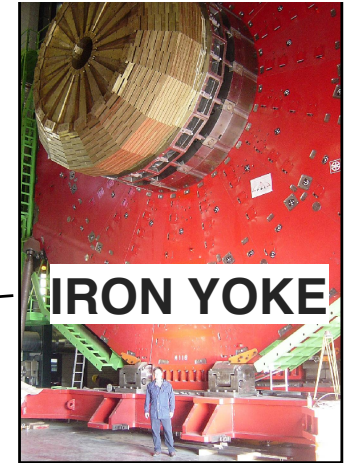
ECAL

Scintillating
PbWO4 crystals
~76k crystals

SUPERCONDUCTING COIL (4T B field)
$\Delta p_T/p_T \sim 1\%@100\text{GeV}$
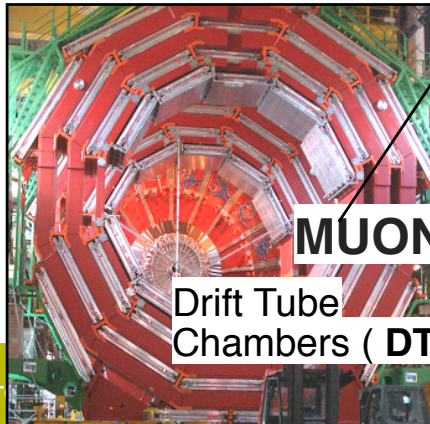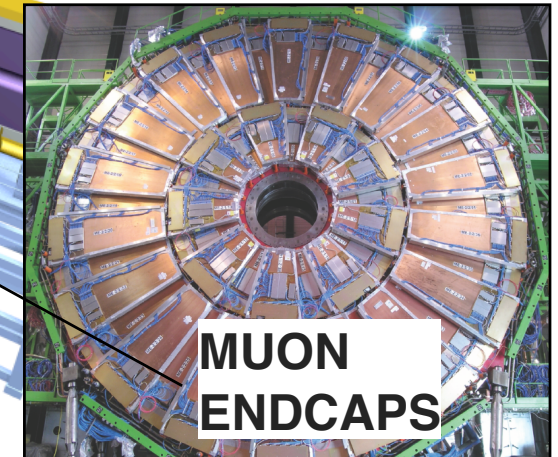
HCAL

Plastic scintillator/brass sandwich

IRON YOKE

TRACKER

Silicon Microstrips
Pixels
~220m² of Si

MUON BARREL

MUON ENDCAPS

Drift Tube Chambers (**DT**)

Resistive Plate Chambers (**RPC**)
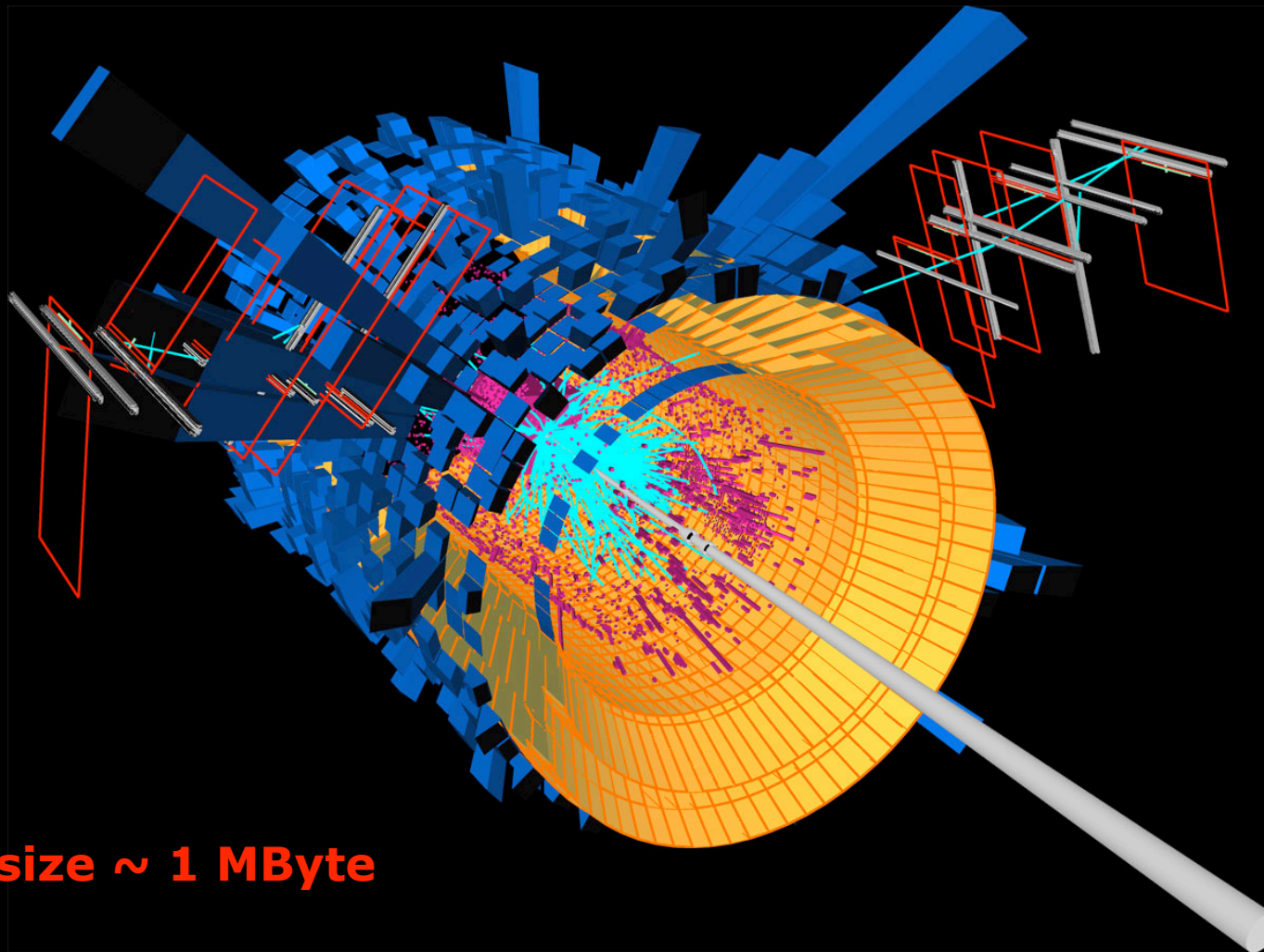
Cathode Strip Chambers (**CSC**)

Resistive Plate Chambers (**RPC**)

Enormous amount of data sources

sub-detector	number of channels	number of FE chips	number of detector data links	number of data sources (FEDs)	number of DAQ links (FRLs)
Tracker pixel	$\approx$ 66 M	15840	$\approx$ 1500	40	40
Tracker strips	$\approx$ 9.3 M	$\approx$ 72 k	$\approx$ 36 k	440	250 (merged)
Preshower	144384	4512	1128	56	56
ECAL	75848	$\approx$ 21 k	$\approx$ 9 k	54	54
HCAL	9072	9072	3072	32	32
Muons CSC	$\approx$ 500 k	$\approx$ 76 k	540	8	8
Muons RPC	192 k	$\approx$ 8.6 k	732	3	3
Muons DT	195 k	48820	60	10	10
Global Trigger	n/a	n/a	n/a	3	3
CSC, DT Track Finder	n/a	n/a	n/a	2	2
Total	$\approx$ 77 M			626	458

Each data source will deliver on average ~2kByte per pp collision.

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
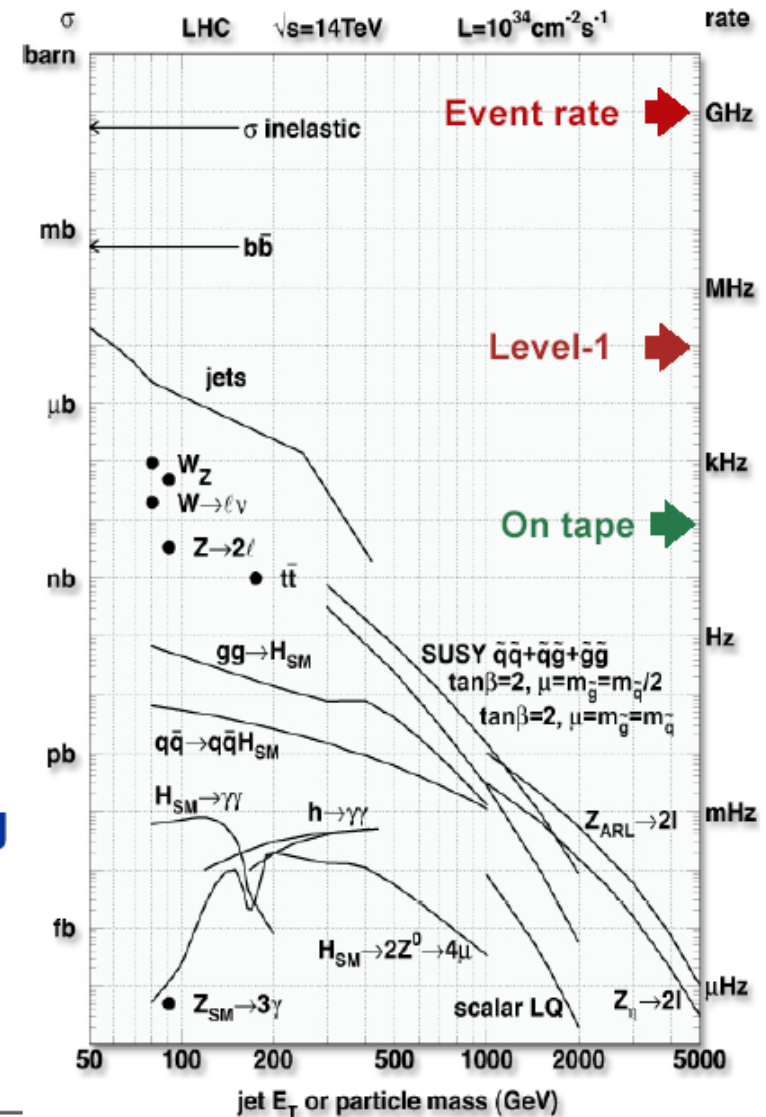Master Level Course 2007-2008

Pag. 48
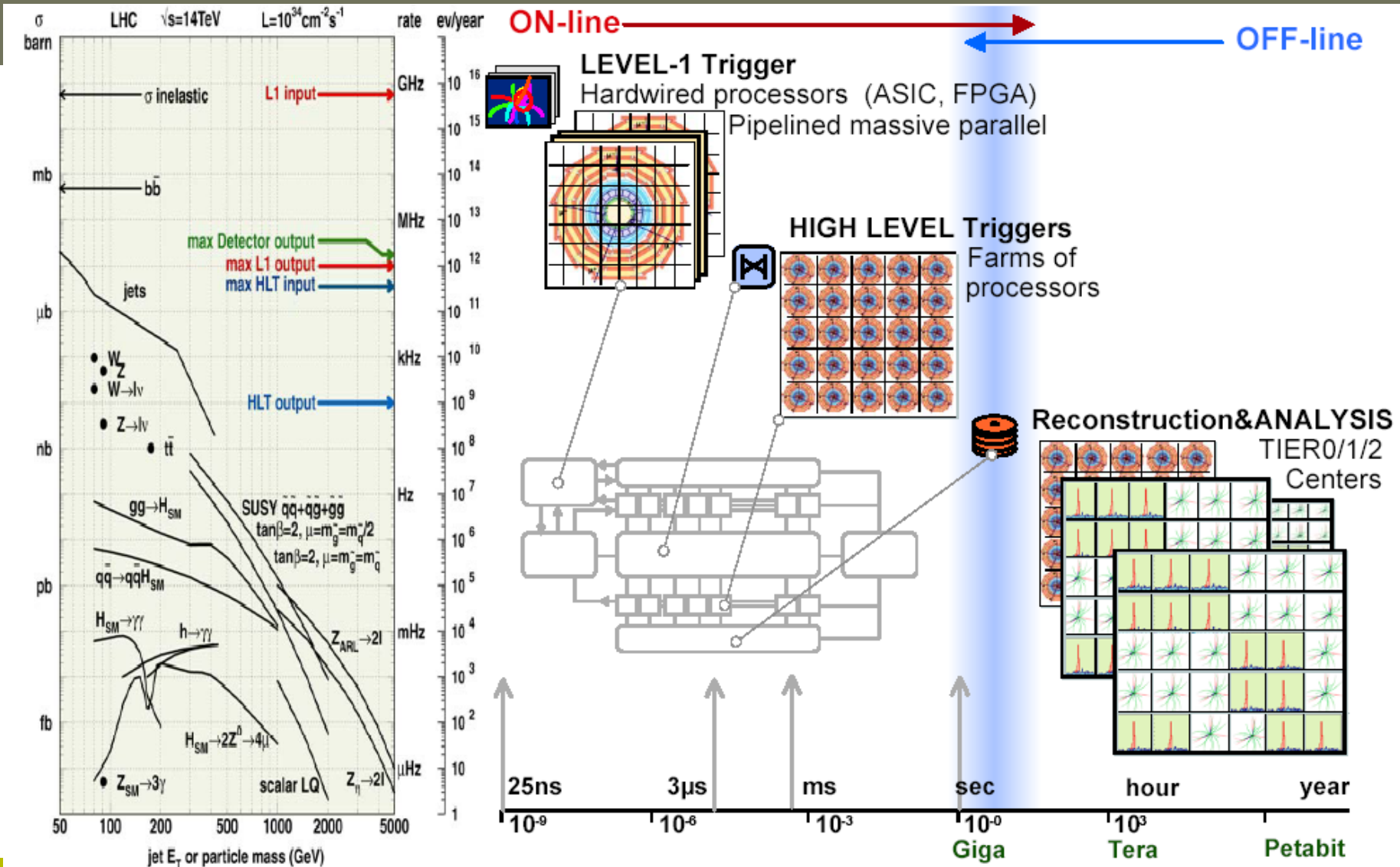
Example of a final state event

Event size ~ 1 MByte

Selectivity of the physics signals

- **Cross sections of physics processes vary over many orders of magnitude**
  - ◆ Inelastic: 10^9 Hz
  - ◆ $W \to \ell\, \nu$: 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 $\to$ electron bkg
  - ◆ Decays of K, π, b $\to$ muon bkg
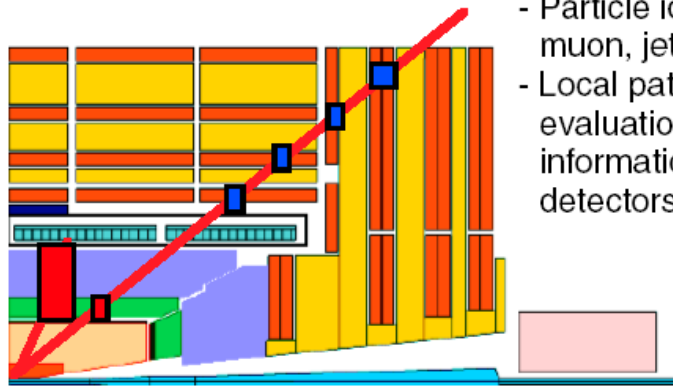- **Selection needed: $1{:}10^{10-11}$**
  - ◆ Before branching fractions...

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 50

Physics selection at the LHC

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
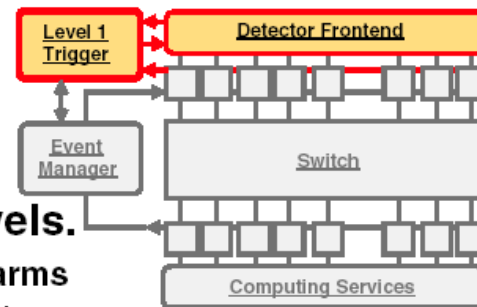
Master Level Course 2007-2008

Pag. 51

Physics selection at the LHC

40 MHz

Level-1. **Specialized processors**
- Particle identification: high p_T electron, muon, jets, missing E_T
- Local pattern recognition and energy evaluation on prompt macro-granular information from calorimeter and muon detectors
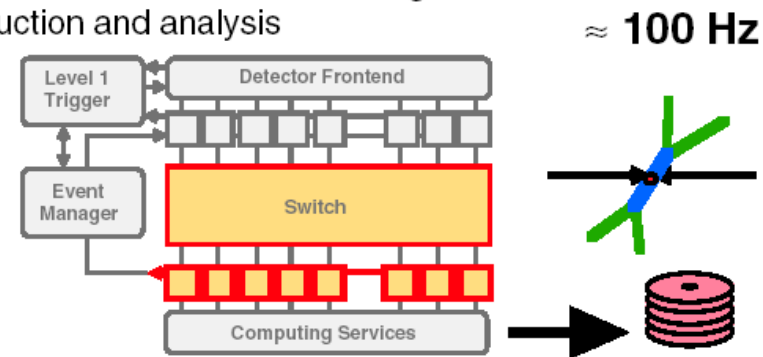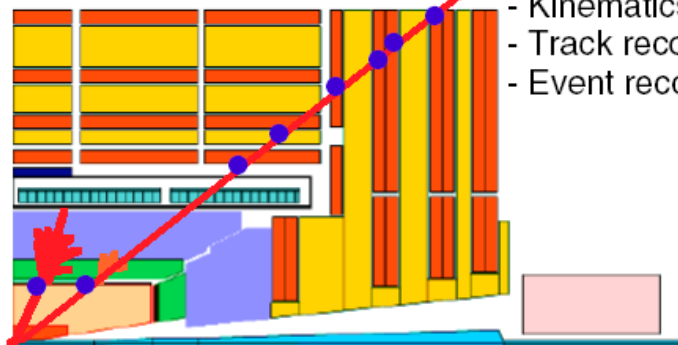
Up to 100 kHz

High trigger levels.

Network and CPU farms
- Clean particle signature
- Finer granularity precise measurement
- Kinematics. effective mass cuts & event topology
- Track reconstruction and detector matching
- Event reconstruction and analysis

$\approx$ **100 Hz**

Level 1 Trigger
Detector Frontend
Event Manager
Switch
Computing Services

Reference for Level-1 in CMS

Public manuscript written in 2000 and contains lots of information on the Level-1 Trigger of the CMS experiment. Almost 600 pages. Can be found at

http://cms.cern.ch/iCMS/jsp/page.jsp?mode=cms&action=url&urlkey=CMS_TDRS

LABORATOIRE EUROPEEN POUR LA PHYSICQUE DES PARTICULES

CERN EUROPEAN LABORATORY FOR PARTICLE PHYSICS

CERN/LHCC 00-xx
CMS TDR 6.1
November 2000

C M S

The TriDAS project. Volume I
The Trigger Systems

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 53

Reference for HLT in CMS

Public manuscript written in 2002 and contains lots of information on the High Level Trigger of the CMS experiment. Almost 500 pages. Can be found at

http://cms.cern.ch/iCMS/jsp/page.jsp?mode=cms&action=url&urlkey=CMS_TDRS
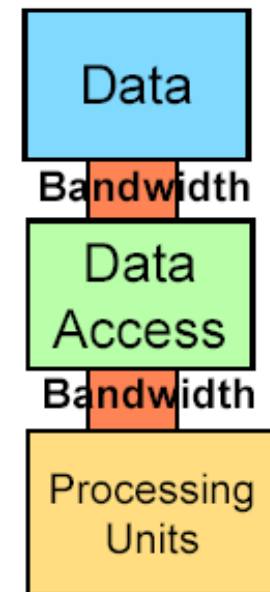
Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 54
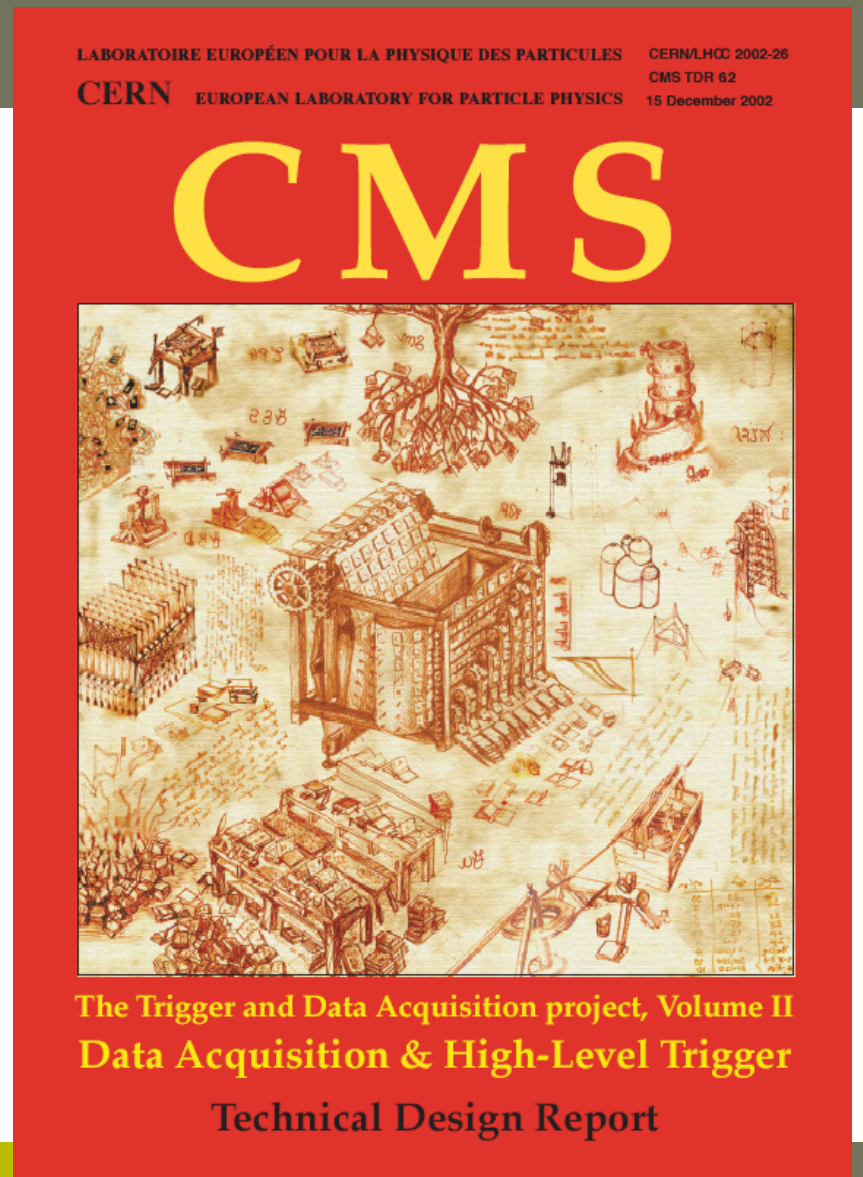
Trigger structures

First L1 reduction from 40MHz to 10^5Hz (cfr 10^2Hz final rate to tape/disk) is always there, followed by 1 or 2 extra steps.
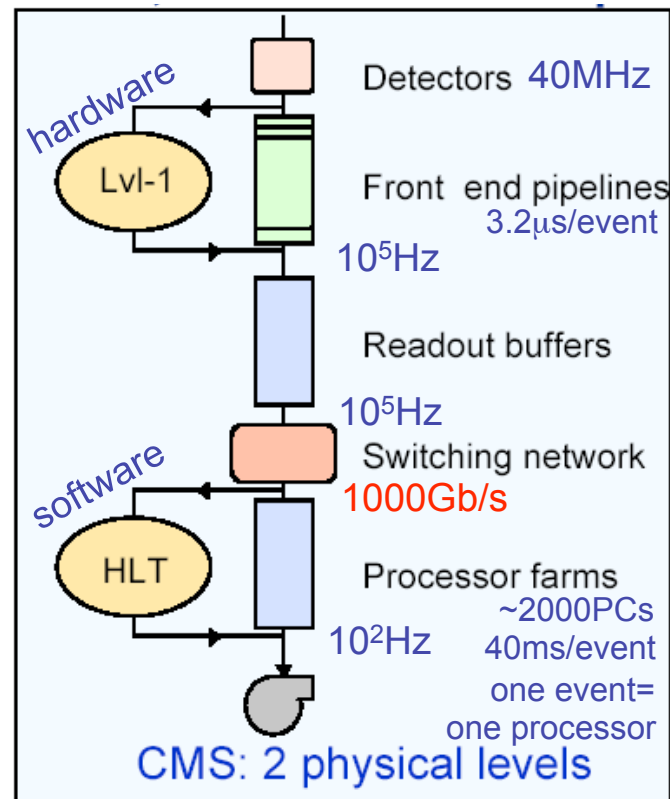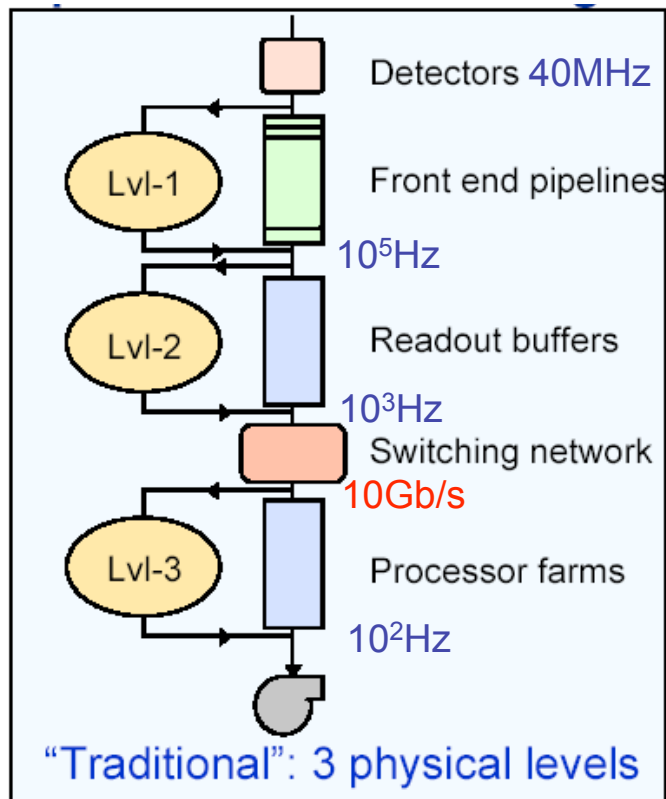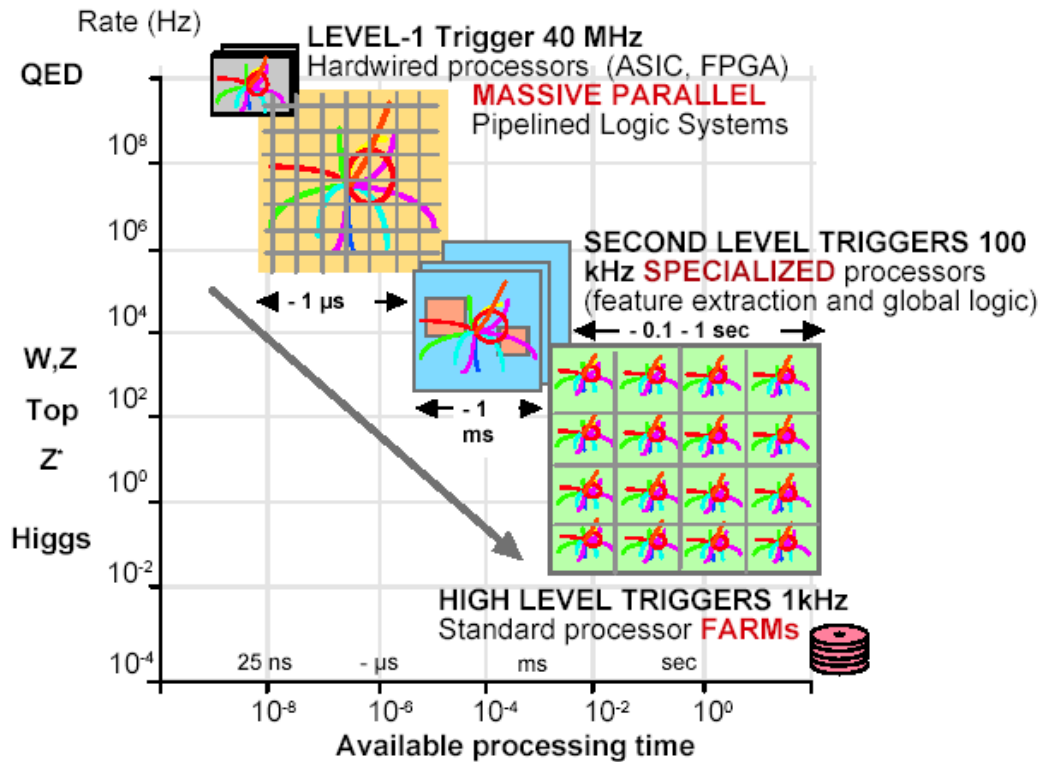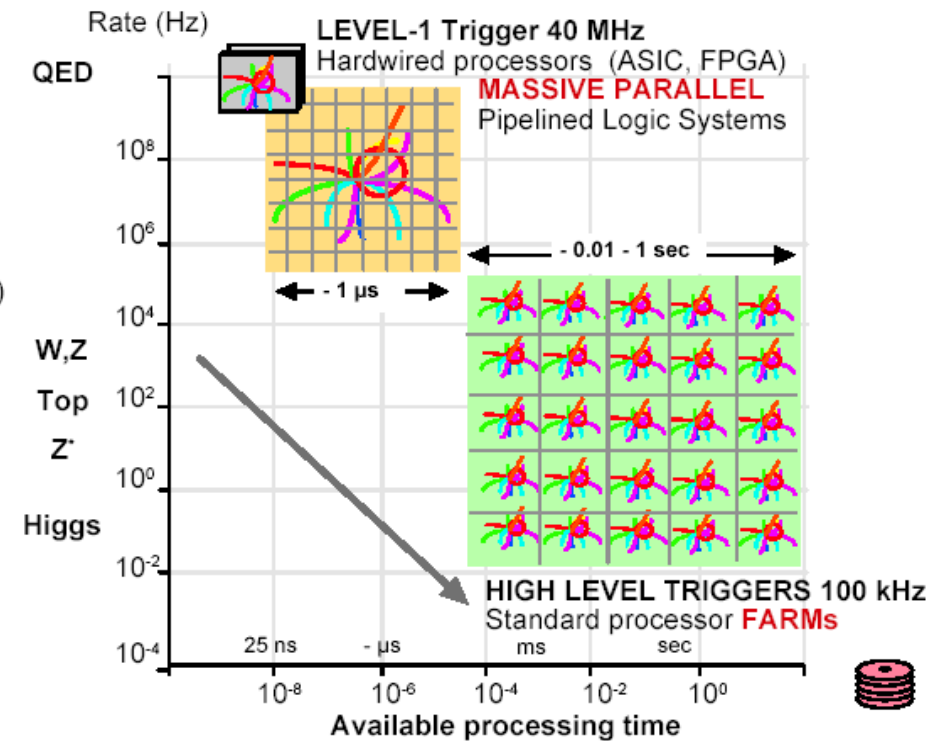
Trigger structures

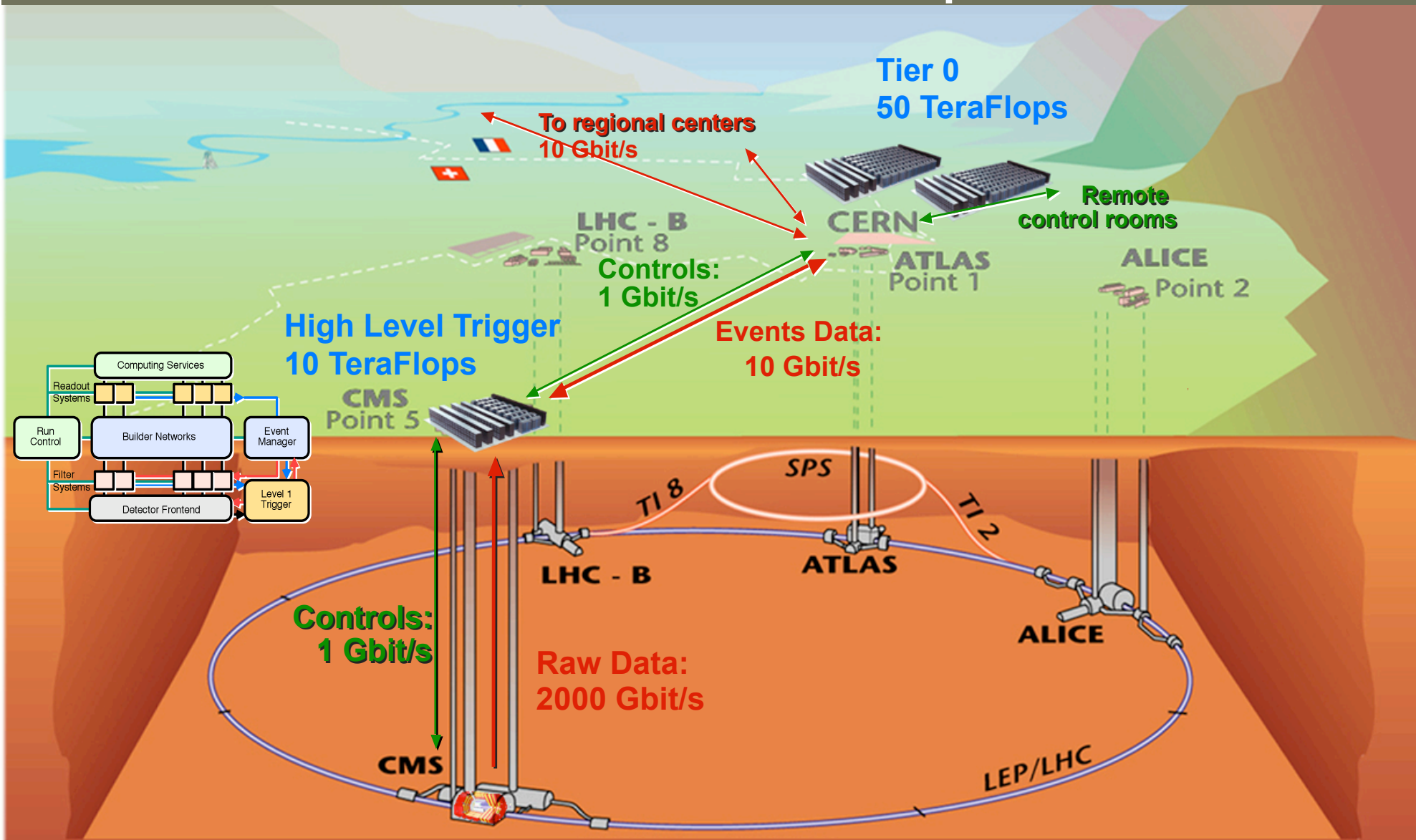**3-step approach**
- Control logistics
- Specialized processors

2-step approach
- Bandwidth important
- Commercial processors

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 56

Data flow in the CMS experiment

Tier 0
50 TeraFlops

To regional centers
10 Gbit/s

Remote
control rooms

LHC - B
Point 8

Controls:
1 Gbit/s

CERN

ATLAS
Point 1

ALICE
Point 2

High Level Trigger
10 TeraFlops

Events Data:
10 Gbit/s

Computing Services

Readout
Systems

Run
Control

Builder Networks

Event
Manager

Filter
Systems

Detector Frontend

Level 1
Trigger

CMS
Point 5

SPS

TI 8

TI 2

LHC - B

ATLAS

ALICE

Controls:
1 Gbit/s

Raw Data:
2000 Gbit/s

CMS

LEP/LHC

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 57

The Level-1 trigger

The Level-1 trigger

- The Level-1 trigger algorithms have to analyze every bunch crossing

- To get a quasi-deadtime-free operation, pipelines are installed at the detectors to cover the 3.2 μs time needed for the trigger decision

- The decision is based on algorithms which are calculated by specific processors (=hardware)

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 59

Level-1 trigger algorithms

- **Physics facts:**
  - ◆ pp collisions produce mainly hadrons with $P_T \sim 1$ GeV
  - ◆ Interesting physics (old and new) has particles (leptons and hadrons) with large transverse momenta:
    - • $W \rightarrow e\nu$: $M(W) = 80$ GeV/c^2; $P_T(e) \sim 30\text{-}40$ GeV
    - • $H(120$ GeV$) \rightarrow \gamma\gamma$: $P_T(\gamma) \sim 50\text{-}60$ GeV

- **Basic requirements:**
  - ◆ Impose high thresholds on particles
    - • Implies distinguishing particle types; possible for electrons, muons and "jets"; beyond that, need complex algorithms
  - ◆ Typical thresholds:
    - • Single muon with $P_T > 20$ GeV (rate ~ 10 kHz)
      - → Dimuons with $P_T > 6$ (rate ~ 1 kHz)
    - • Single e/γ with $P_T > 30$ GeV (rate $\sim 10\text{-}20$ kHz)
      - → Dielectrons with $P_T > 20$ GeV (rate ~ 5 kHz)
    - • Single jet with $P_T > 300$ GeV (rate $\sim 0.2\text{-}0.4$ kHz)

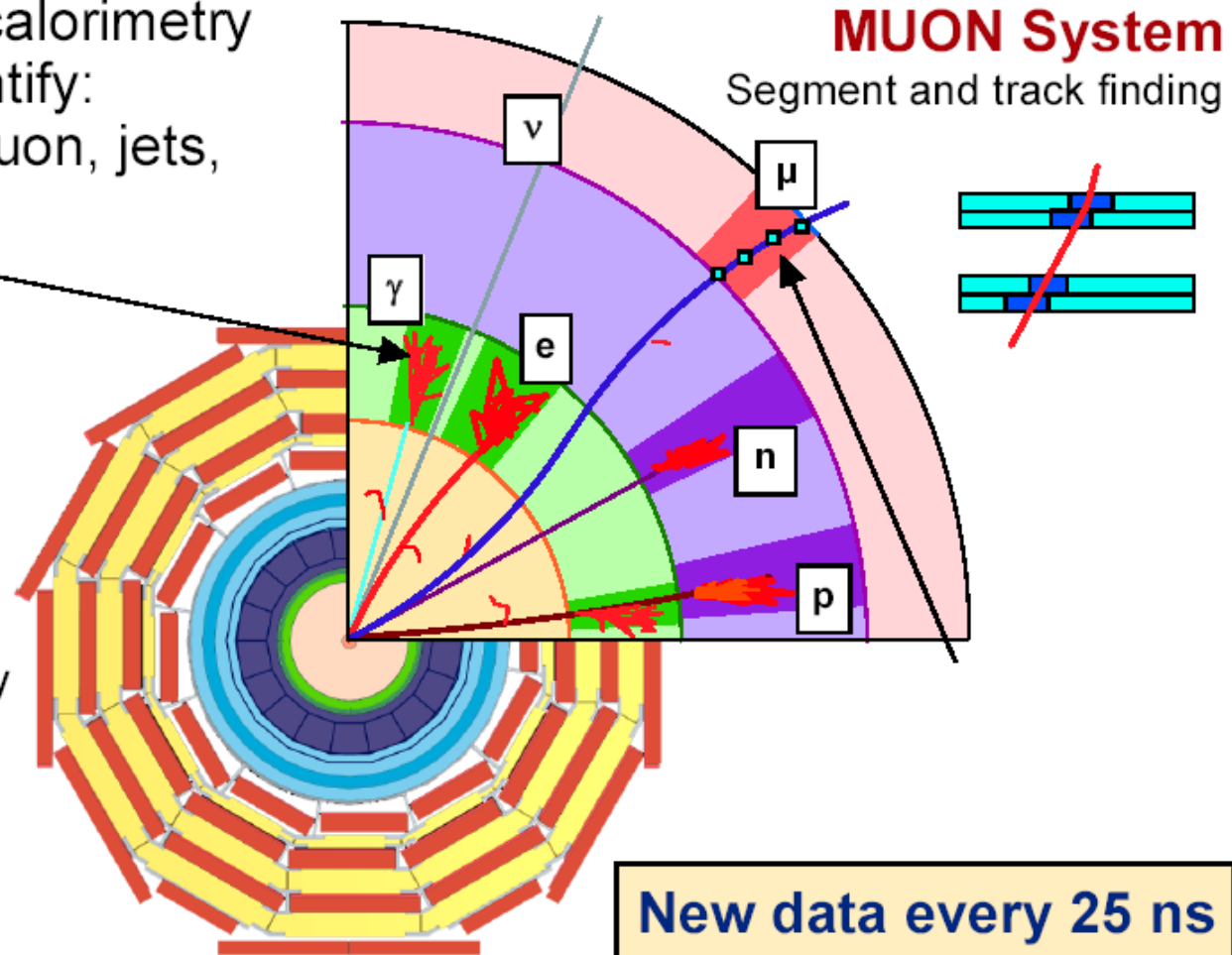Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 60

Signatures in the detector(s)

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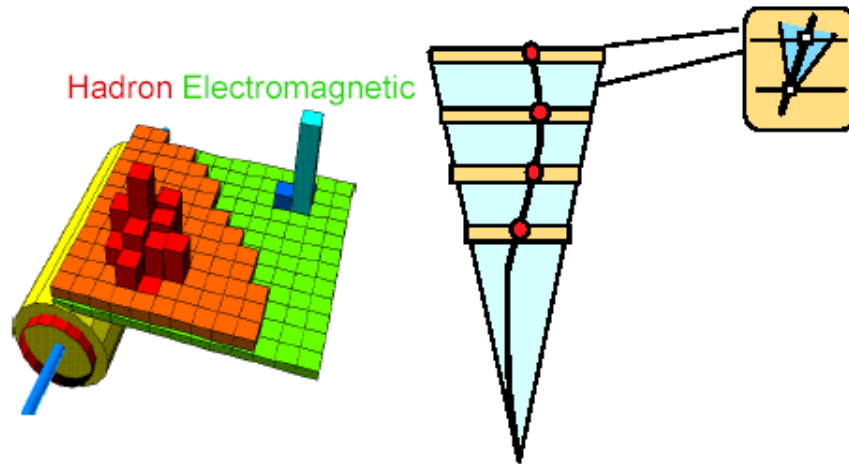High p_t electron, muon, jets, missing E_T

MUON System
Segment and track finding
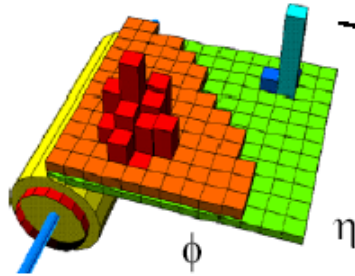
CALORIMETERs
Cluster finding and energy deposition evaluation

New data every 25 ns
Decision latency ~ µs

Signatures in the detector(s)

Pattern recognition much faster/easier

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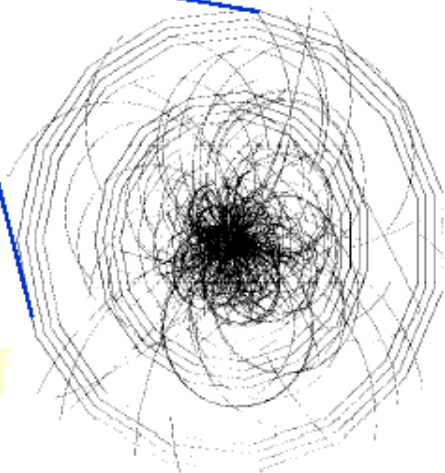
- **Simple algorithms**
- **Small amounts of data**
- **Local decisions**

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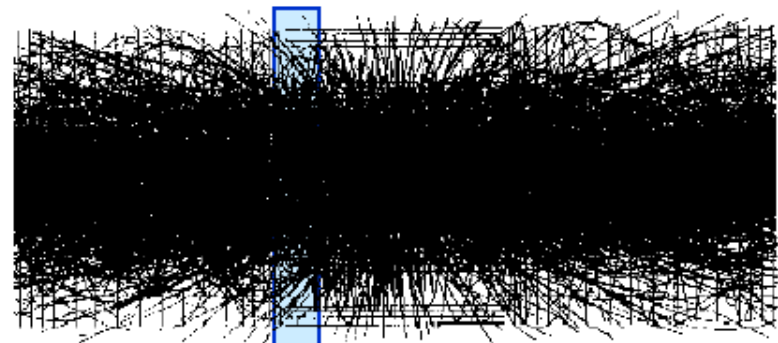
- Complex algorithms
- Huge amounts of data
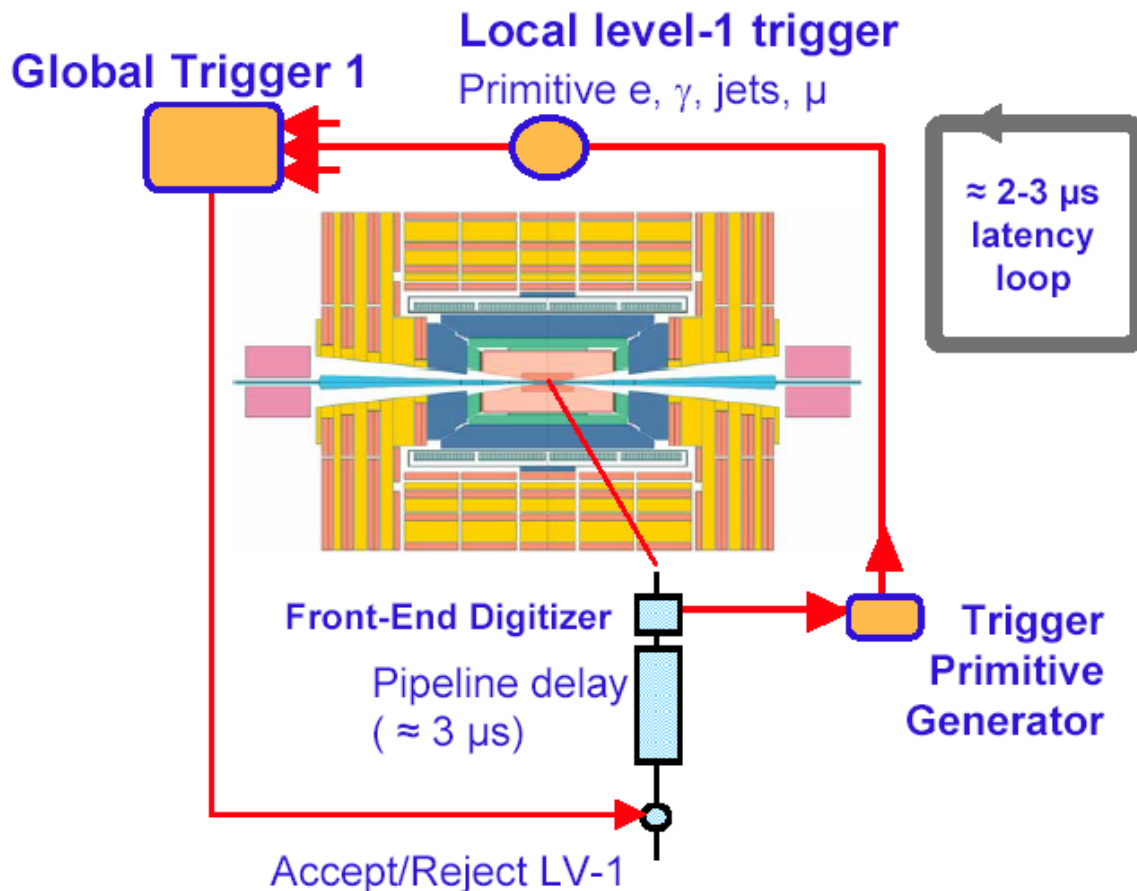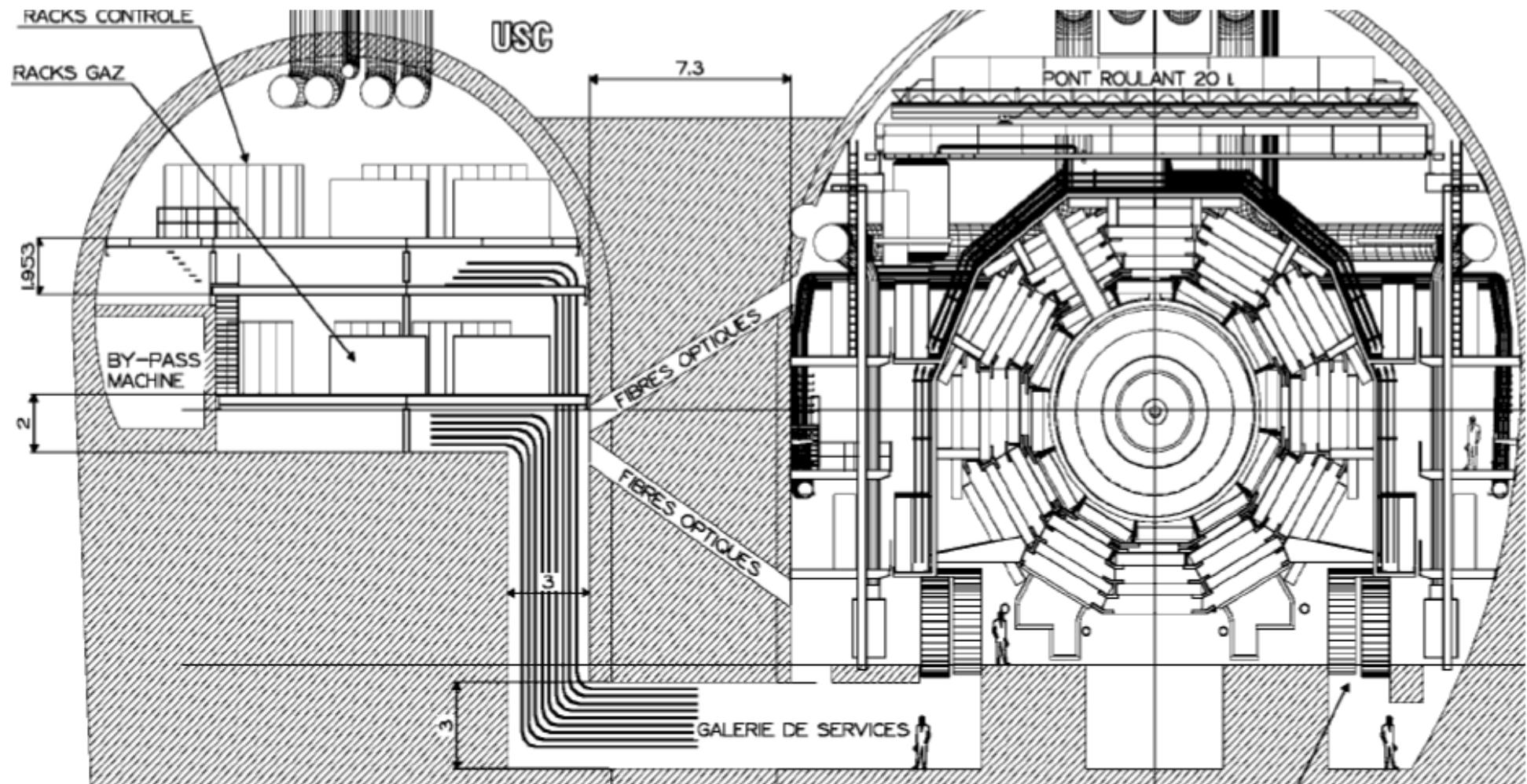- Need to link sub-detectors

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 62

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 1

Local level-1 trigger
Primitive e, γ, jets, μ

≈ 2-3 μs latency loop

Front-End Digitizer

Pipeline delay (≈ 3 μs)

Trigger Primitive Generator

Accept/Reject LV-1

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 63

Level-1 is hardware

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 64

The path of the signal...
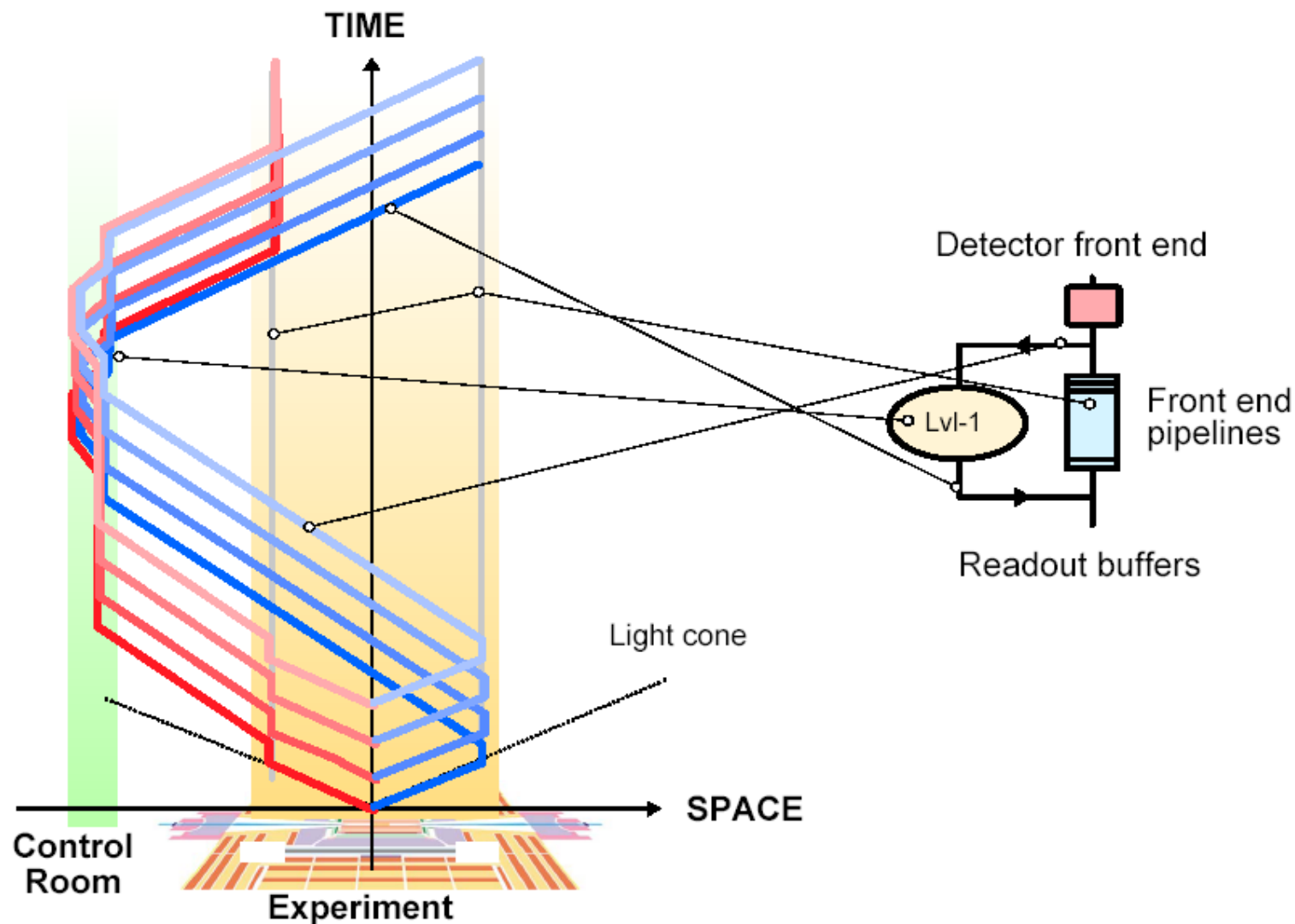
Detector readout: front-end

DIGITAL Asynchronous

LVL 1   Bunch#   Discr.

Bunch#

ANALOG pipeline

40 MHz → Shaper ASP

Pipeline

MUX

ASP

ADC

DSP

DIGITAL Synchronous

40 MHz ← Shaper

ADC

DSP

Pipeline
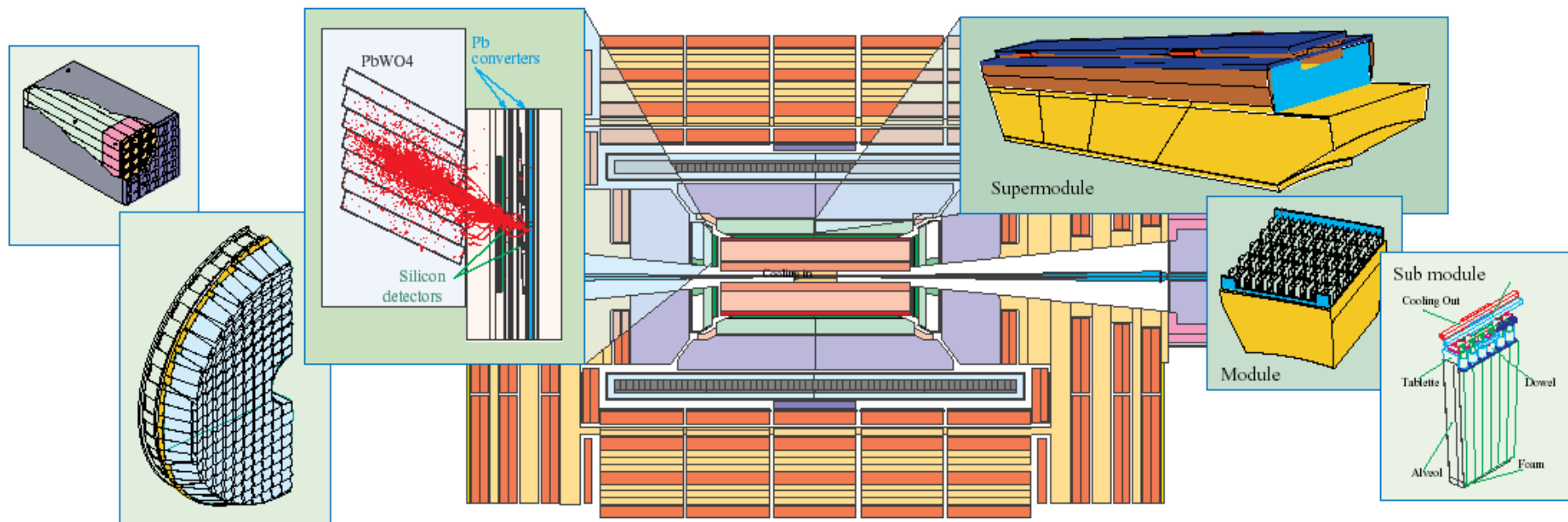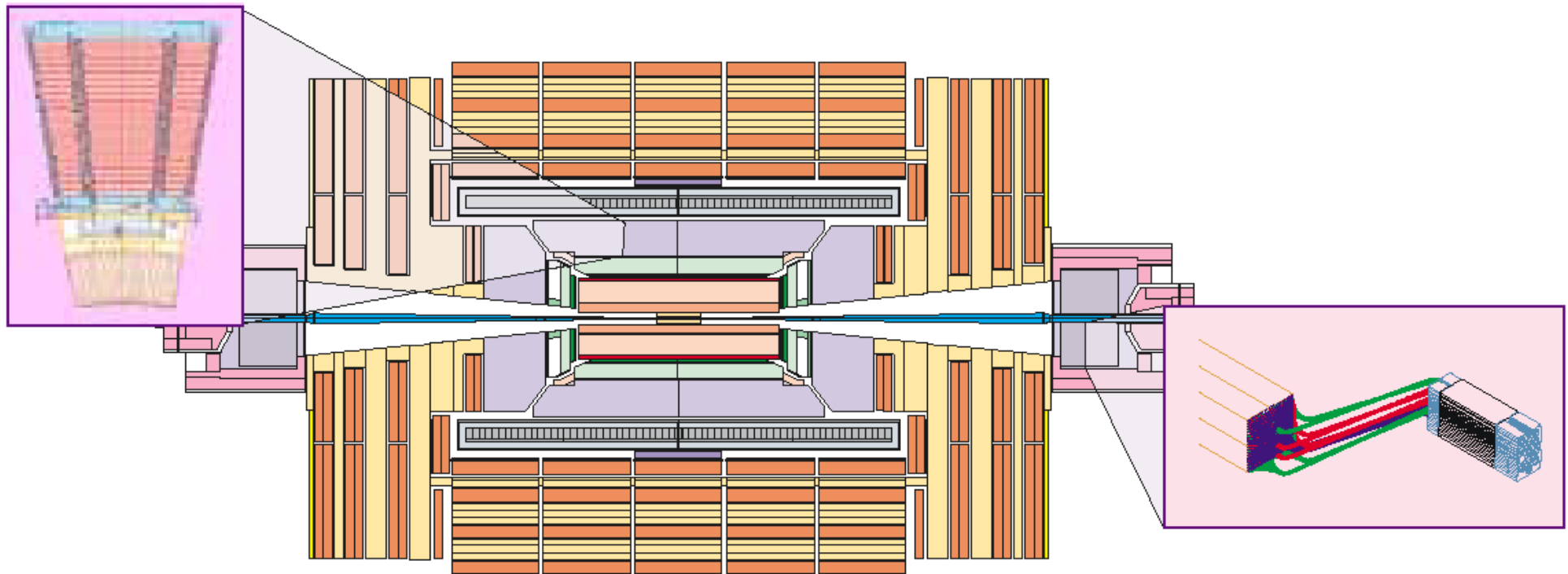
Readout
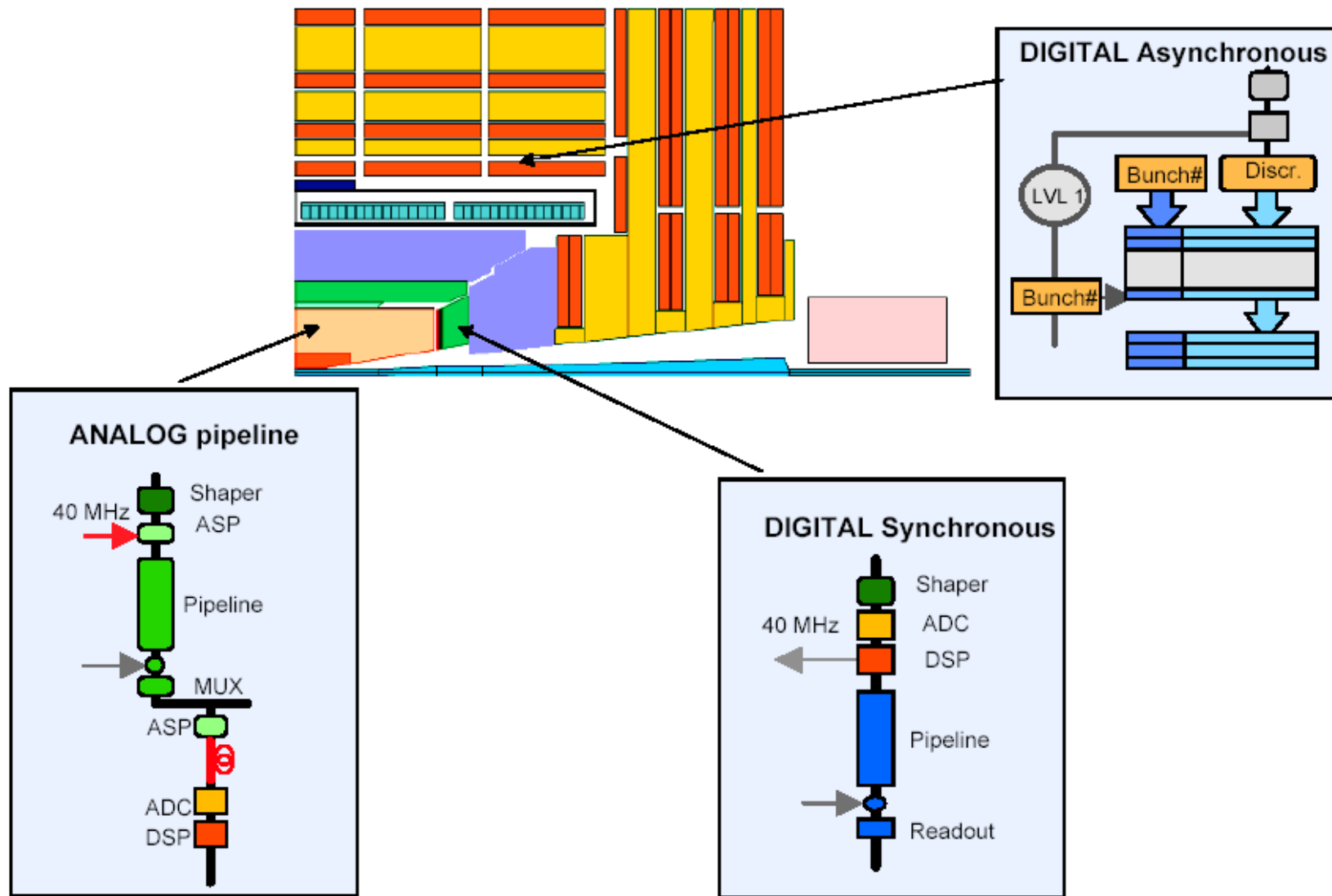
Calorimeter system in CMS

The electro-magnetic calorimeter (ECAL) will help in the reconstruction of electrons and photons. It consist out of ~77k lead-tungstate crystals equipped with avalanche photodiodes in the barrel or vacuum phototriodes in the endcap.

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
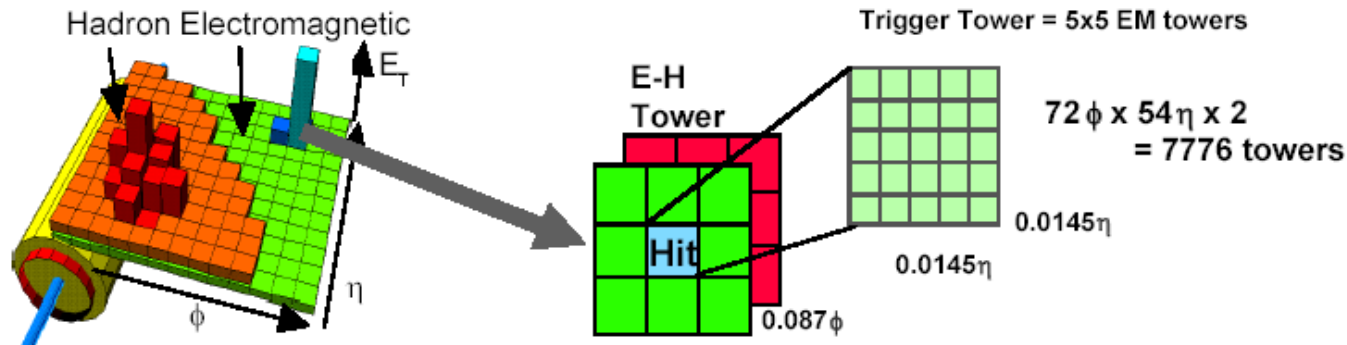Master Level Course 2007-2008

Pag. 67

Calorimeter system in CMS

The hadron calorimeter (HCAL) is important for the detection of hadronic jets from quarks, gluons or new particles, and also for the measurement of the missing energy resulting from a neutrino or a new very weakly interacting particle. These are sampling calorimeters of copper absorber plates interleaved with thick scintillator sheets.

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
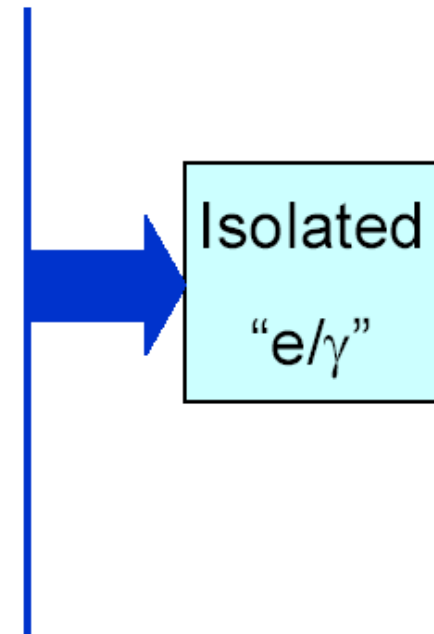Master Level Course 2007-2008

Pag. 68

Level-1 Calorimeter trigger

Level-1 Calorimeter trigger

- The efficiency of triggering an event with a reconstructed electron of a given transverse momentum is important (trigger object not necessary the offline object as then we have more time to reconstruct all pieces in the detector).

- To select the signal events (eg. Higgs bosons) among all proton collisions we apply cuts on the transverse momentum of the objects. High mass resonances produce on average objects like electrons with larger transverse momenta.

- But we need to make sure that if we need electrons above $p_T=25$GeV/c that we get them on disk after triggering.

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 70
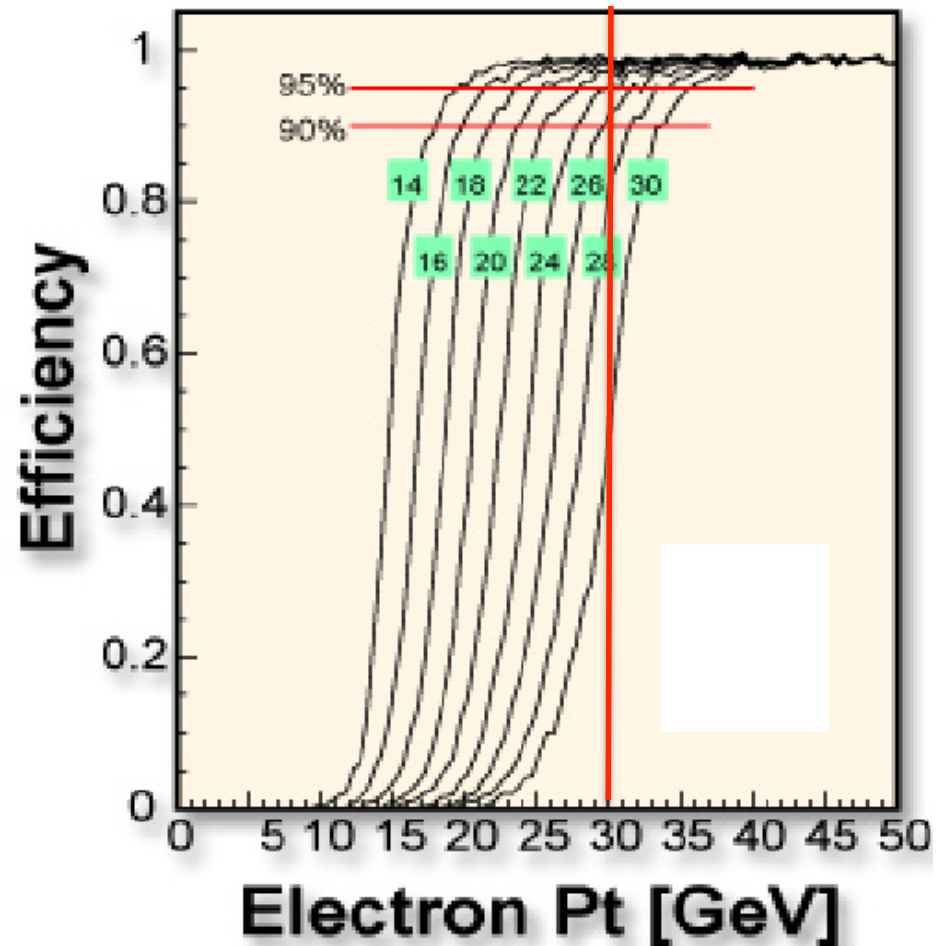
Level-1 Jet and τ trigger

- **Issues are jet energy resolution and tau identification**
  - ◆ Single, double, triple and quad thresholds possible
  - ◆ Possible also to cut on jet multiplicities
  - ◆ Also ETmiss, SET and SET(jets) triggers

Trigger Tower

"τ-like" shapes identified for τ trigger

HCAL

ECAL

PbWO4 Crystal

$\Delta\eta, \Delta\phi = 0.348$

$\Delta\eta, \Delta\phi = 1.04$
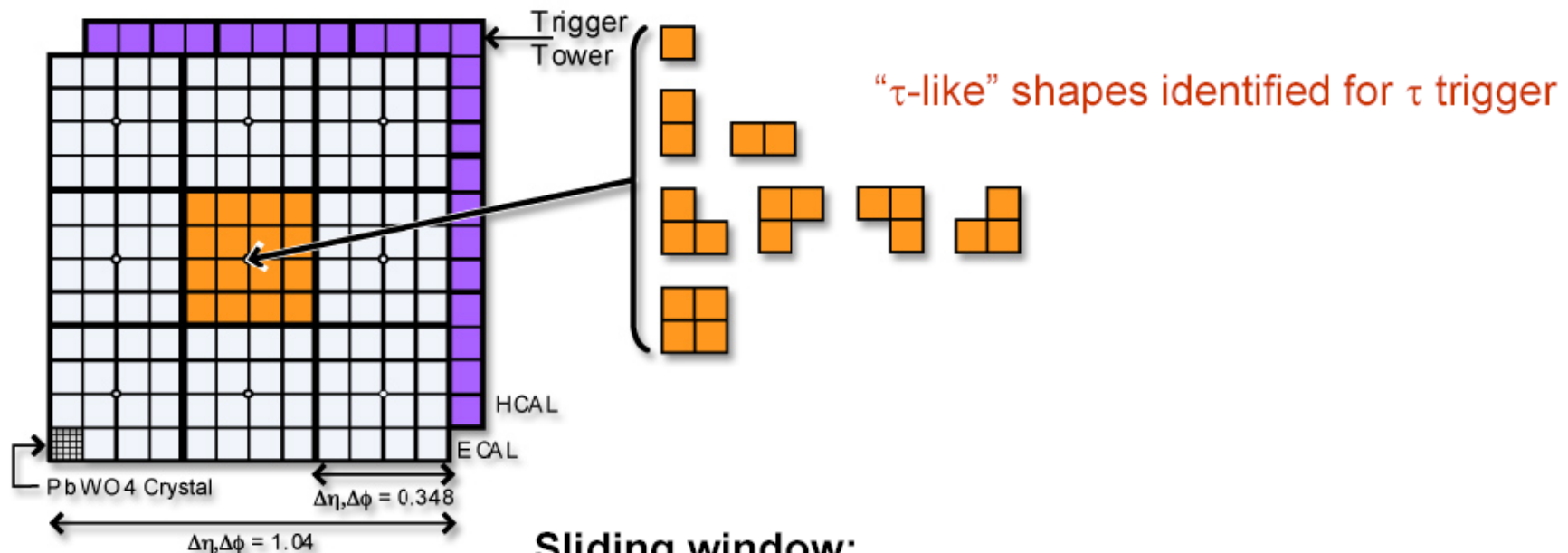
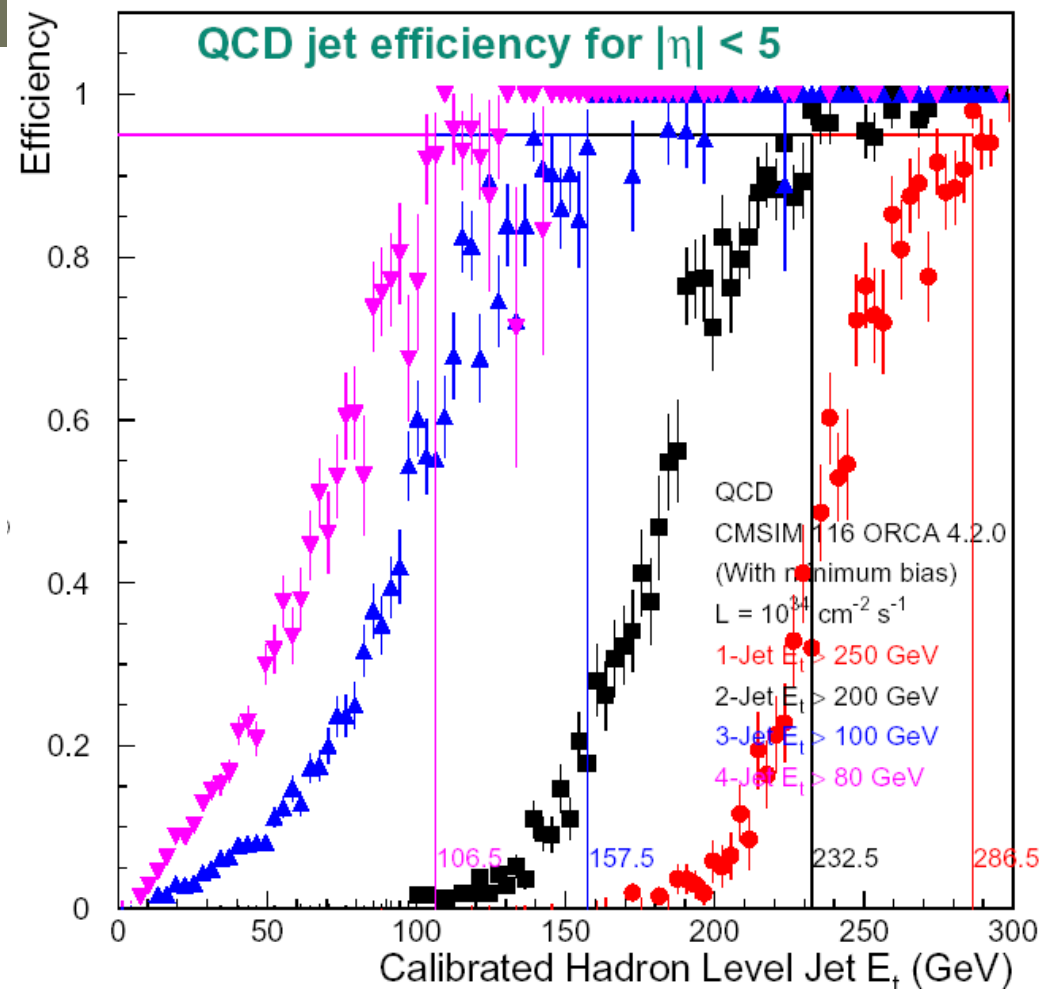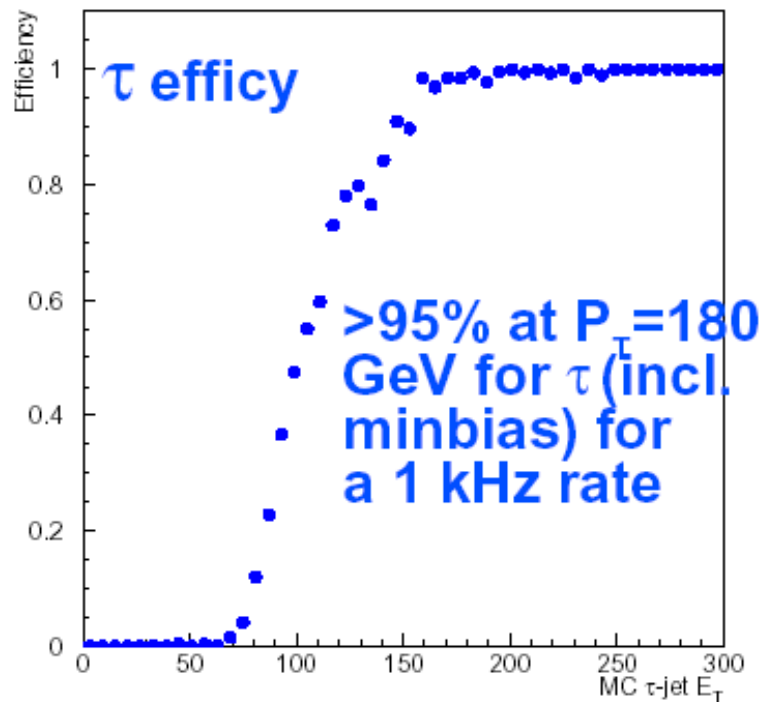**Sliding window:**
- granularity is 4x4 towers = trigger region
- jet E_T summed in 3x3 regions $\Delta\eta, \Delta\phi = 1.04$

Level-1 Jet and τ trigger

Similar efficiencies for τ and jets as for electrons. The resolution of these objects is however worse, hence the less sharp turn-on curve.

τ efficy

>95% at P_T=180 GeV for τ (incl. minbias) for a 1 kHz rate

QCD jet efficiency for $|\eta|$ < 5

QCD
CMSIM 116 ORCA 4.2.0
(With minimum bias)
$L = 10^{34}$ cm^{-2} s^{-1}
1-Jet E_t > 250 GeV
2-Jet E_t > 200 GeV
3-Jet E_t > 100 GeV
4-Jet E_t > 80 GeV

106.5   157.5   232.5   286.5

Calibrated Hadron Level Jet E_t (GeV)

>95% at P_T=286, 232, 157, 106 GeV for individual 1,2,3,4 jet triggers (incl. minbias) (~0.5 kHz rate each totalling ~2 kHz)

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 72

Muon system in CMS

- The muon system consists out of several complementary techniques for example to suppress the amount of "ghost" muons.
- Tracking algorithms are developed to construct the path of the muon candidate.
- The RPCs or Resistive Plate Chambers are dedicated to the trigger system due to their good time resolution to check the bunch crossing time. They are complemented with the high resolution provided by the DT and CSC.
- Hence we have two trigger subsystems which deliver independent information to the Global Muon Trigger.
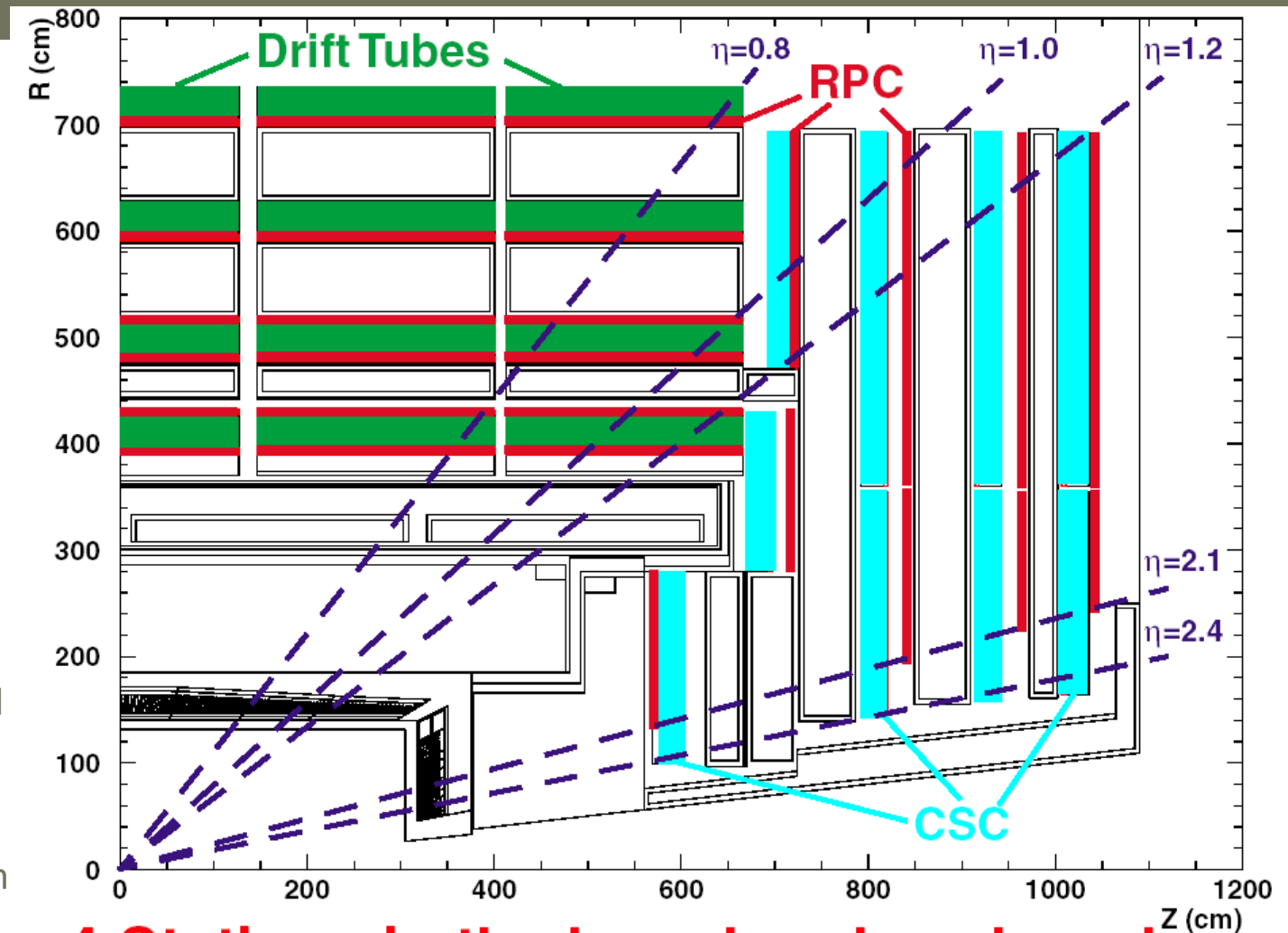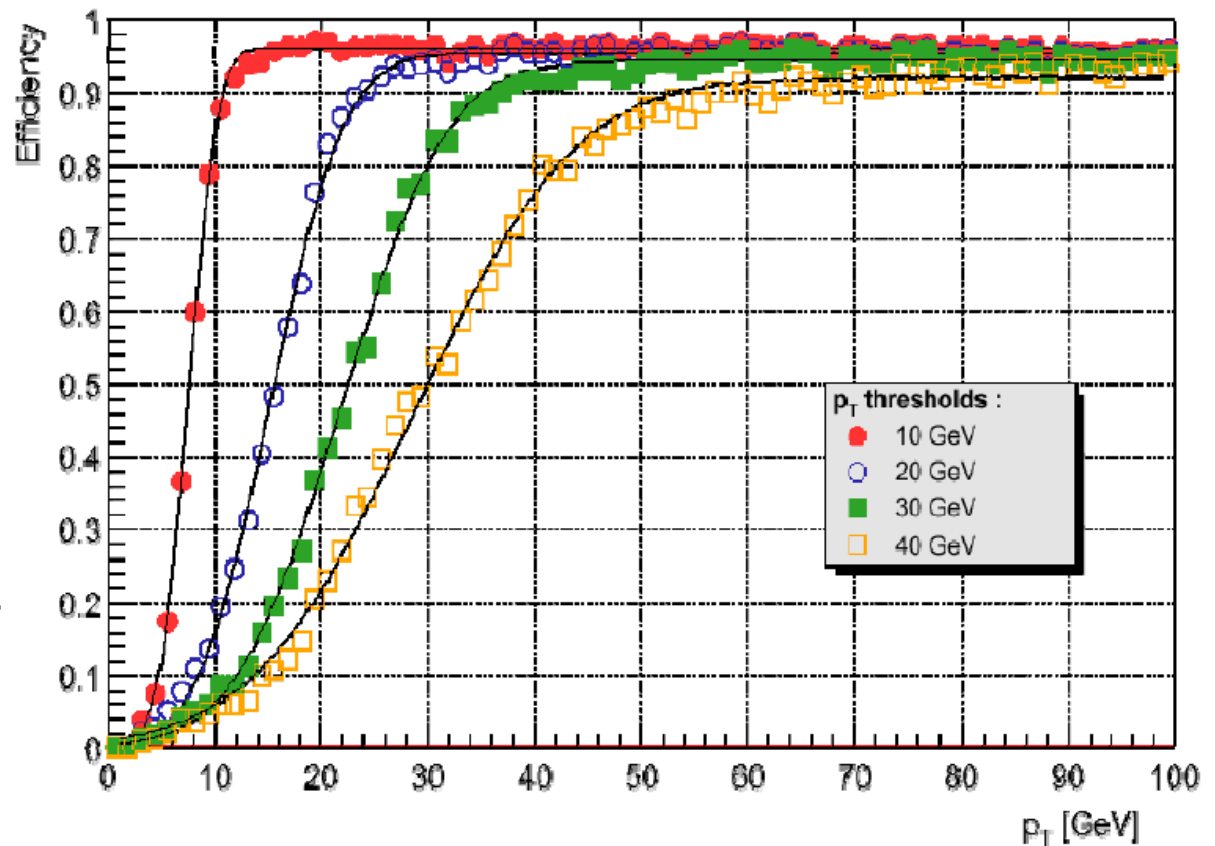

4 Stations in the barrel and each endcap

Data-analysis at colliders - Prof. J. D'Hondt (VUB)

Master Level Course 2007-2008

Pag. 73

Level-1 muon trigger

Similar efficiency function for muons as for electrons

- For the muon Level-1 trigger several sub-detectors are combined to reduce the "ghost" muon candidates which are also triggering the system.

- At a threshold at pT>25GeV/c a "ghost" rate of 0.07% and an efficiency for real muons of 90% is reached by taking the "AND" logic of the three sub-detectors (in $|\eta|<2.4$).

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 74

Level-1 trigger rates

The overall rate of online selected events is dropping fast with increasing thresholds. Need to tune the thresholds in order to fill the complete bandwidth of 100kHz with useful data events.

Data-analysis at colliders - Prof. J. D'Hondt (VUB)

Master Level Course 2007-2008   Pag. 75
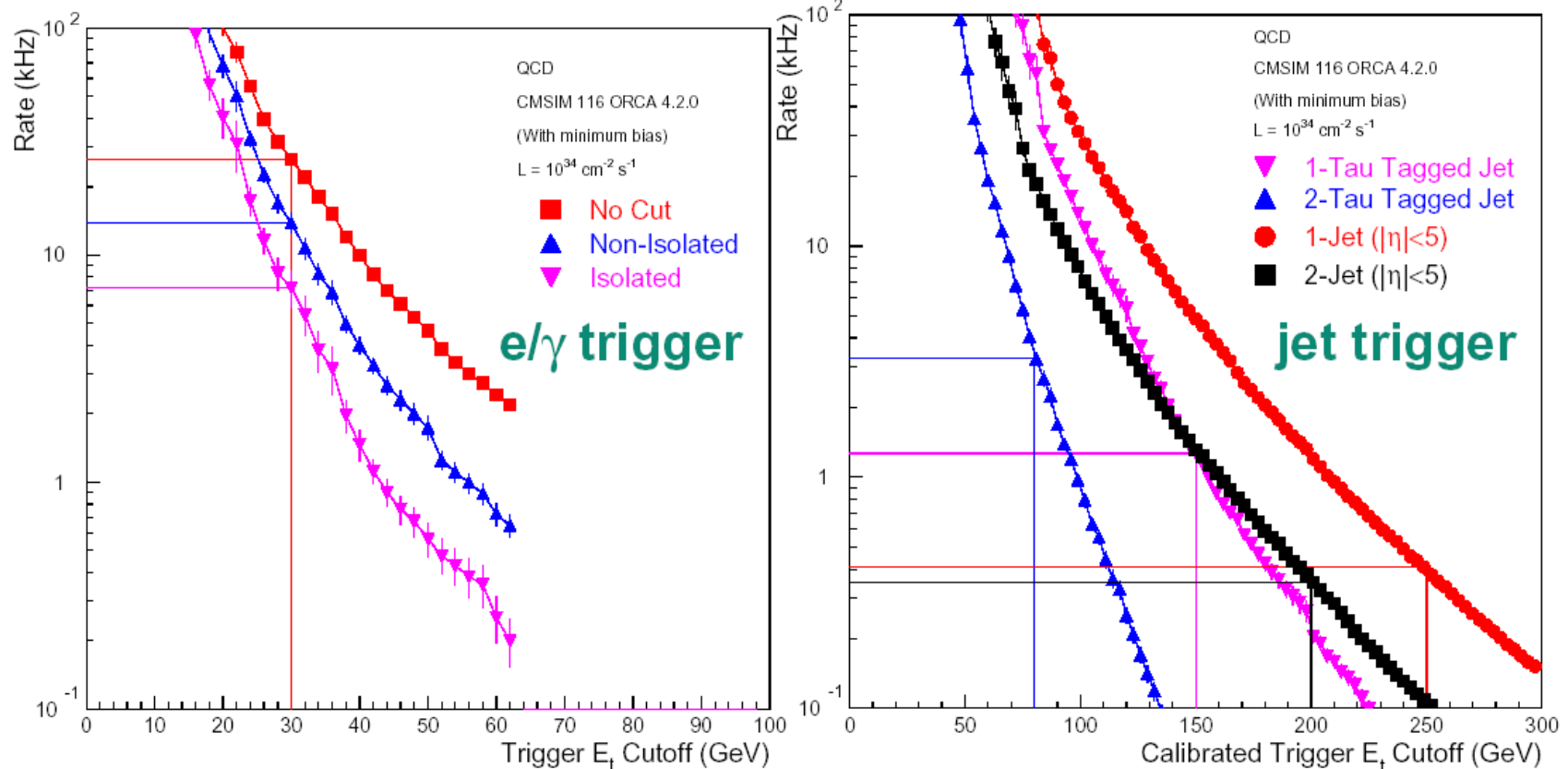
Level-1 trigger rates

Same challenges as at the Tevatron but higher energies, much higher luminosity & more interactions/crossing

Cut on E_T and p_T to discriminate against QCD backgrounds

- Higher E_T cuts than Tevatron needed
- More boost & don't loose efficiency

Examples of unprescaled high p_T trigger thresholds and rates:

	CDF L1		CDF L2		LHC L1	
	P_T Cut	Rate (HZ)	P_T Cut	Rate (Hz)	Pt Cut	Rate (Hz)
Single μ	4 GeV/c	280	12 GeV/c	25	20 GeV	10k
Single e	8 GeV	240	16 GeV	30	30 GeV	20k
Single γ	8 GeV	2400	18 GeV	60	30 (GeV)	20k
Single Jet	10 GeV	10K	90 GeV	10	300 (GeV)	200

CDF Rates for $2 \times 10^{32} cm^{-2} s^{-1}$, scaled from $3 \times 10^{31} cm^{-2} s^{-1}$

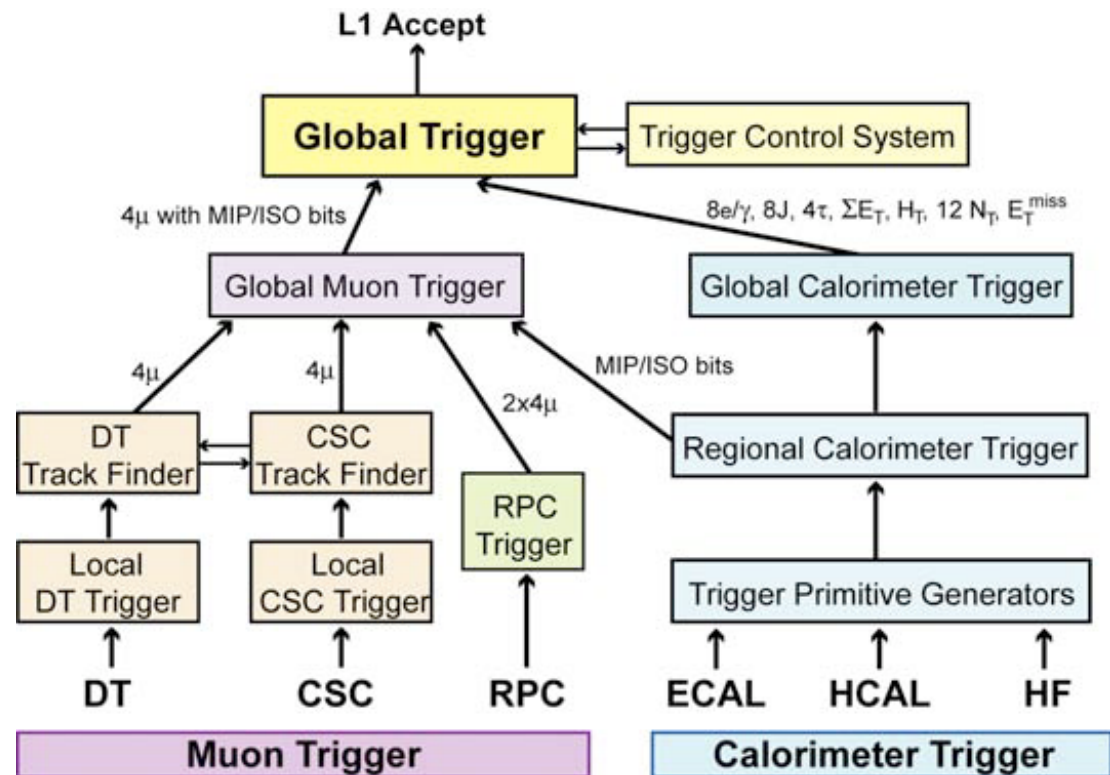LHC rates for $10^{34} cm^{-2} s^{-1}$, from N. Ellis, LECC Workshop 2002 at Colmar

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 76

Global Level-1 trigger (CMS)

- The decision is based on the algorithm calculations and on the readiness of the sub-detectors and the DAQ system, which is determined by the Trigger Control System (TCS).

Level-1 Trigger Objects used for the decision:

- **Jets**: 4 central, 4 forward, 4 tau-tagged & multiplicities
- **Electrons**: 4 isolated, 4 non-isolated
- 4 **muons**
- Missing E_T and total E_T

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 77

Level-1 is hardware

This selection requires hardware very near to the detector, hence new technologies or evolving technologies when we go to larger interaction rates (ASICs for radiation tolerant on-detector electronics and FPGAs which are more flexible, and communication technologies).

Zeus Calorimeter Trigger:
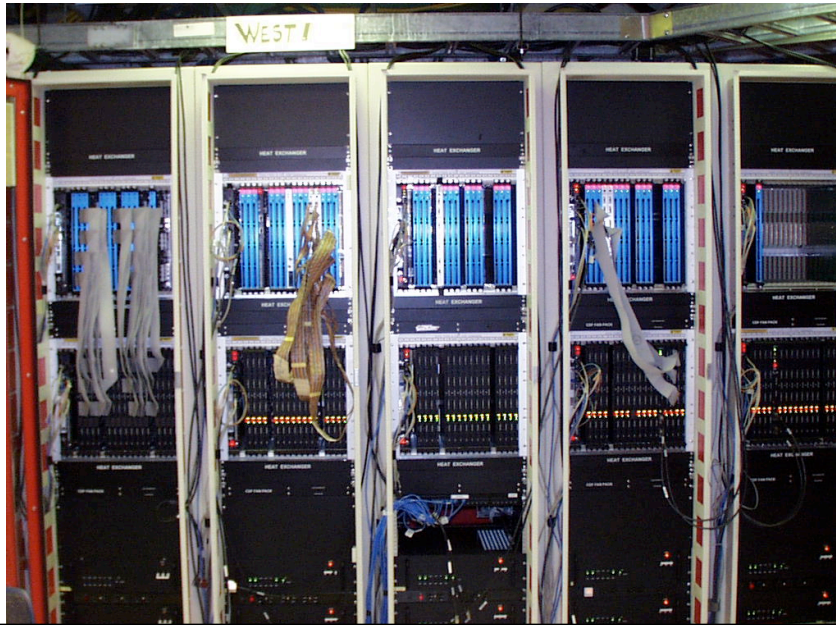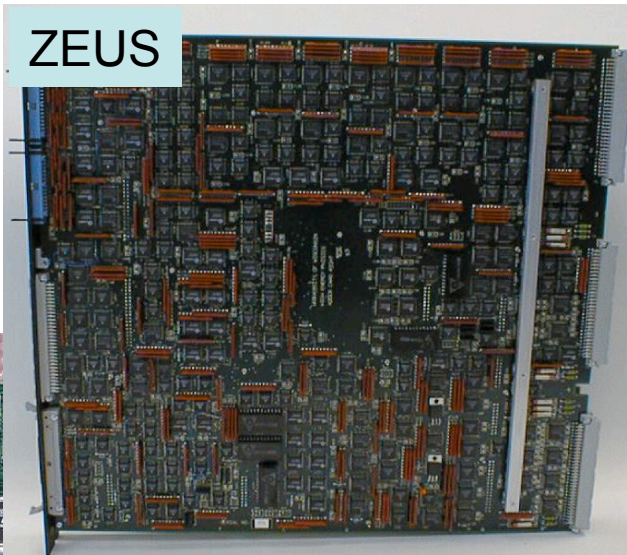
16 9U Crates

CDF L1&L2 Calorimeter Triggers: 12 9U Crates

ASIC: Application-Specific Integrated Circuits
FPGA: Field-Programmable Gate Arrays

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
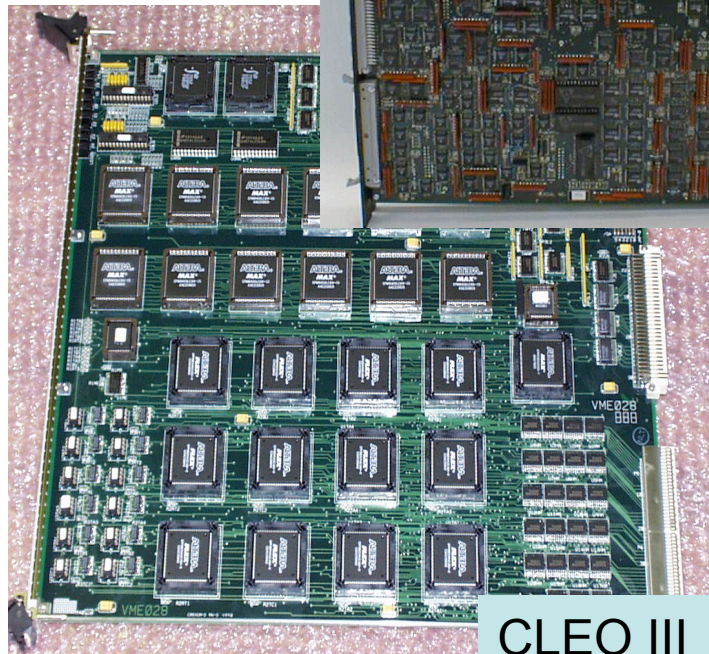Master Level Course 2007-2008

Pag. 78
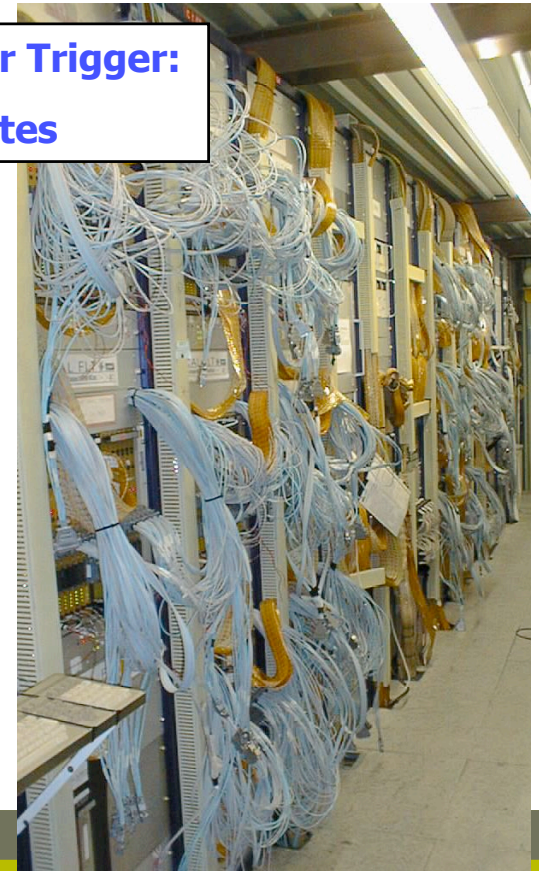
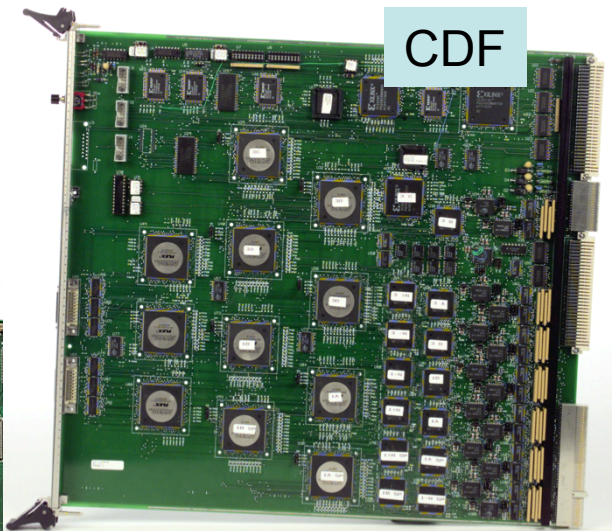# Level-1 is hardware

ZEUS

CDF

CLEO III
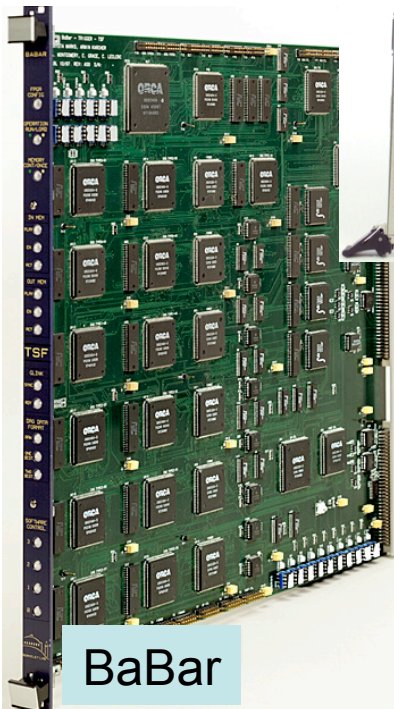
BaBar

They look all the same, but use different techniques and serve different purposes

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 79

Trigger and DAQ trends

- Detectors become more and more complex, hence more and more read-out channels. Hence the event size in #byte is increasing demanding for larger bandwidths.

- The luminosity of the accelerators is increasing as we search for more and more rare processes. Hence the Level-1 rate in Hz is increasing demanding for faster processors.

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 80
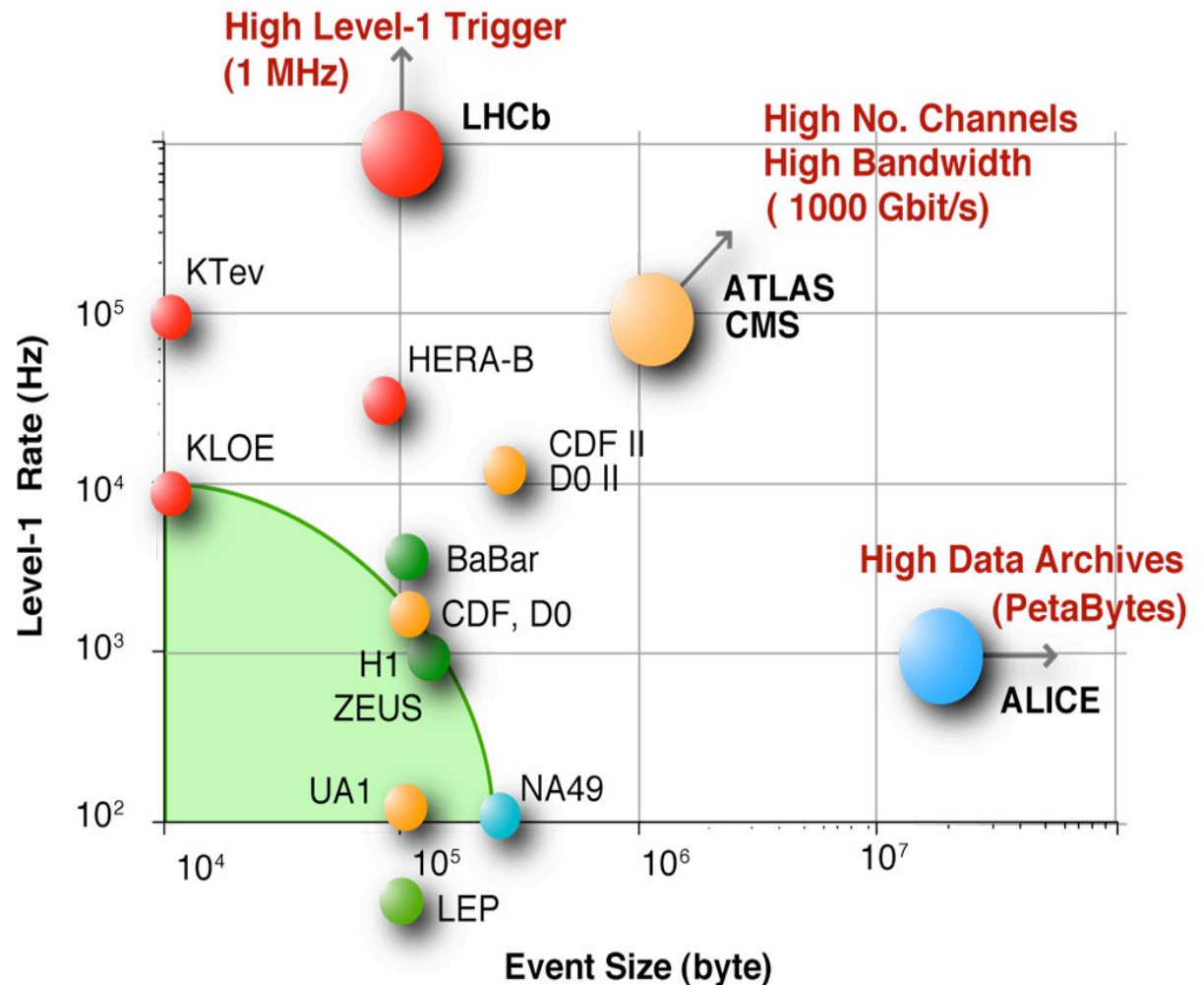
Level-1 summary

- First step in the online selection of the collision events to be saved on disk, based on hardware deployed very near to the detector.

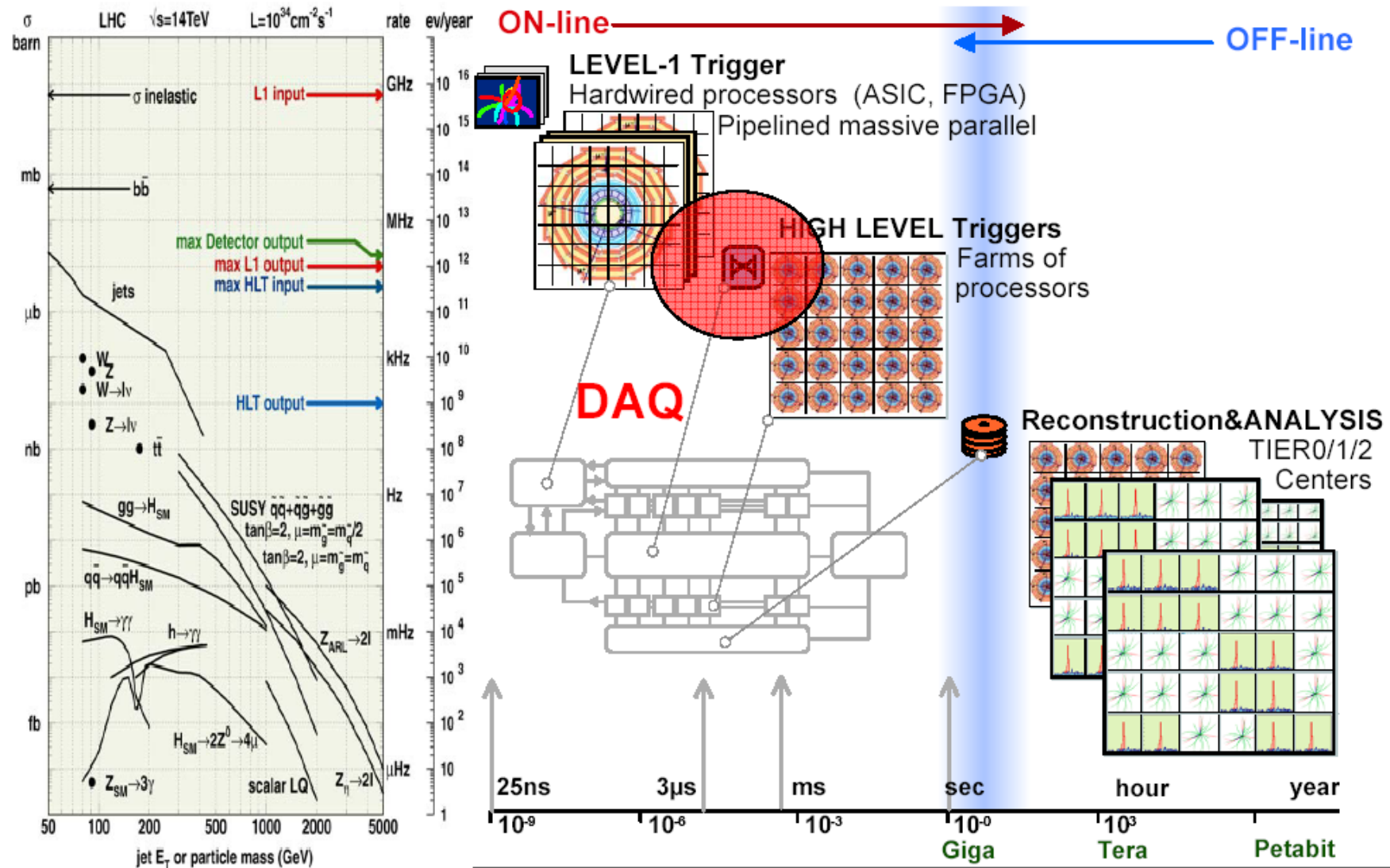- Large challenges for the LHC which deals with very high interaction rates (40MHz) and the detectors with a huge amount of read-out channels.

- Pipelined system for $3\mu s$ to decide on the event, therefore uses only part of the detector information (calorimeter & muon systems). Remember most of the time is spend on the transportation of the signal through cables.

- Reduce the rate from 40MHz to about 100kHz.

- Complex process to optimize the use of the bandwidth (which is limited) to the next trigger step. Tuning of threshold requirements for each stream (eg. 1 electron, 1 electron isolated, 2 muons, 3 jets, …) results in a "trigger table" of the experiment. One should take into account the cross sections of the signal processes to be observed or discovered versus the rate of the background processes or the detector background consisting out of "fake" objects (eg. "ghost" muons).
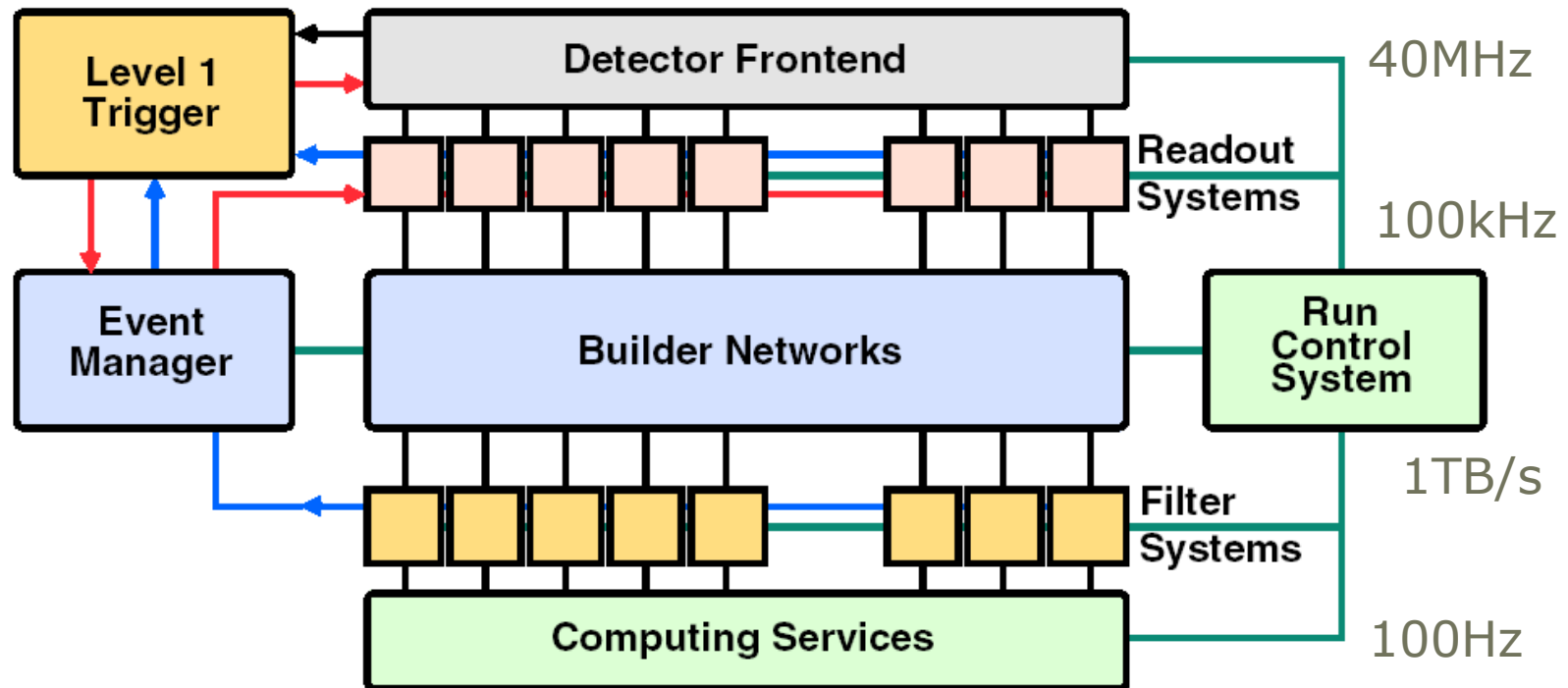
Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 81

DAQ system

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 82

The "Event Builder"

- The main task of the DAQ system is to read the data corresponding to each event selected by the Level-1 trigger out of the detector Front-End Drivers (= the readout electronics specific to each type of detector). Beyond that to concentrate this into one data structure, called the physical "event", and to forward this event to the filter farm for the final High Level decision if the event is to be written on disk or not for permanent storage. This complex system is called the "Event Builder".

- The main parameters of this system is dictated by the physics program of the experiment
  - maximum Level-1 trigger rate of 100kHz
  - maximum rate after the High-Level Trigger of 100Hz
  - average event size of 1MB (dictated by the occupancy of the CMS detector channels)

- This is implemented in several stages.

Main view of the DAQ system

Detector Frontend — 40MHz

Level 1 Trigger

Readout Systems — 100kHz

Event Manager

Builder Networks

Run Control System

Filter Systems — 1TB/s

Computing Services — 100Hz

Collision rate	40 MHz	I-O units bandwidth (512+512)	400 MByte/s
LV1 Maximum trigger rate	100 kHz	Builder network (512x512 port)	≥ 500 Gbit/s
Average event size	≈ 1 Mbyte	Event filter computing power	≈ 5 10^6 MIPS
Data production	≈ Tbyte/day	High Level Trigger acceptance	1 - 10 %
Event Flow Control	≈ 10^6 Mssg/s	Overall dead time	≤ 2%

Data flow in proton collisions

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 85
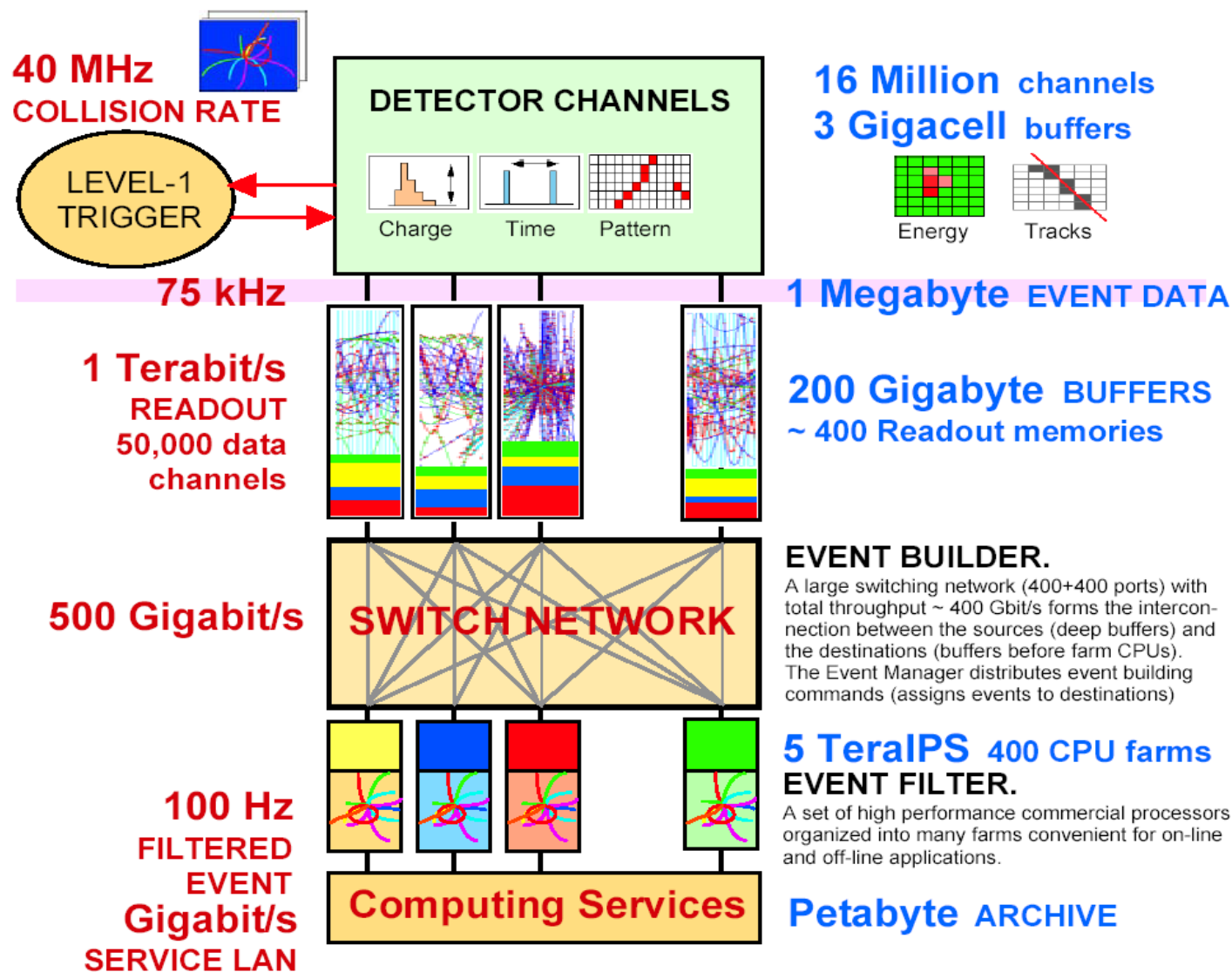
Main view of the DAQ system

From fragments of the event in different read-out systems, to full events.
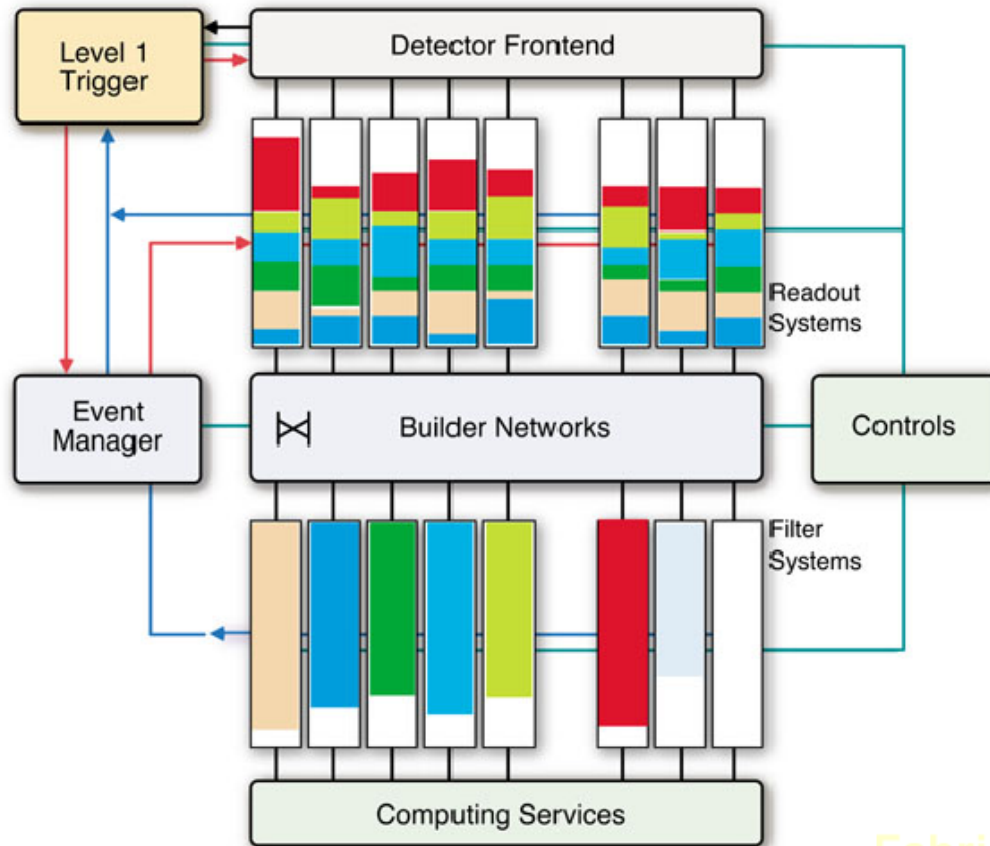
Event fragments :
Event data fragments are stored in separated physical memory systems

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

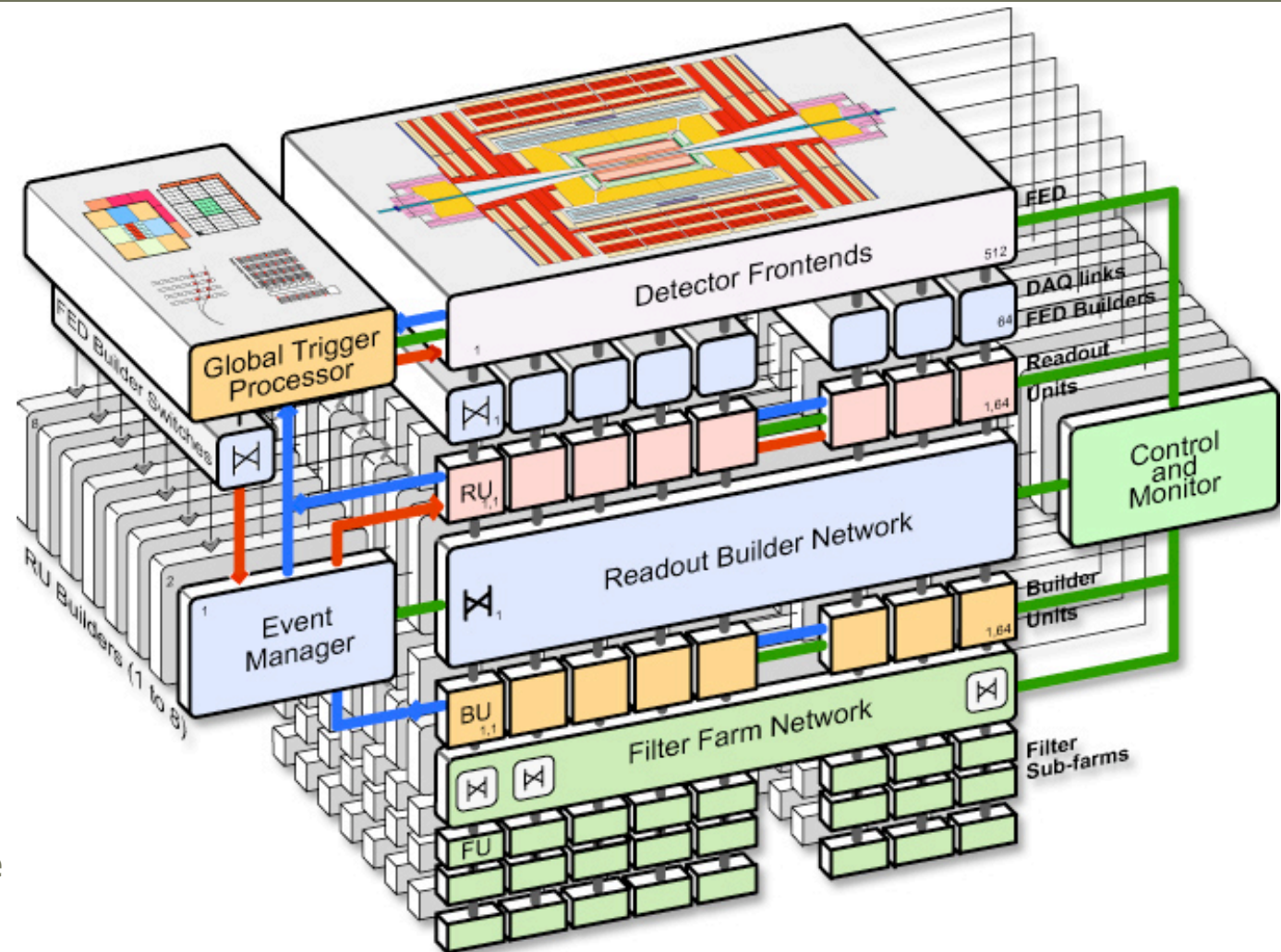**Hardware:**
Fabric of switches for builder networks
PC motherboards for data Source/Destination nodes

Main view of the DAQ system

The factorization of the DAQ functions into four tasks facilitates the design. Deep buffers allow to match the widely different operating rates of the elements at each stage.
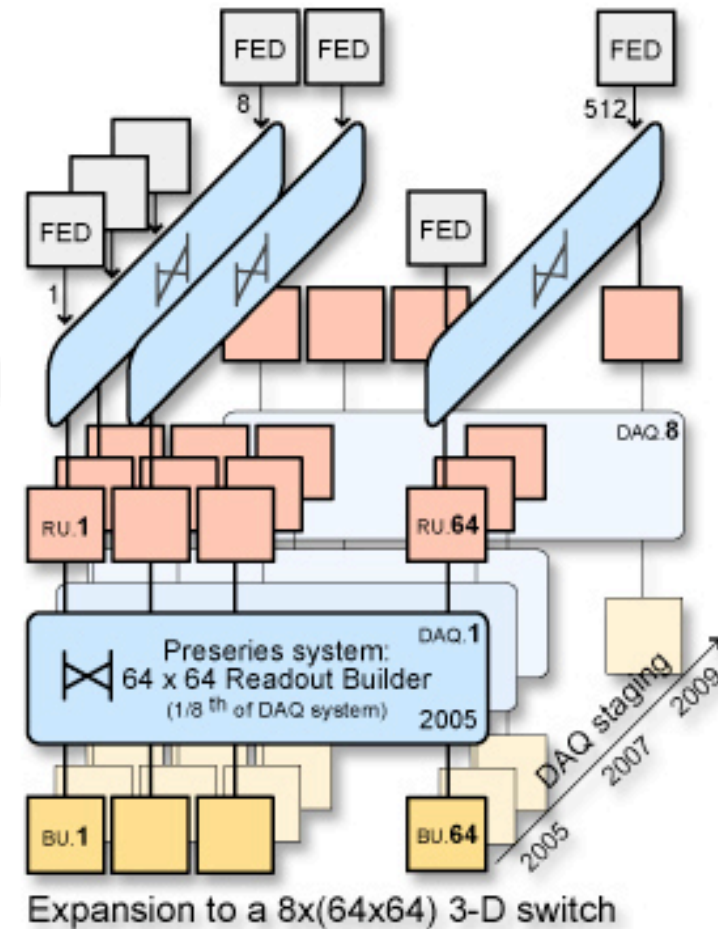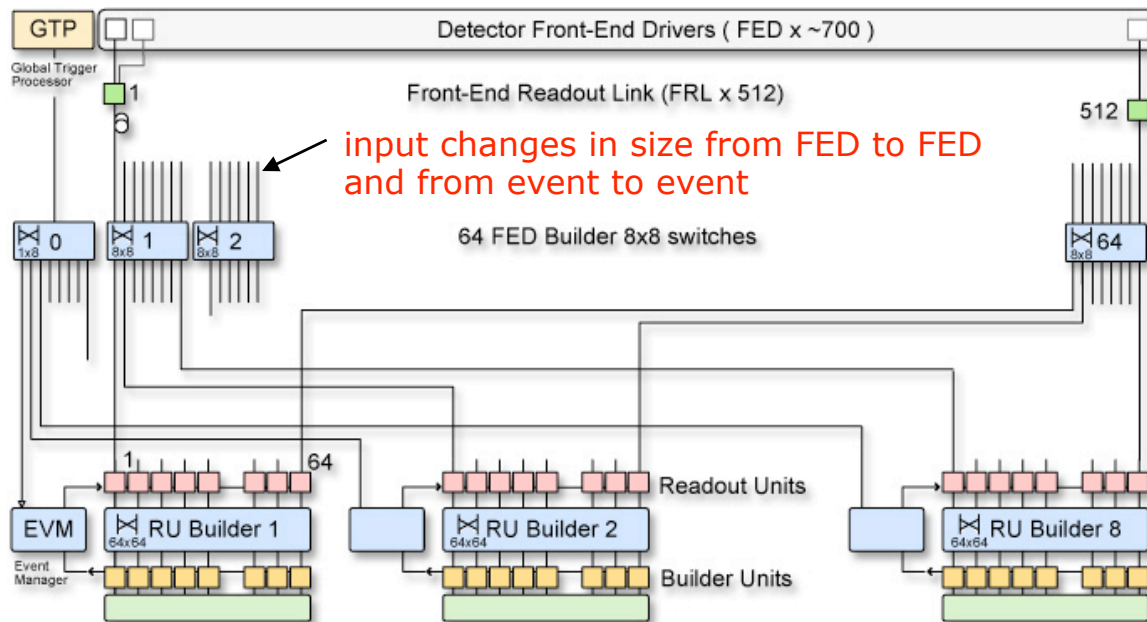
- Detector read-out (fragments)
- Event building (switch network)
- Selection stage (High Level Trigger or HLT)
- Analysis/storage stage (computing services)

For example the FED delivers at a rate of 100kHz (every 10μs), but it takes the event building stage 1MB/2Gb/s=4ms to read the complete event.

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 87

Scalability of the DAQ system

- The DAQ system is duplicated 8 times for which the deployment is staged (=flexibility). They are connected with a 8x8 switch network (FED Builder), hence for each event in the system there are 8 potential destinations as far as the Front-End information is concerned.
- Basically one extra layer of switches allows to increase the DAQ performance when duplicating the full system which can come whenever the physics program requires.

input changes in size from FED to FED and from event to event

Expansion to a 8x(64x64) 3-D switch

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 88

FED Builder specifications

- The main functions of the FED Builder are:
  - transport the fragments from the underground area to the surface
  - assemble event fragments of variable size from some ~700 FEDs into 64 super-fragments of roughly equal size (16kB each)
  - these super-fragments are given one of the 64 Readout Unit Builders in a particular DAQ slice

2kB output (average!)

2kB x 100kHz = 200MB/s I/O switches (average!)

has an output buffer to select the correct output port for each event

put associate trigger info in the correct Readout Unit Builder
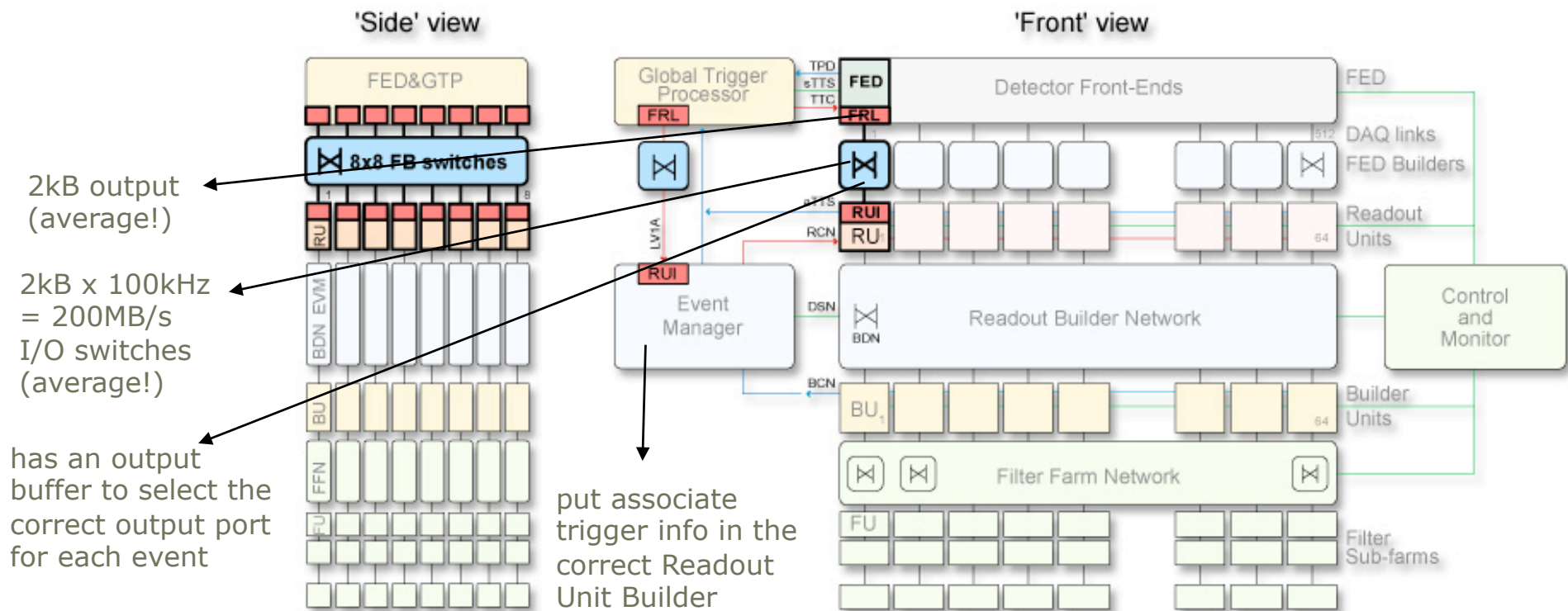
Main view of the DAQ system
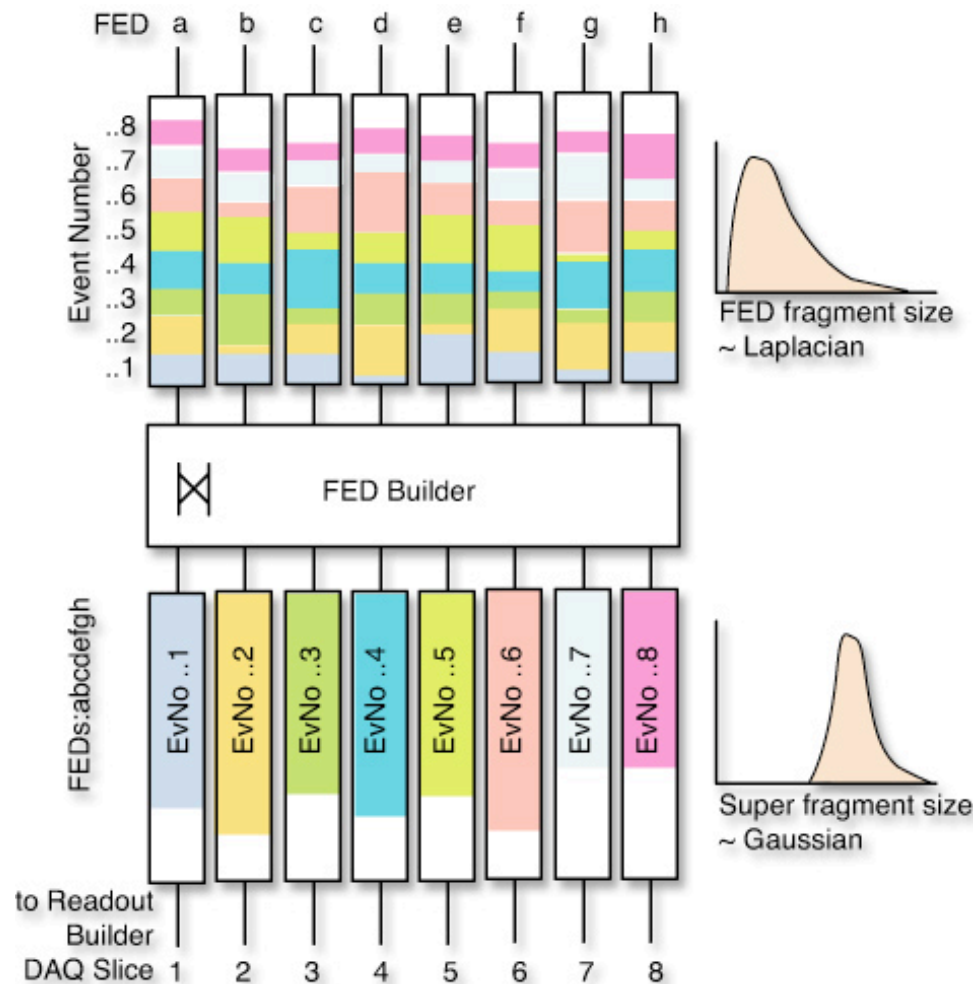
Close-up illustration between the FED and the Readout Units :
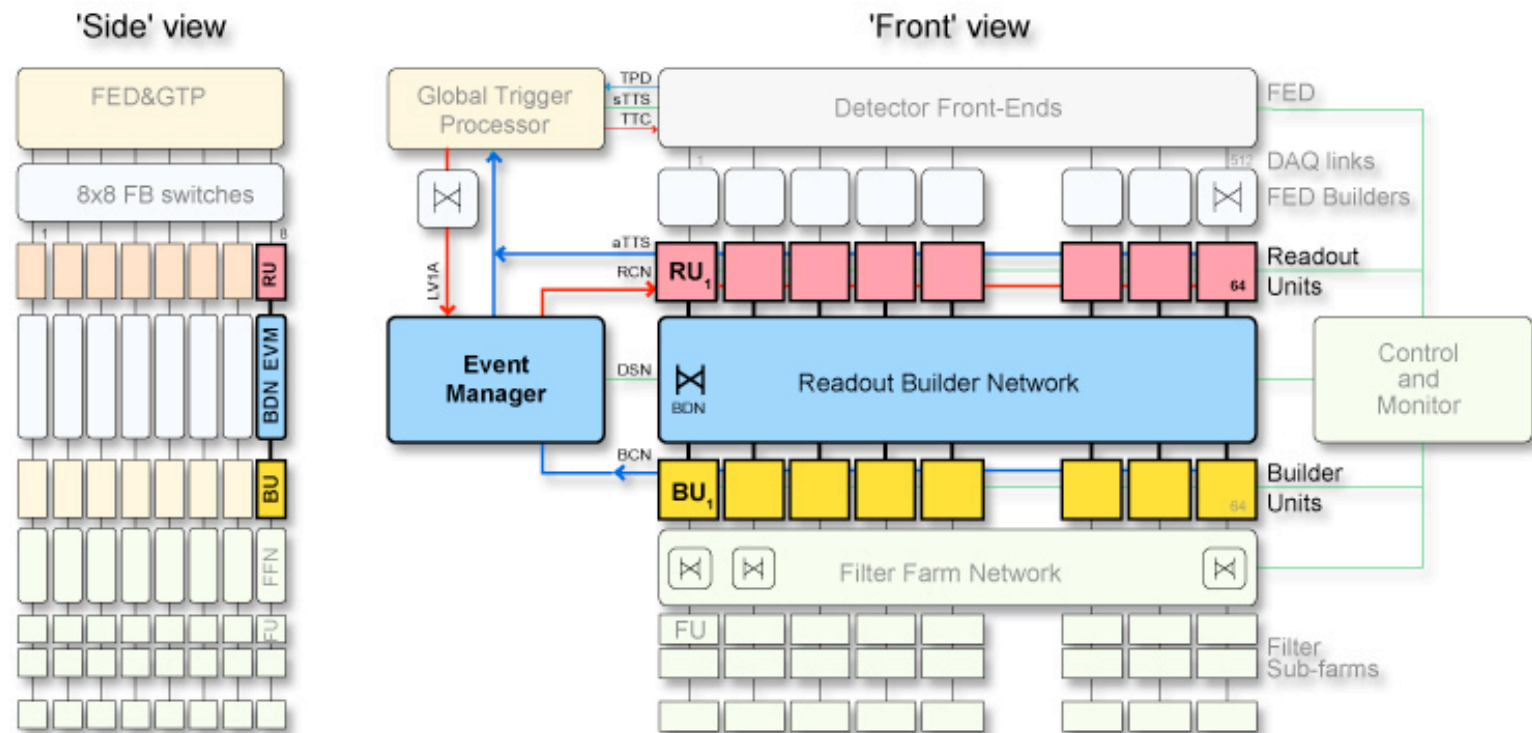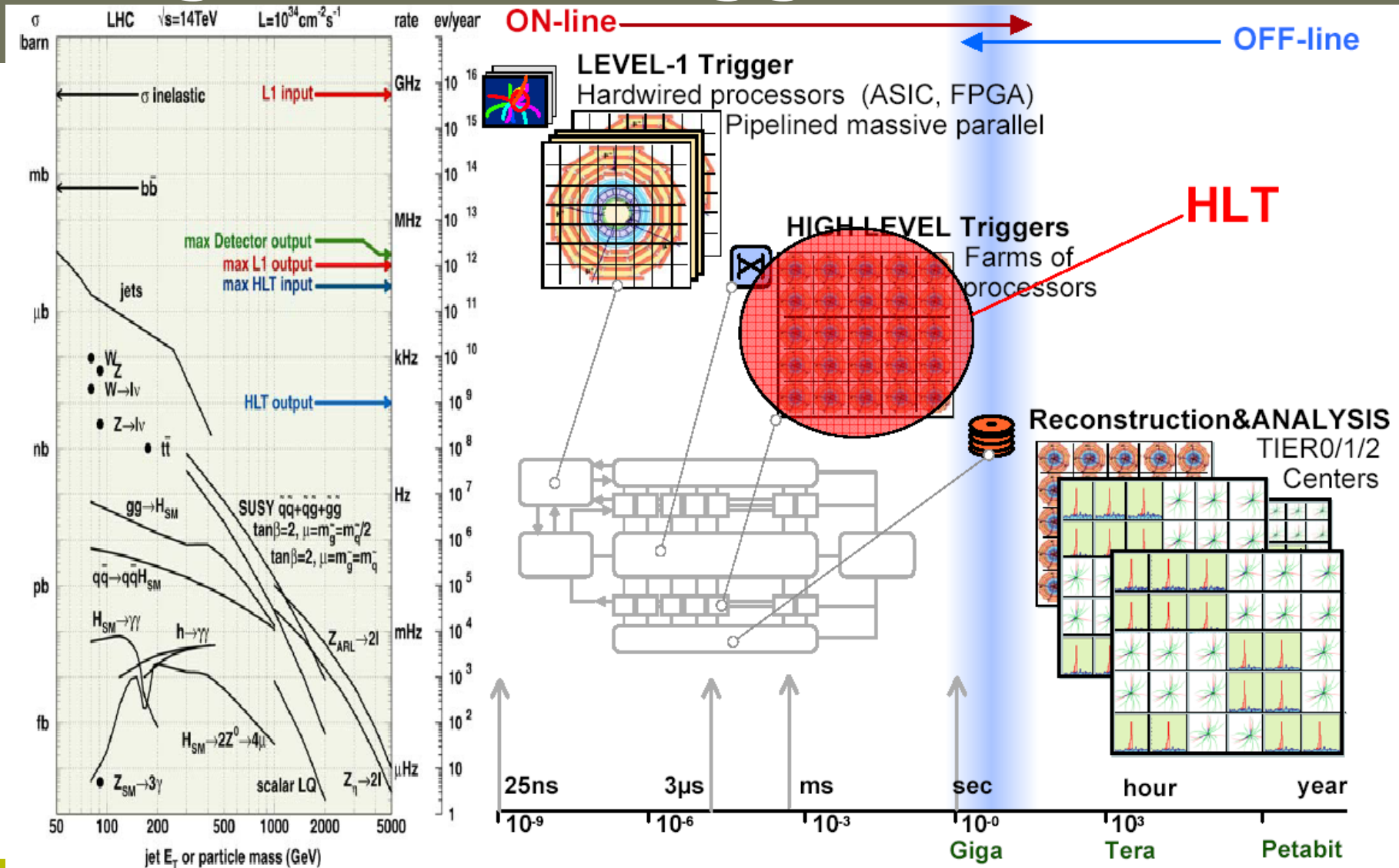
- The size of the fragments of the event which are being formed is increasing (FED fragments of 2kB to super-fragments of 16kB on average).
- The amount of links is being reduced.

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 90

Readout Unit Builder

Via a network of switches the event is built from the fragments deliver from the Front-End electronics of the detectors. At a rate of 12.5kHz (= 100kHz / 8 DAQ slices) and a mean event size of 16kB from 64 RUs = 1MB, the total sustained throughput of the switch fabric should be about 200MB/s (including event-by-event fluctuations).

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 91

High-Level Trigger

High-Level Trigger

Detectors — 40 MHz

Front-end pipelines (10^7 channels) — Level 1

Readout buffers (1000 units) — 75 kHz

Event builder (10^3 x 10^3 fabric switch)

Processor farms ($4 \cdot 10^6$ MIPS) — High Levels

10^2 Hz
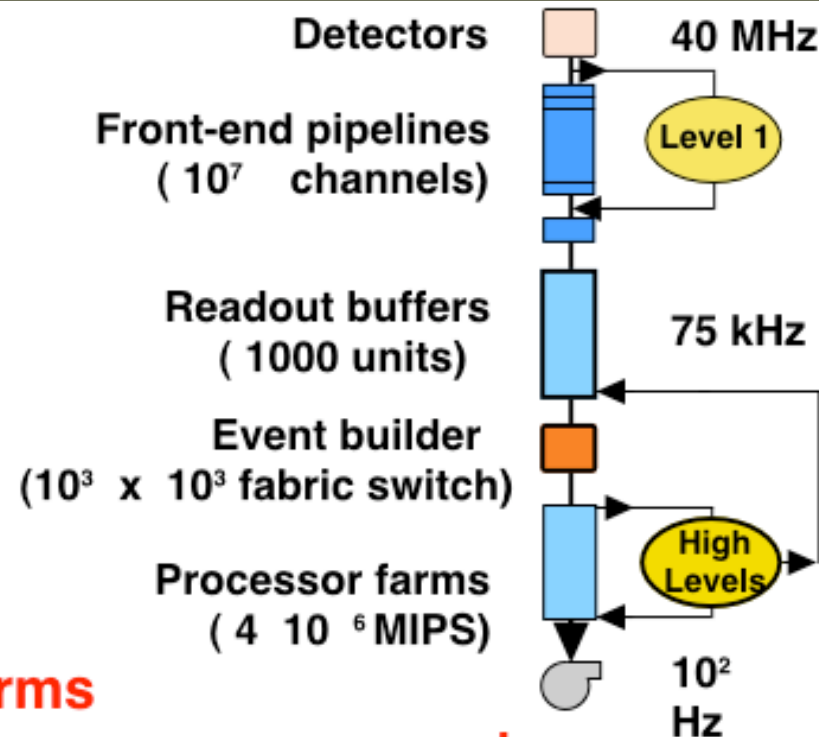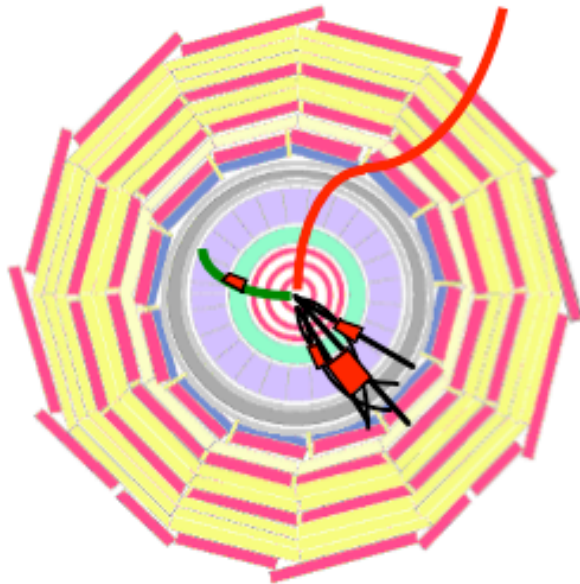
High level triggers. CPU farms

- Finer granularity precise measurement
- Clean particle signature (π^0-γ, isolation, ...)
- Kinematics. Effective mass cuts and topology
- Track reco and matching, b,τ-jet tagging
- Full event reconstruction and analysis

Successive improvements : background event filtering, physics selection

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 93

HLT table

- With each object or multi-object topology an HLT stream is defined based on software loaded on a multi-processor PC farm.
- The tuning of the bandwidth to be allocated to which HLT stream is guided by the physics program of the experiment (eg. Higgs boson, SUSY, ... but also calibration issues).
- It is important to check the trigger efficiency of the signal processes of interest and to adapt the trigger thresholds accordingly. Most of the new physics phenomena of interest are created at larger mass scales and therefore give objects at large transverse momenta.
- At this point we use all relevant information in the event and the level of complexity of the HLT algorithms define the scale of the computing farm.
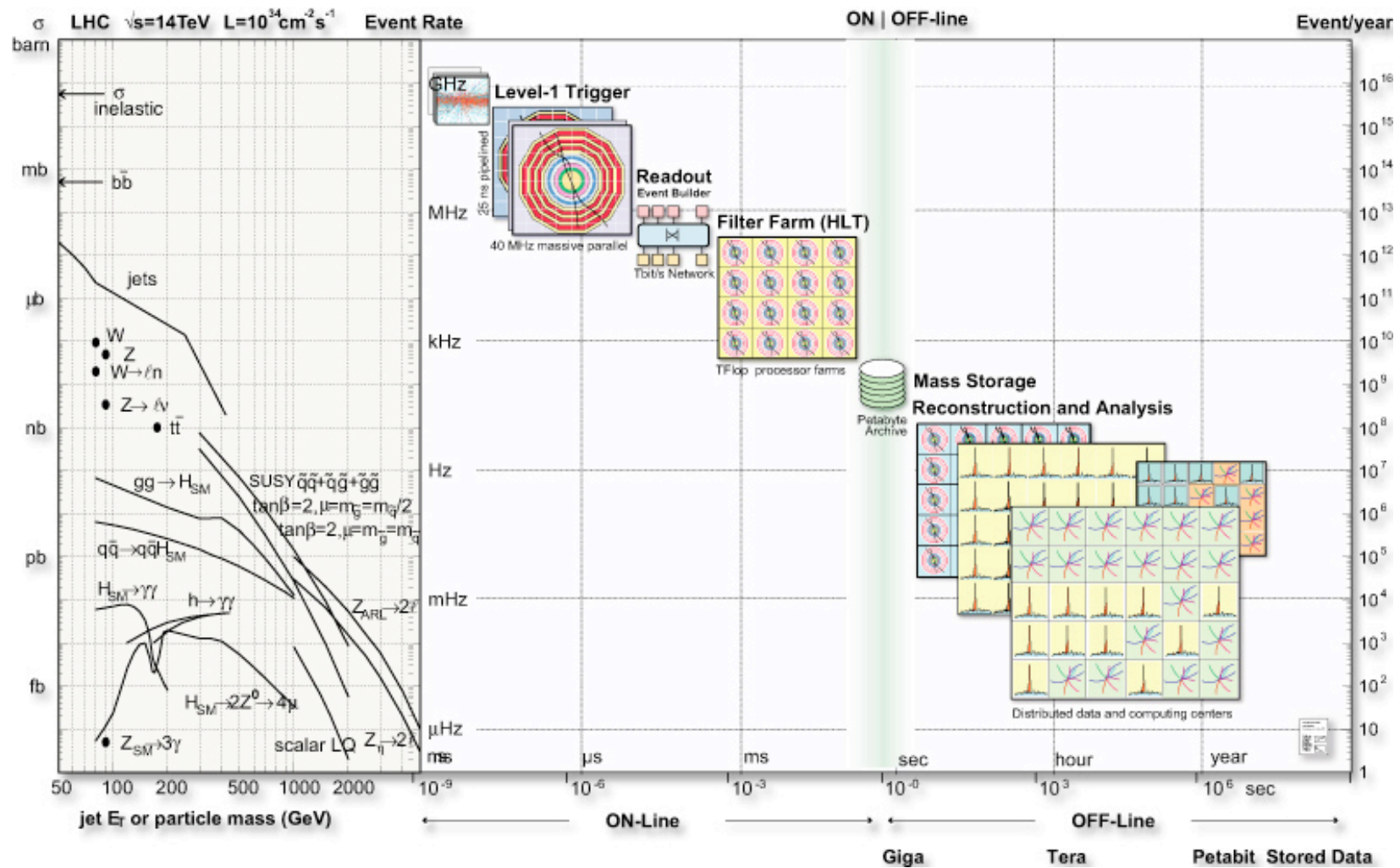- Next course on reconstruction tools.

Referentie:

http://arxiv.org/ftp/hep-ex/papers/0512/0512077.pdf

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 94

HLT table for low luminosity

Trigger	Threshold (GeV or GeV/c)	Rate (Hz)	Cumulative Rate (Hz)
inclusive electron	29	33	33
di-electron	17	1	34
inclusive photon	80	4	38
di-photon	40, 25	5	43
inclusive muon	19	25	68
di-muon	7	4	72
τ-jet * $\not{E}_T$	86 * 65	1	73
di-τ-jets	59	3	76
1-jet * $\not{E}_T$	180 * 123	5	81
1-jet OR 3-jets OR 4-jets	657, 247, 113	9	89
electron * τ-jet	19 * 45	0.4	89.4
muon * τ-jet	15 * 40	0.2	89.6
inclusive b-jet	237	5	94.6
calibration and other events (10%)*		10	105
TOTAL			**105**

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

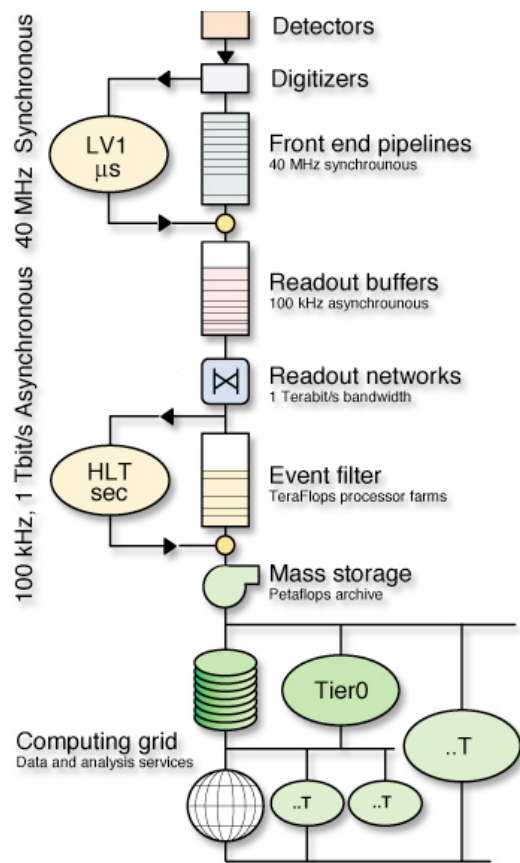Pag. 95

High-Level Trigger

The amount of time needed to select the HLT object and the rate of these events after Level-1 trigger. This gives an estimate on how much computers are needed in the HLT farm for 15.1kHz of Level-1 trigger rate (calculations are of course evolving with the evolution of the computer processors). Today it results in about 2000 CPU's for the start up of the LHC.

Physics Object	CPU time per Level-1 event (ms)	Level-1 Trigger rate (kHz)	Total CPU time (s)
electron/photon	160	4.3	688
muon	710	3.6	2556
tau	130	3.0	390
jets and $\not{E}_T$	50	3.4	170
electron + jet	165	0.8	132
b-jets	300	0.5	150

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 96

Summary of first part

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
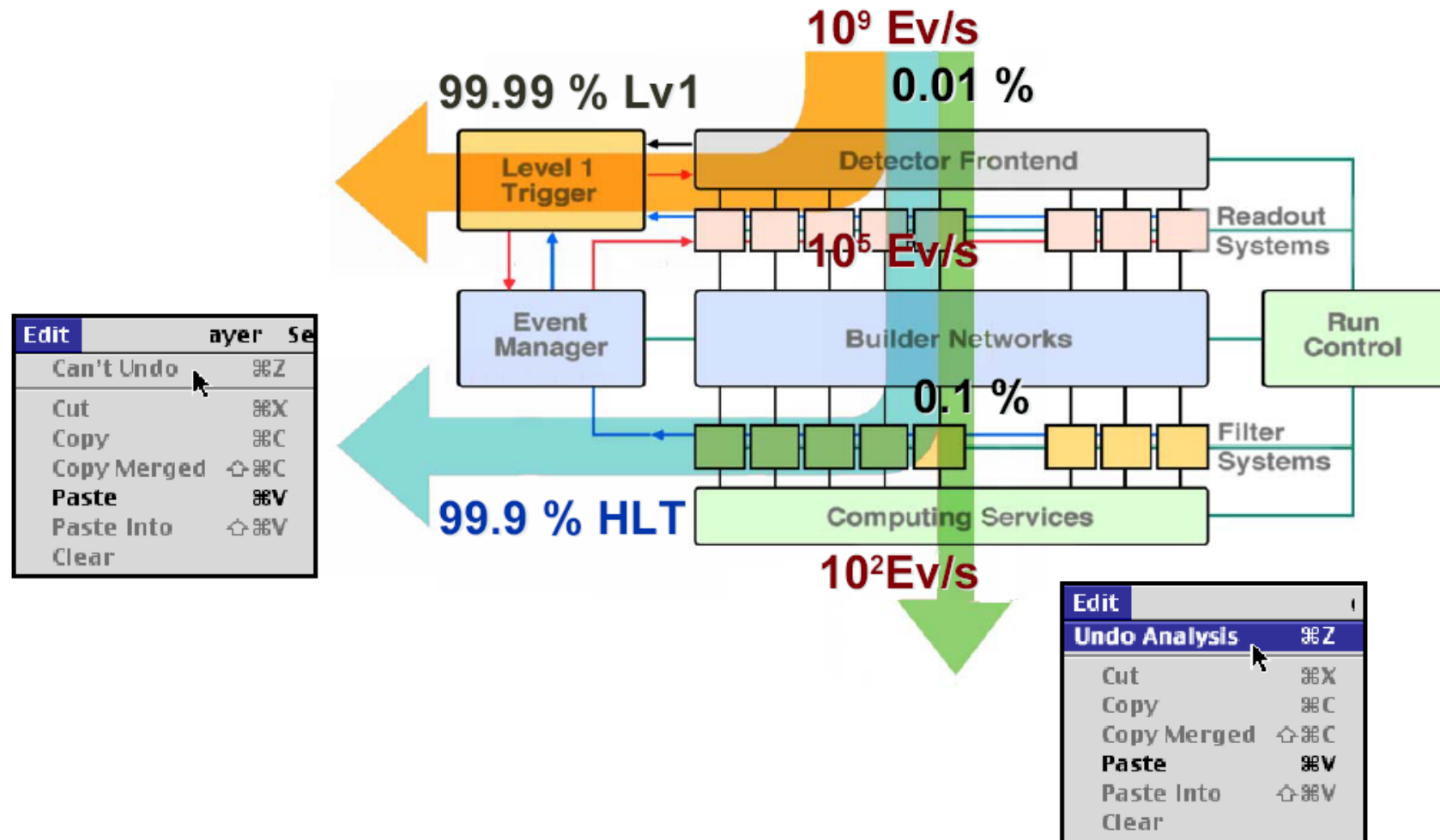Master Level Course 2007-2008

Pag. 97

Summary of first part

- A modern particle physics detector has millions of readout channels and therefore hundreds of Front-End Driver. On top of this the event rate is very high.

- Hence large amount of data has to be processed in a very short time.

- Only a small fraction of the events can be stored on disk, hence very quick a decision has to be made on a complex multi-channel event.

- The online data reduction is performed in several steps: electronics very close to the detector (Level-1 Trigger), a large bandwidth transfer system to built the event (DAQ system) and a more complex layer (High-Level Trigger)
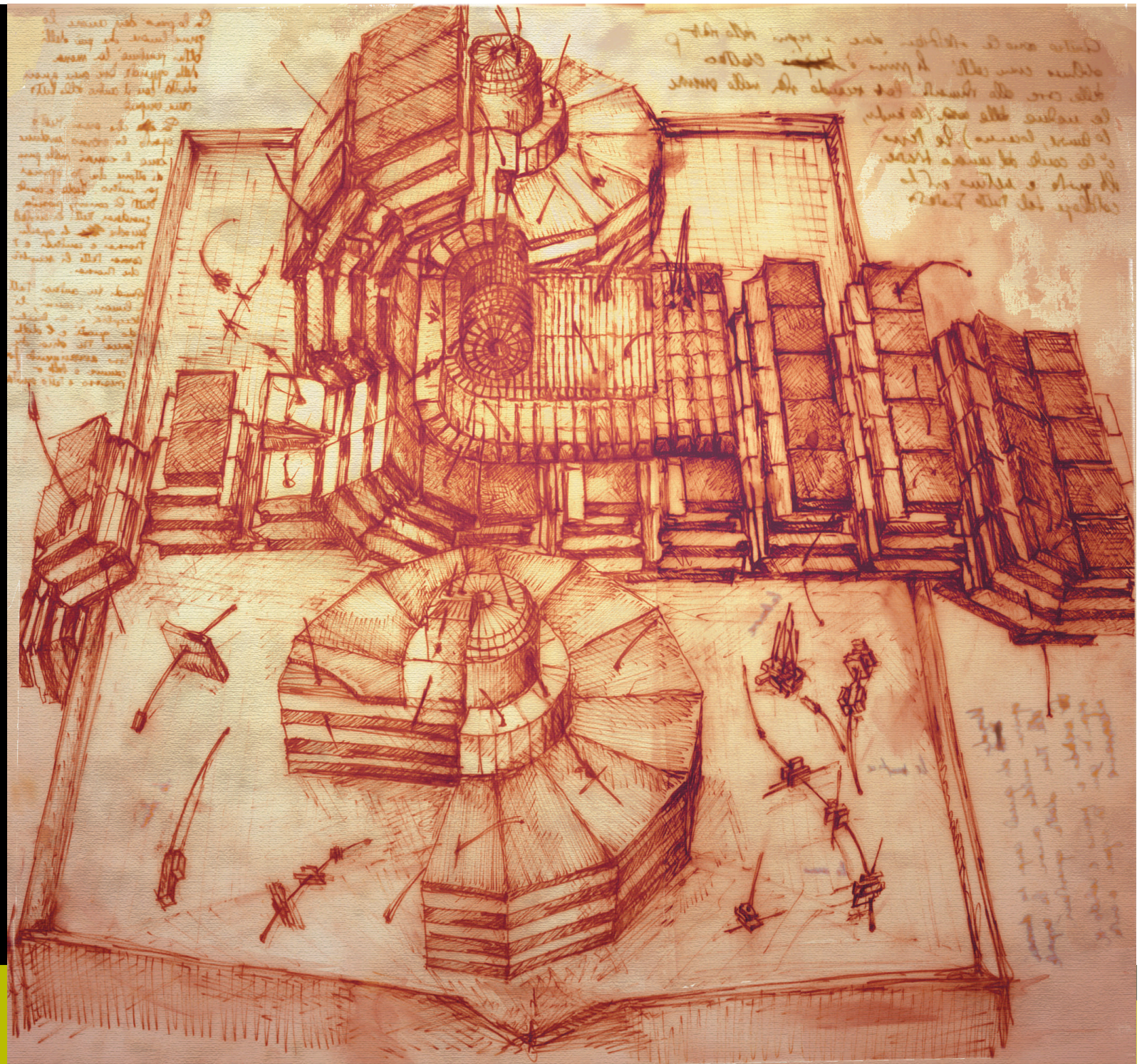
Summary of first part
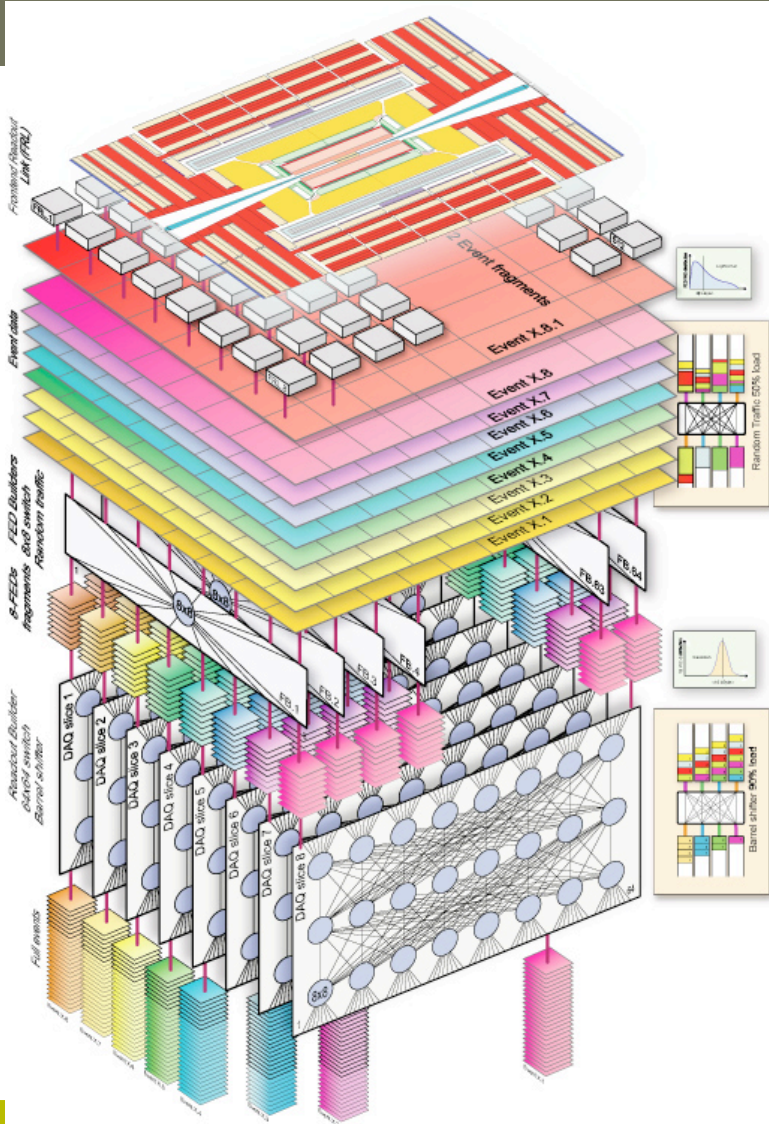
- From signals in the FED's to an event in my ROOT file…

- The DAQ of such experiments is very complex and requires specific electronics and a multi-layer approach to built the events

- The ideas can be applied for other experiments, but the implementation is specific for each (sub-)detector

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 99

Summary of first part

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 100

End
of
Part I

The second part

General concepts of event reconstruction

- Track reconstruction
  - o Note on track finding
  - o General tracking of charged particles in a magnetic field
  - o Multiple scattering effect...
- Track fitting
  - o Track fitting with a least-square procedure
  - o Kalman filter procedure
- Jet reconstruction
  - o Formation of hadronic jets
  - o Different algorithms to cluster objects into jets
  - o Calibration aspects of the energy scale of jets
- Flavour tagging algorithms
  - o Properties of b-quark jets
  - o Different algorithms for b-tagging

Finding tracks

Tracks on the blackboard

Data-analysis at colliders - Prof. J. D'Hondt (VUB)

Master Level Course 2007-2008   Pag. 103

Tracking of charged particles in a magnetic field

Tracks on the blackboard

Multiple scattering effect

Tracks on the blackboard

Putting the pieces together...

Tracks on the blackboard

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 106

Track fitting with a least-square procedure

Tracks on the blackboard

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 107

Kalman filter procedure

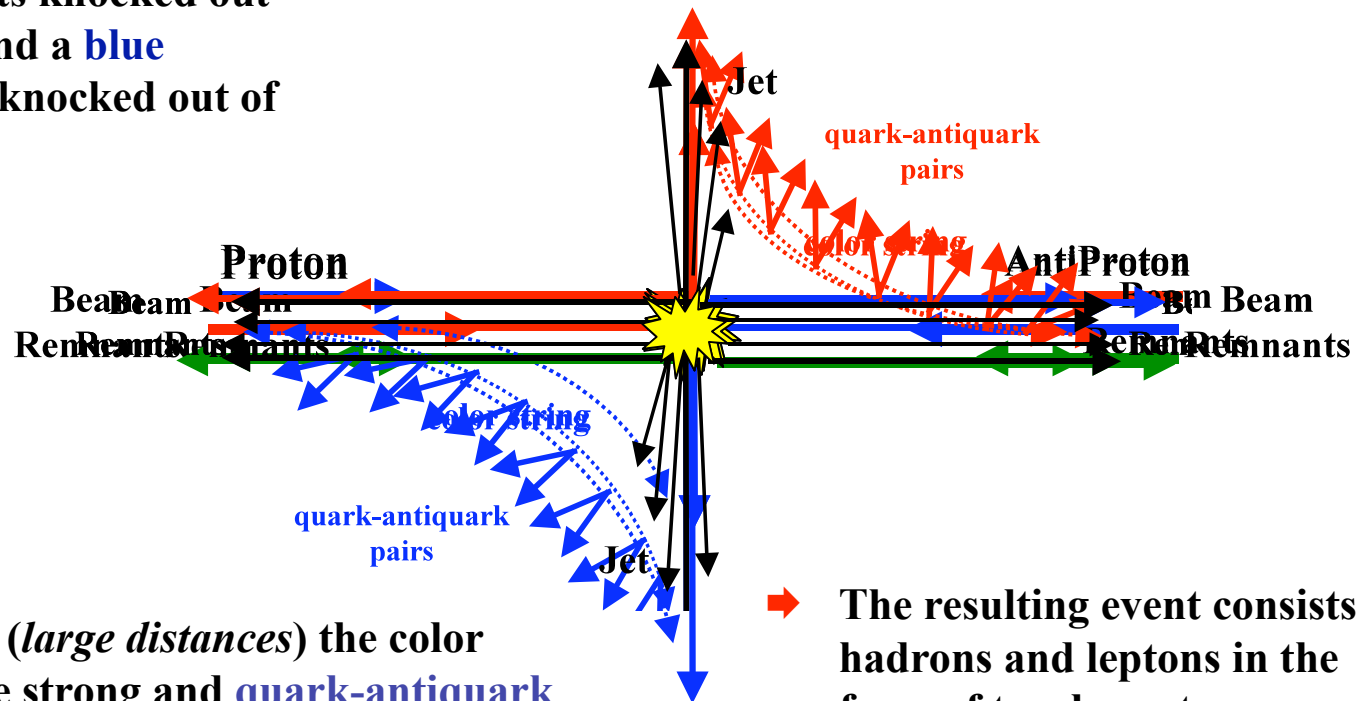Tracks on the blackboard

Formation of hadronic jets

Tracks on the blackboard

Jet reconstruction

Color singlet proton collides with a color singlet proton.

➡ A **red quark** gets knocked out of the proton and a **blue antiquark** gets knocked out of the proton.

➡ At short times (*small distances*) the color forces are weak and the outgoing partons move away from the beam-beam remnants.

Jet

quark-antiquark pairs

color string

Proton

Anti Proton

Beam Remnants

Beam Remnants

Beam Remnants

color string

quark-antiquark pairs

Jet

➡ At long times (*large distances*) the color forces become strong and **quark-antiquark pairs** are pulled out of the vacuum and hadrons are formed.

➡ The resulting event consists of hadrons and leptons in the form of two large transverse momentum **outgoing jets** plus the **beam-beam remnants**.

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 110

String fragmentation

- qq$_{bar}$ pairs are linked by a color string

- While q and q$_{bar}$ move away the energy of the string increases. At a given distance r we have:

$$V(r) = -\frac{A(r)}{r} + \kappa r \qquad \kappa = 1\,\text{GeV/fm}$$

- Above a given distance it becomes energetically favourable to break the string and produce a new q q$_{bar}$ pair (light quarks typically)

- the string brakes and a pair of hadrons is formed
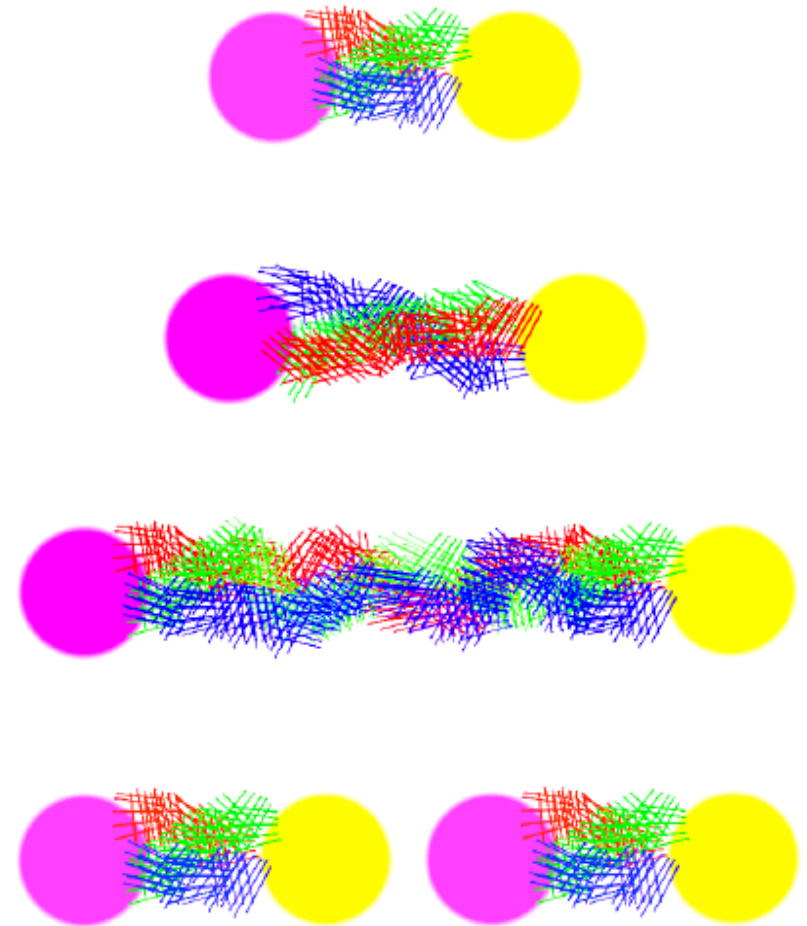
Alternatives

ISAJET: Evolve the parton-shower from Q2 (*virtual photon invariant mass*) to Qmin ~ 5 GeV. Use a complicated fragmentation model to evolve from Qmin to outgoing hadrons.
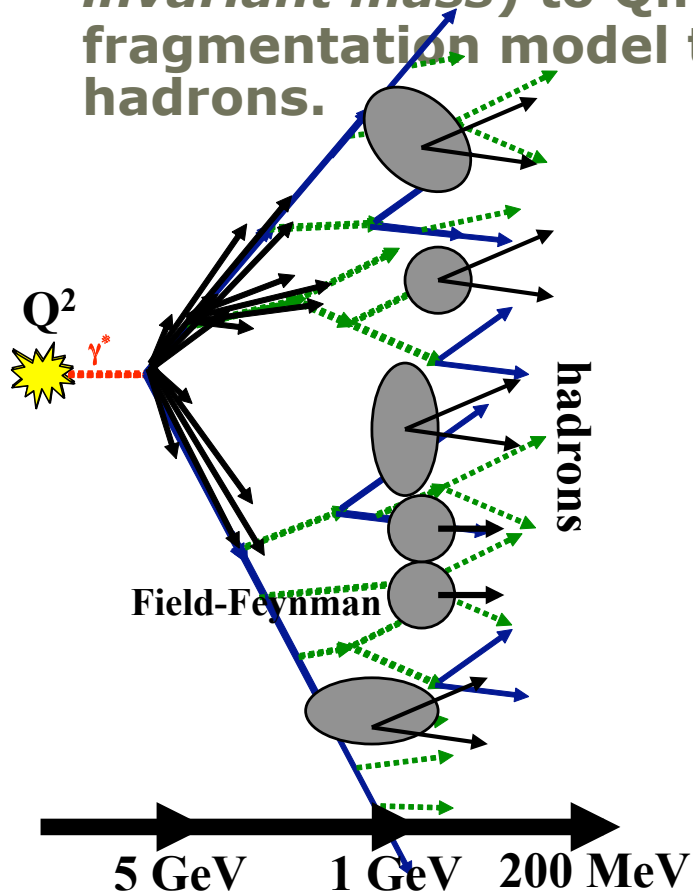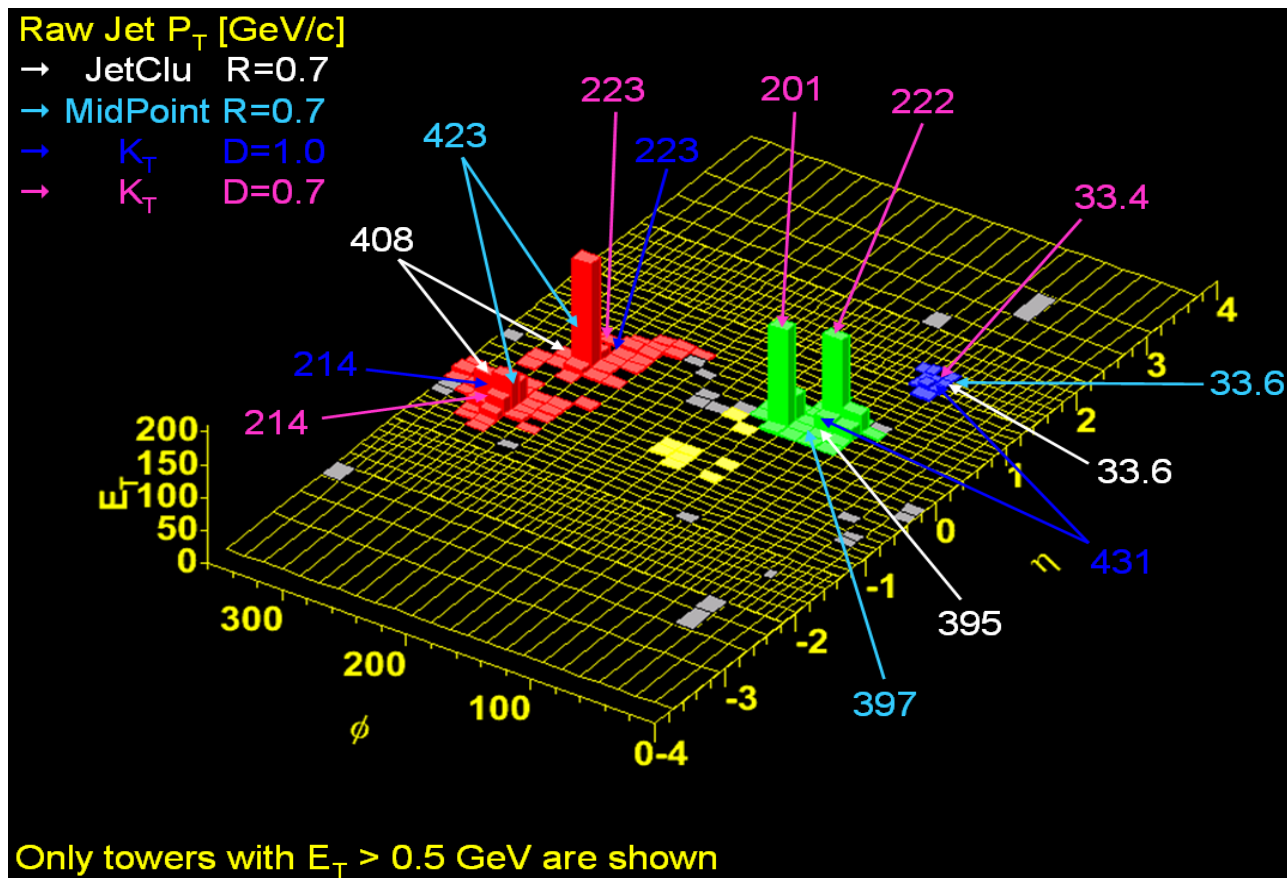
HERWIG: Evolve the parton-shower from Q^2 (*virtual photon invariant mass*) to Q_{min} ~ 1 GeV. Form color singlet clusters which "decay" into hadrons according to 2-particle phase space.

MLLA: Evolve the parton-shower from Q^2 (*virtual photon invariant mass*) to Q_{min} ~ 230 MeV. Assume that the charged particles behave the same as the partons with $N_{chg}/N_{parton} = 0.56$!

Q^2

γ^*

hadrons

Field-Feynman

5 GeV   1 GeV   200 MeV

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 112

Alternatives

"Experimental Jets"

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 113

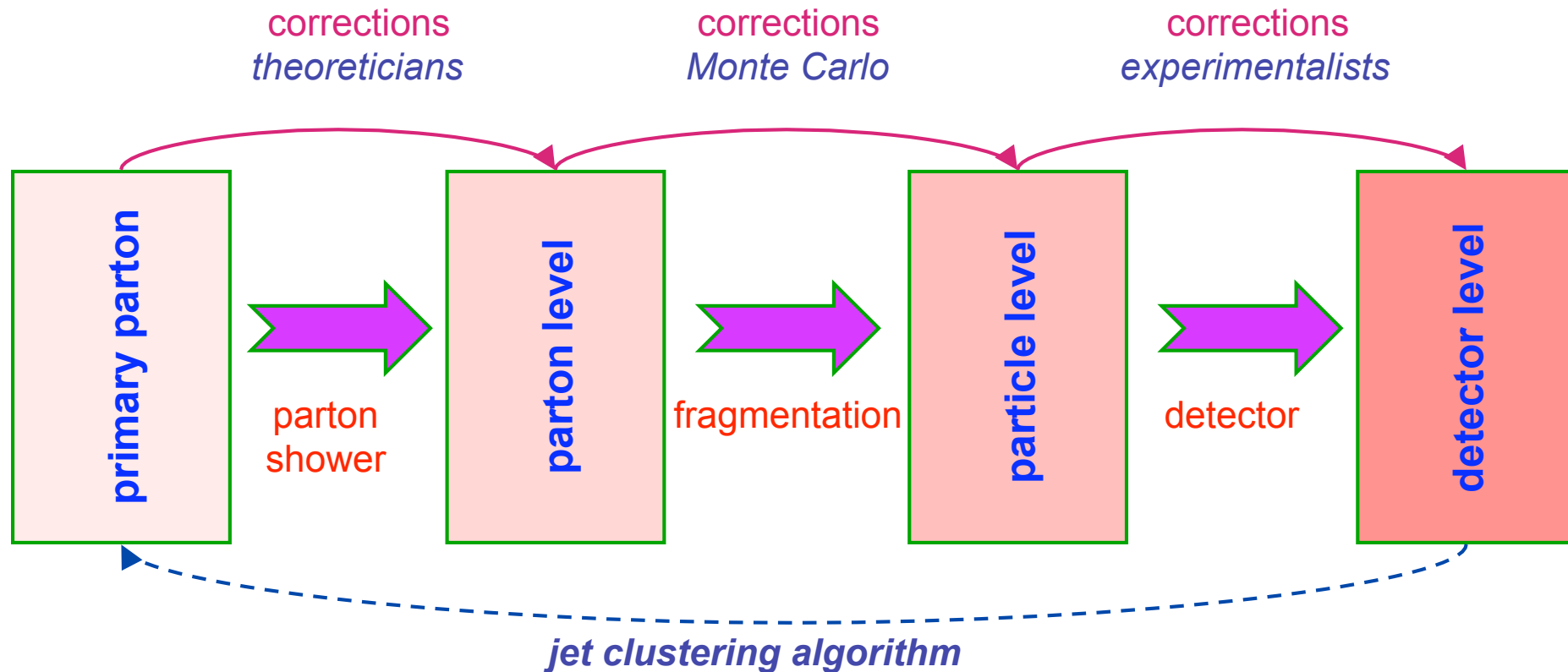# Jet reconstruction

❖ A jet of particles is a very complex object !!   Tevatron paper hep-ex/0005012

  ♦ *and we will see many of these particles*
 ♦ *jet algorithms can have both theoretical and experimental difficulties*

corrections   corrections   corrections
theoreticians   *Monte Carlo*   *experimentalists*

primary parton → parton shower → **parton level** → fragmentation → **particle level** → detector → **detector level**

jet clustering algorithm

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 114

Jet reconstruction

<u>Aim :</u> optimize the resolution on the reconstructed primary parton kinematics

♦ *use several jet clustering algorithms and study their features*
♦ *understand the influence of each step in the production of the the detected jet*
♦ *important steps : input object → clustering algorithm → calibration*

❖ Theoretical requirements

♦ *infra-red safety : jets insensitive to soft radiation*
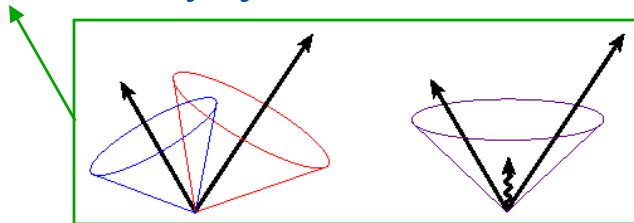
♦ *collinear safety : jets insensitive to collinear radiation*
♦ *boost invariance along the beam axis*
♦ *order dependence : equivalent at parton, particle and detector level*
♦ *straightforward implementation for perturbative calculations*

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 115
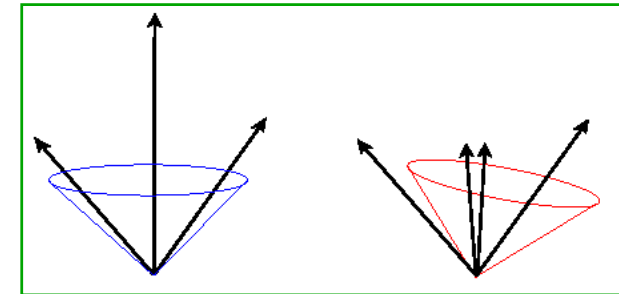
Jet reconstruction

❖ Experimental requirements
- ♦ *independent of granularity of the detector, energy responds or resolution (input objects)*
- ♦ *the algorithm should result in an optimized resolution*
- ♦ *stability versus the presence of pile-up events (more collisions in the same beam crossing)*
- ♦ *when used in the trigger it should aim for a minimal CPU time*
- ♦ *the algorithm should identify all physically interesting jets*
- ♦ *the jet definition should allow for a simple jet calibration method*

➢ All this makes a study of jets very complicated

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 116

Jet reconstruction

❖ Cone algorithms
- ♦ *construct a cone → always needs a seed to start with (seed finding mechanism)*
- ♦ *algorithm parameter : the opening angle of the cone ΔR (to be optimized)*
- ♦ *dealing with overlapping jets : splitting and merging procedures*
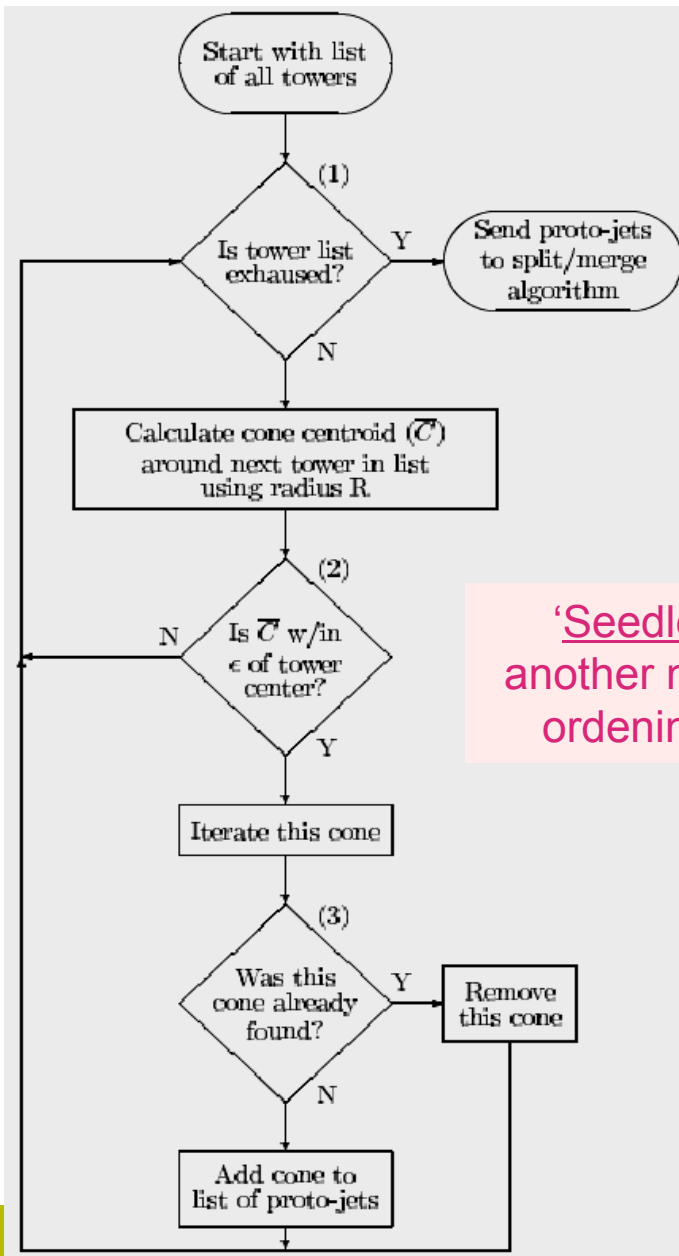- ♦ *Iterative, 'Seedless', 'Midpoint' Cone, …*

❖ Cluster algorithms
- ♦ *define a metric in the object space → 'distance' between each pair of objects*
- ♦ *closer to perturbative QCD phenomenology*
- ♦ *no problem with overlapping jets (no splitting and merging methods needed)*
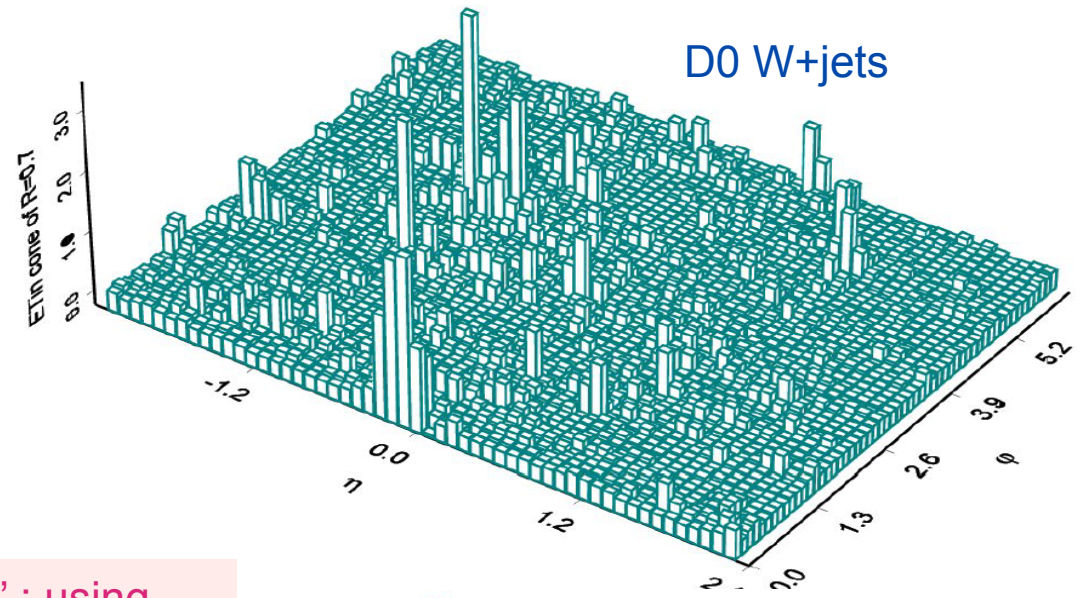- ♦ *algorithm parameters : ordening scheme and definition of metric*

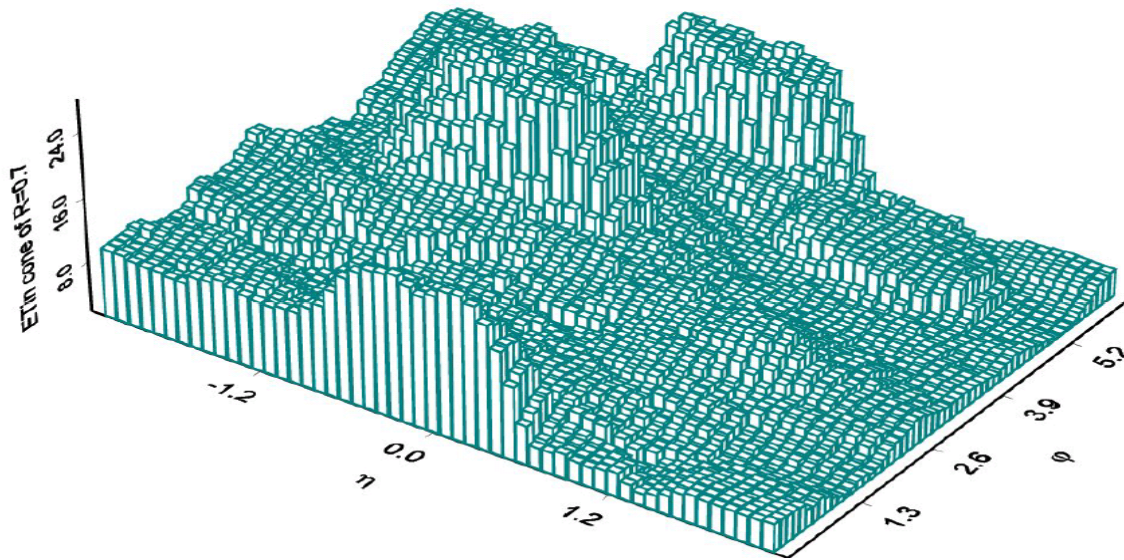❖ Continuous jet algorithms (*not available yet in CMS reconstruction*)
- ♦ *'new' algorithms based on statistical information (cfr. Minimum Variance Bound)*
  - F.V.Tkackov hep-ph/9901444 or hep-ph/9912415   L.Angelini *et al.* 'Jet Analysis by Deterministic Annealing'
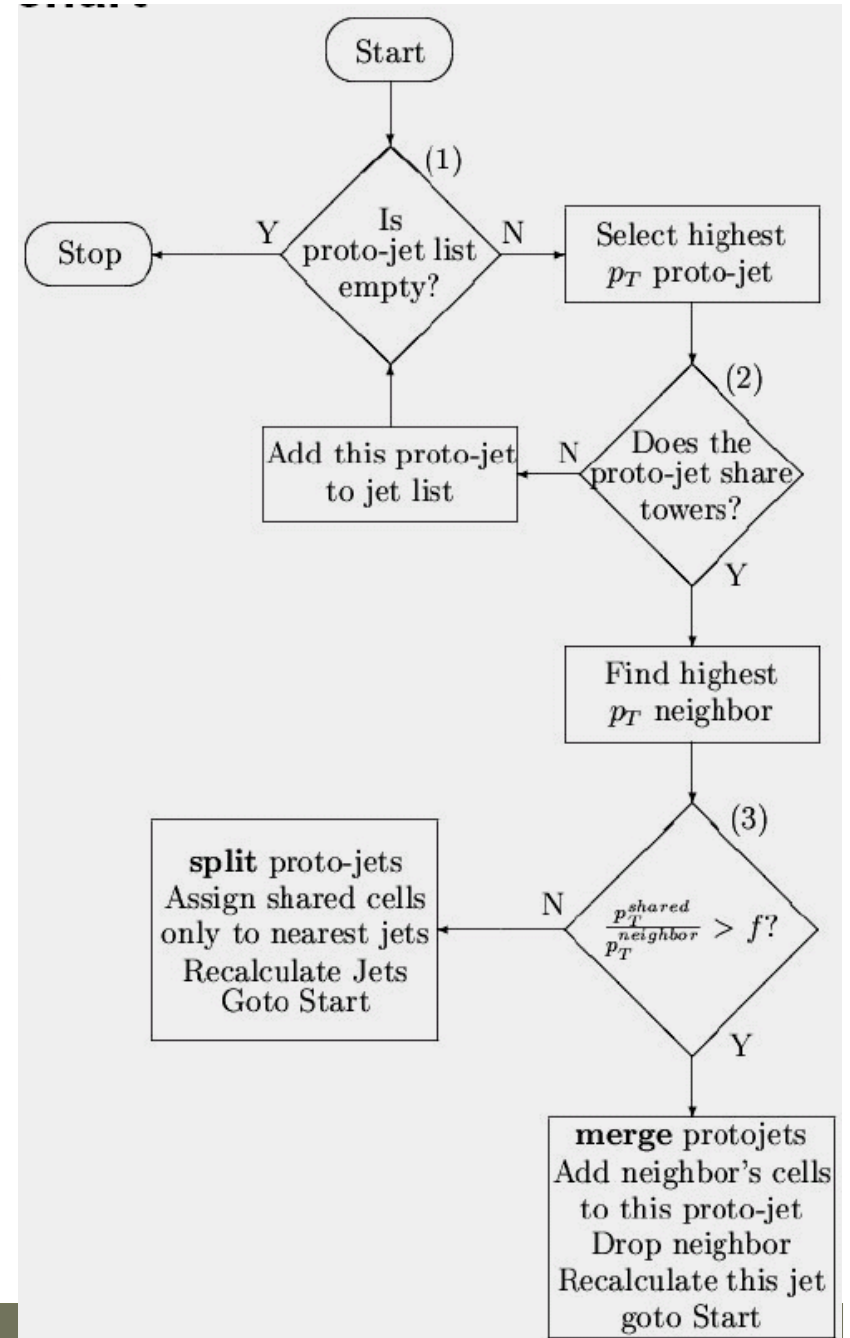- ♦ *some algorithms are based on 'entropy' maximization (cfr. statistical mechanics)*

Start with list of all towers

(1) Is tower list exhausted? → Y → Send proto-jets to split/merge algorithm

N

Calculate cone centroid ($\overline{C}$) around next tower in list using radius R.

(2) Is $\overline{C}$ w/in ϵ of tower center? → N

Y

Iterate this cone

(3) Was this cone already found? → Y → Remove this cone

N

Add cone to list of proto-jets
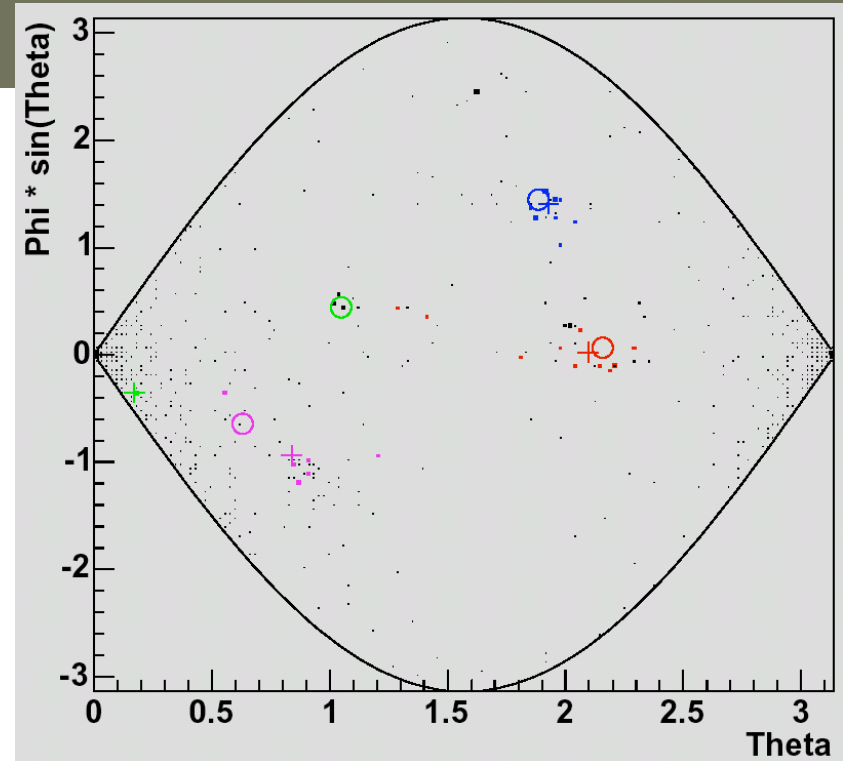
'Seedless' : using another mechanism for ordening the seeds

D0 W+jets

ΔR=0.7
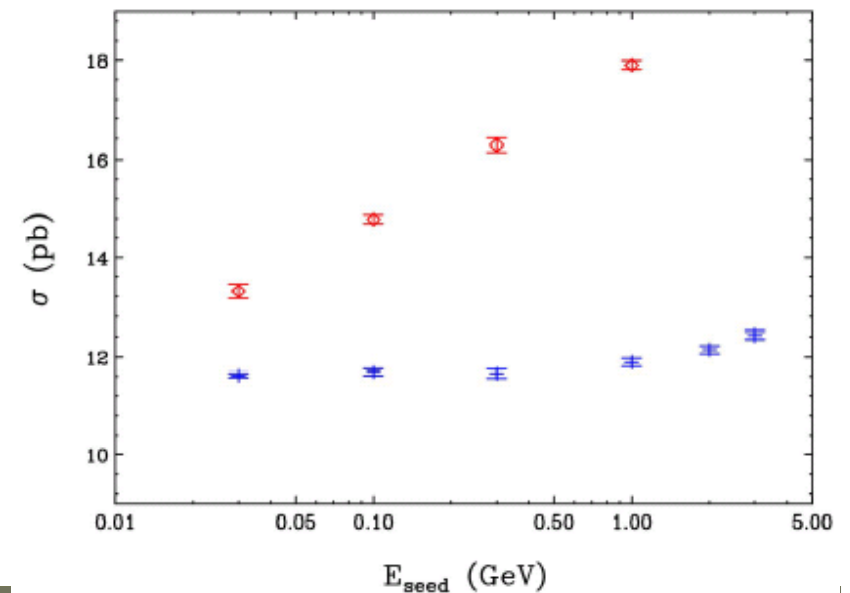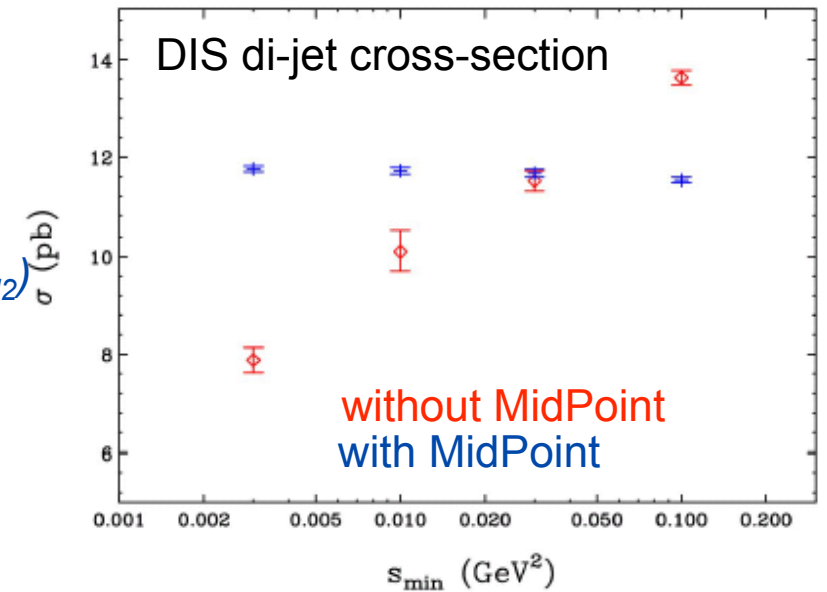total E$_T$ in cone
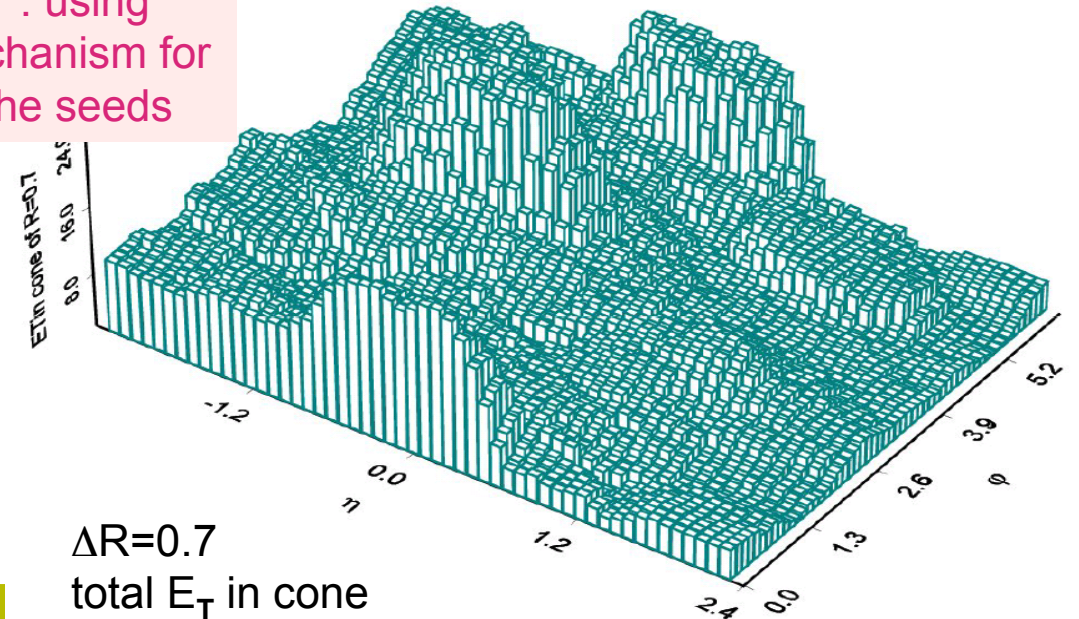
Merging and splitting procedure of proto-jets

♦ *seed-based algorithms with extra feature*
♦ *add 'MidPoints' to the list of proto-jets (eg. $p_{seed1}+p_{seed2}$)*
♦ *approximates seed-less algorithms but does not uses all possible towers as seeds*
♦ *improvement when factorizing the procedure :*

Generate E_T ordered list of towers

Find protojets around towers with E_T > threshold

Generate midpoint list from protojets

Find protojets around midpoints

Goto split/merge
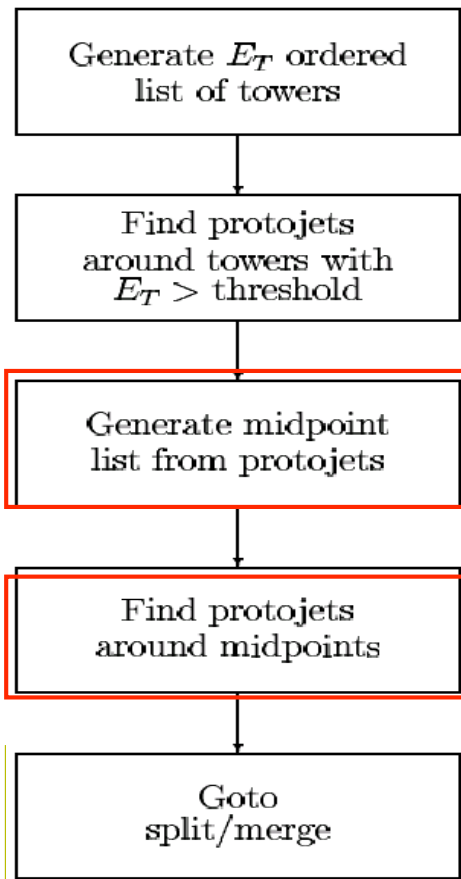
♦ *use R/2 to search cone*
♦ *use R to calculate kinematics (MidPointFix)*

hep-ph/0111434

DIS di-jet cross-section

σ (pb)

without MidPoint
with MidPoint

s_{min} (GeV^2)

σ (pb)

E_{seed} (GeV)

For each object and pair of objects:

order all d_{ii} and d_{ij}:

$$d_{iB} = k_{T,i}^2$$
$$d_{ij} \underset{*}{=} \min(k_{T,i}^2, k_{T,i}^2) \cdot \Delta R_{ij}^2$$

If $\mathbf{d_{min}=d_{ij}}$

$\Rightarrow$ merge objects

If $\mathbf{d_{min}=d_{iB}}$

$\Rightarrow$ merge with beam-jet

or add to list of jets

❖ Two modes of operation
  ▪ Inclusive mode : everything into jets
  ▪ Exclusive mode : 2 beam jets in each direction + hard jets   hep-ph/0210022
❖ Different jet resolution variables
  ▪ angular scheme, $\underline{\Delta R\ scheme}$ (*), QCD emission scheme
❖ Different recombination schemes
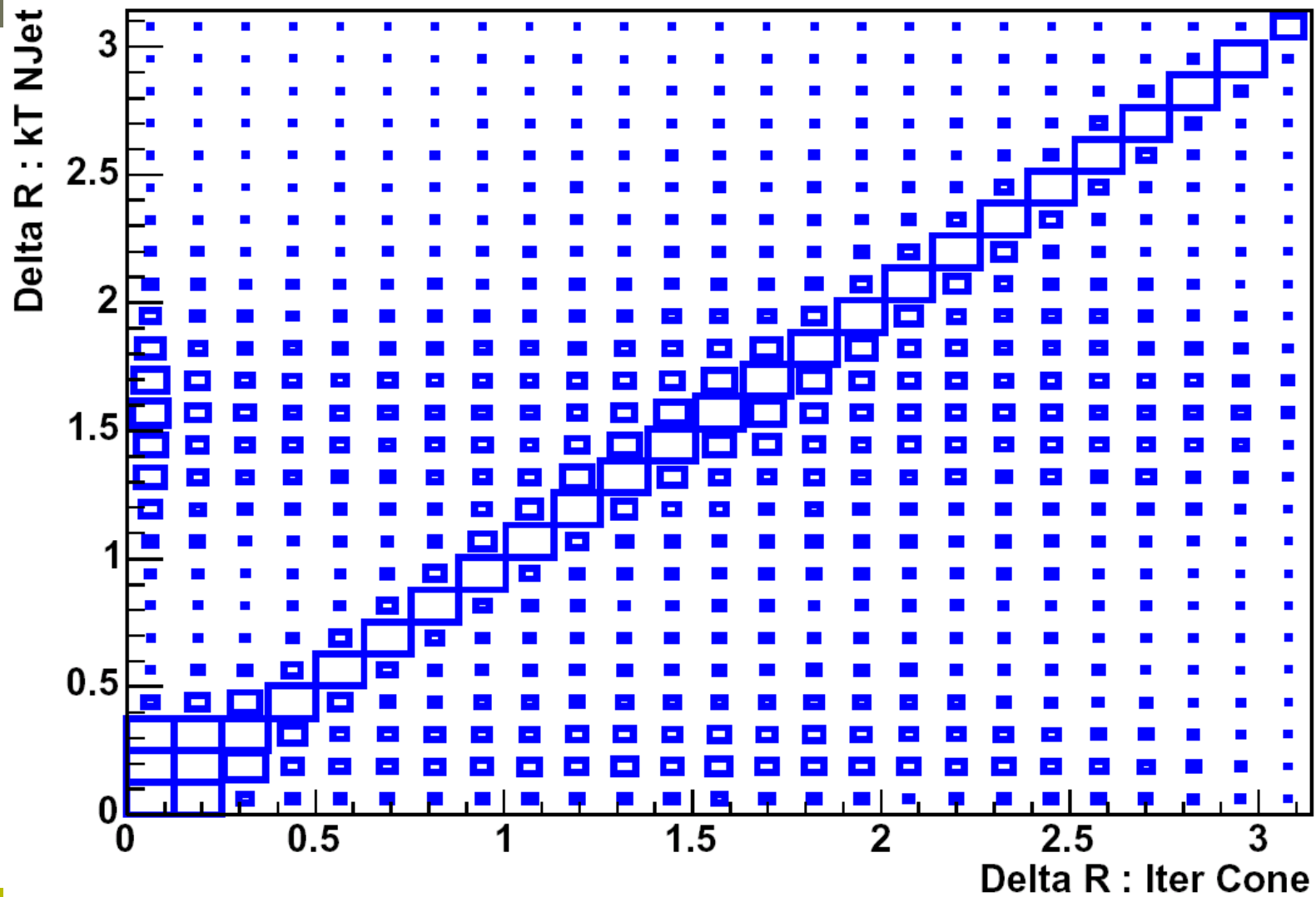 ▪ $\underline{E\ scheme}$, p_t scheme, p_t^2 scheme, E_t scheme, E_t^2 scheme

❖ Features of the k_T algorithms :
  ▪ *small angle behaviour is infra-red safe and collinear safe*
  ▪ *no overlapping jets and not influences by seed towers*
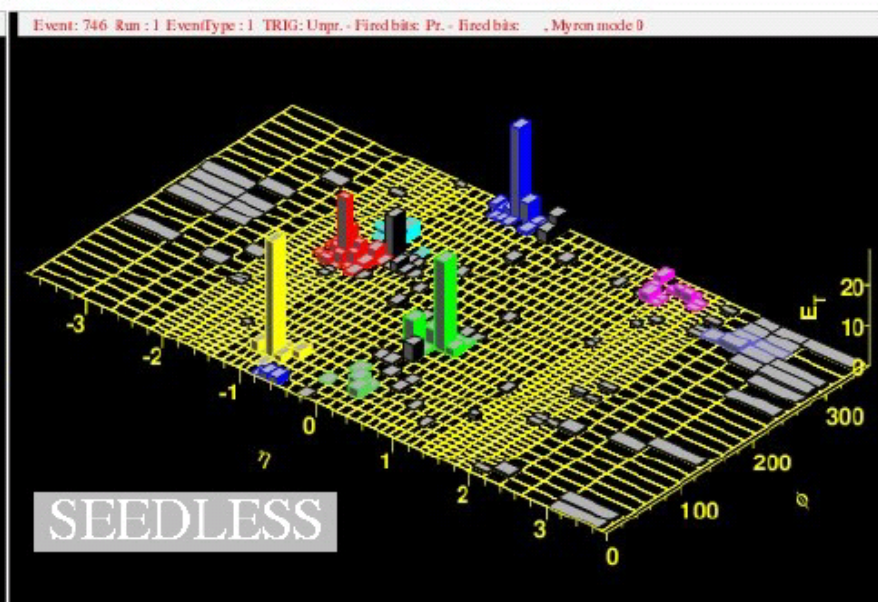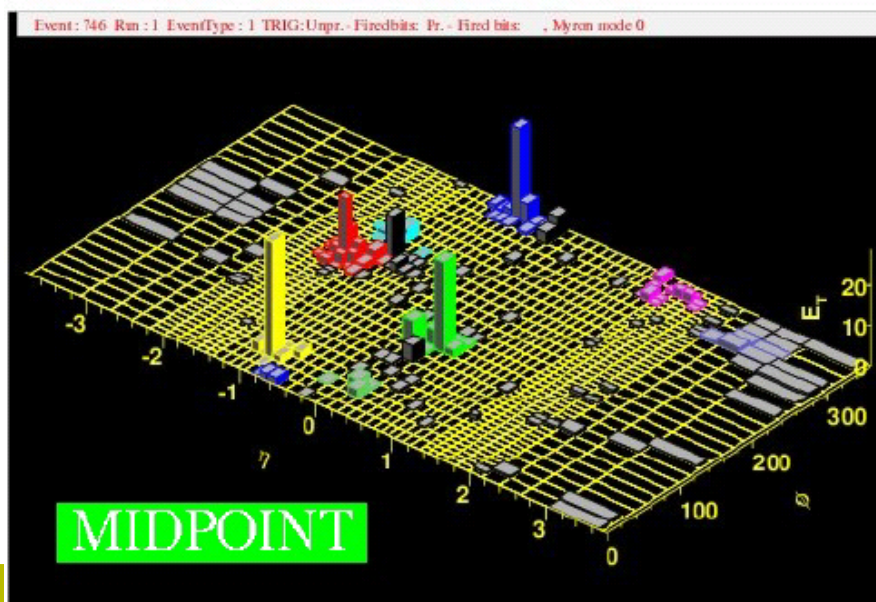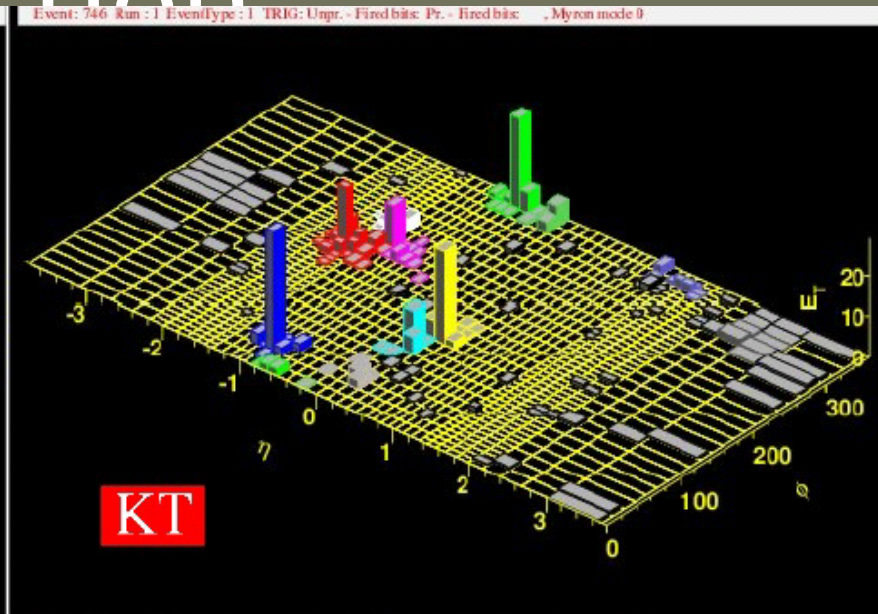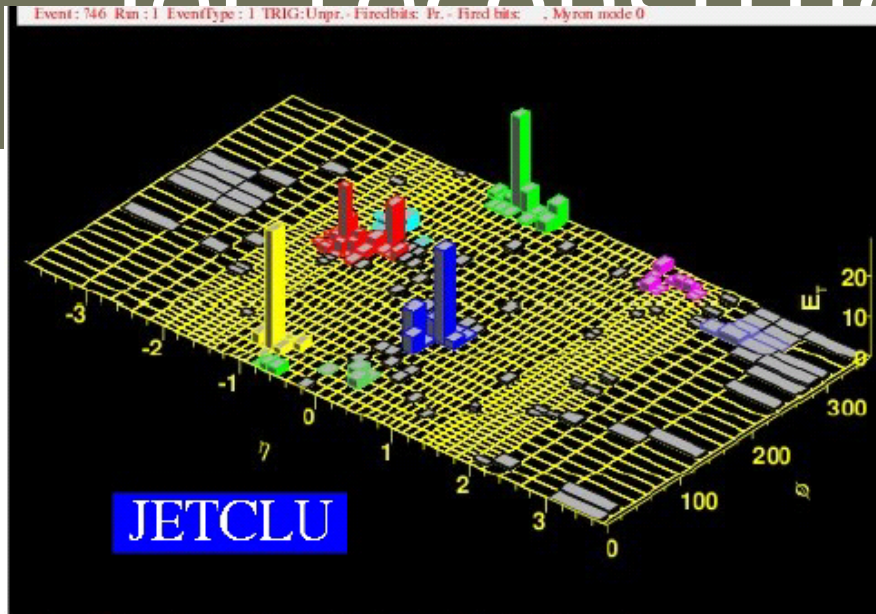 ▪ *closer to perturbative QCD*

Jet reconstruction
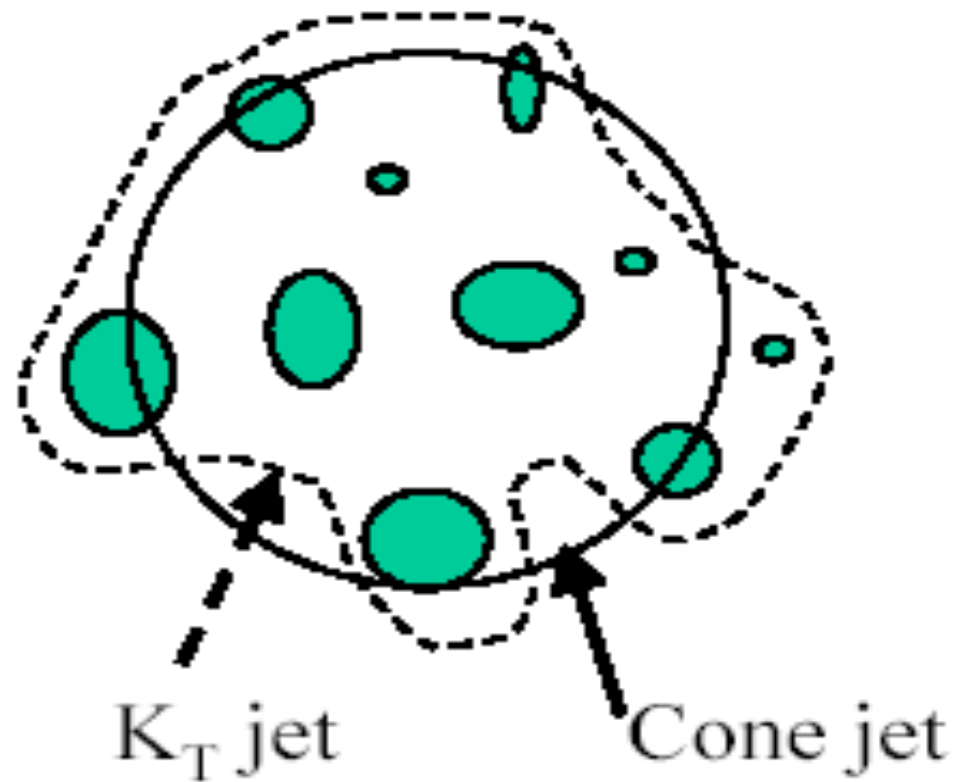
K_T jet   Cone jet

Jet reconstruction

Data-analysis at colliders - Prof. J. D'Hondt (VUB)

Master Level Course 2007-2008

Pag. 123

Jet reconstruction

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 124

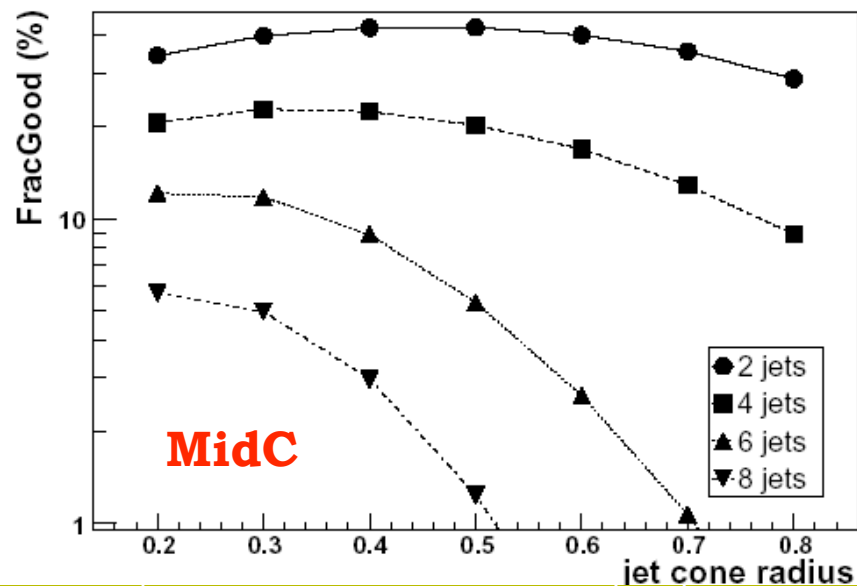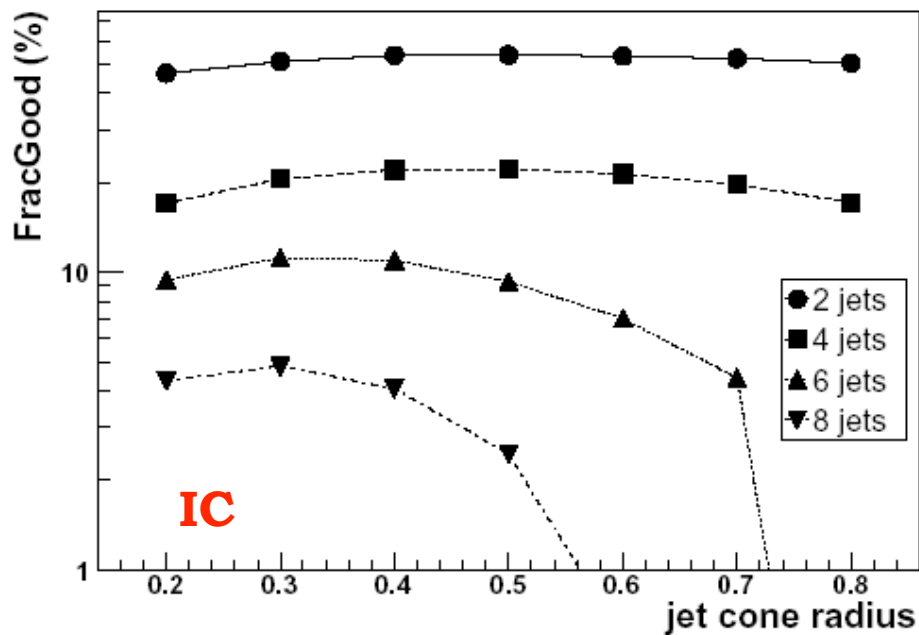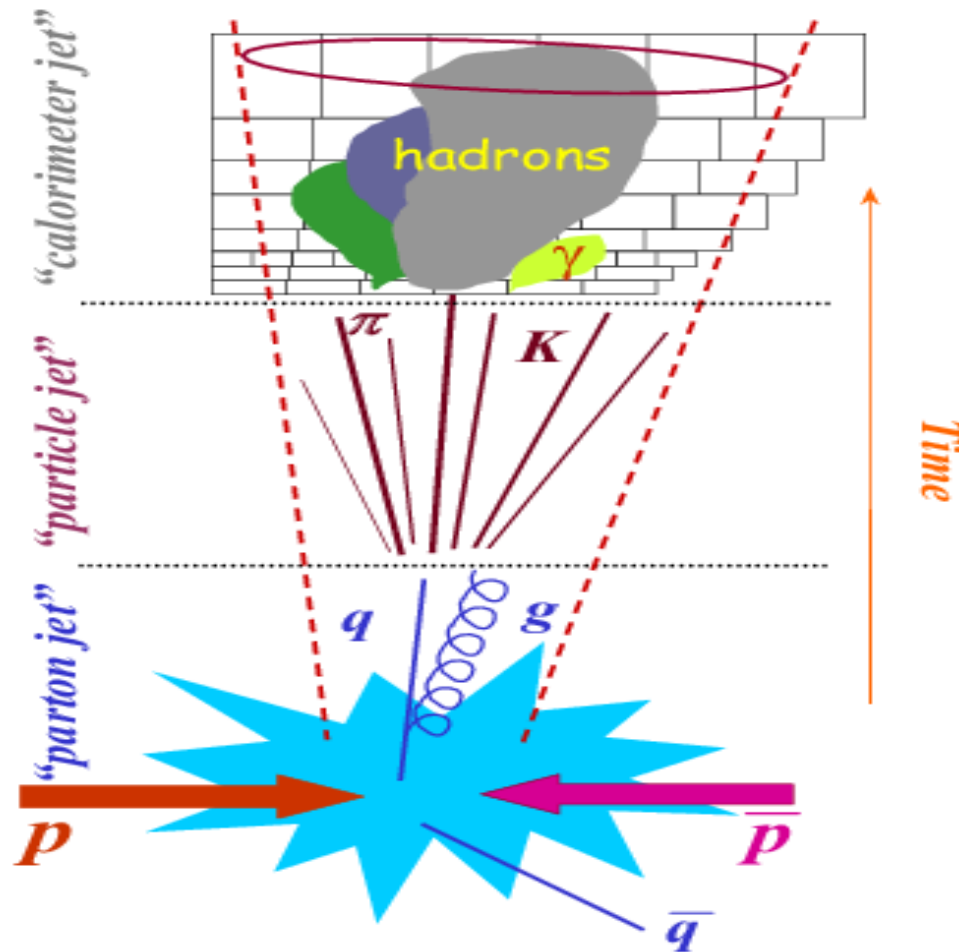# Jet reconstruction

For top quark physics we would like to reconstruct the four-momentum of the initial or primary partons. Hence the reconstructed jets have to match the parton in angle and energy. This can be quantified and optimized for the parameters of the algorithm.
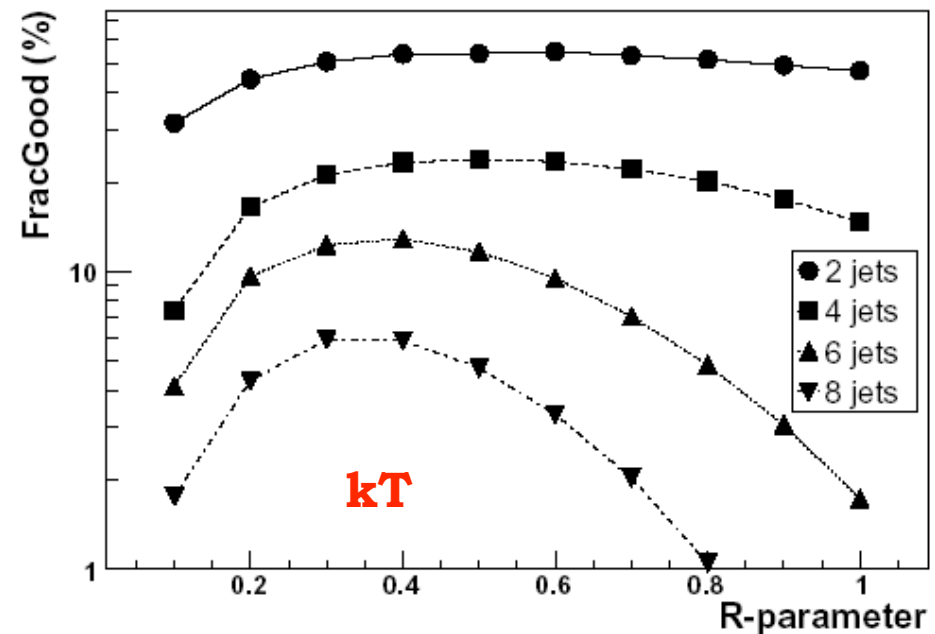
Jet calibration

Jet Calibration

{ <u>Aim :</u> know the absolute jet energy to the level of at least 1%

} Complex issue (needed in kinematic fit)

- ♦ *<u>physics</u> : ISR/FSR, fragmentation, underlying event, jet algorithm, …*
- ♦ *<u>detector</u> : calorimeter responds, non-linear effects, …*
- ♦ *all of them have to be understood at the 0.2% level to reach the 1% global goal*
- ♦ *the method has to cover a wide range of jet energies and directions*

Tevatron based method :

$$E = [E^{raw} \cdot f_{rel} - MI] \cdot f_{abs} - UE + JAC$$

- ♦ *f_{rel} : relative energy scale correction (detector related corrections)*
- ♦ *MI : multiple interactions in the event*
- ♦ *f_{abs} : absolute jet energy scale (to be differentiated in the observable space)*
- ♦ *UE : energy of the underlying event*
- ♦ *JAC : corrections dependent on the clustering algorithm (eg. out of cone energy)*

All these parameters must be estimated with real data (*control samples*)

- ♦ *$W \rightarrow jj$ ($\delta m_W \sim 30$ MeV) in single-lepton tt events*
- ♦ *Z + jets ($Z \rightarrow e^+e^-$ or $Z \rightarrow \mu^+\mu^-$) also good for b-jet calibration*
- ♦ *γ + jets : large systematics of background from hadronic jets misidentified as photons*
- ♦ *…*

{ <u>Aim :</u> know the absolute jet energy to the level of at least 1% directly from data

} Light-flavoured jets in W→jj

- ♦ *use the W decays in single-lepton tt events (large purity of sample)*
- ♦ *compare generated parton energy with reconstructed jet energy*
- ♦ *apply the constraint $m_W = m_{jj}$*
- ♦ *main systematic uncertainty : QCD final state radiation (optimize the jet algorithm)*
- ♦ *momentum range : ~50 GeV < p_T < ~200 GeV*
- ♦ *take ½ of the events to estimate the jet corrections, and apply them on the other ½*

Flavour dependent measurement in Z+jets

- ♦ *major work : design some efficient event selection for this process*
- ♦ *typical flavour mixture :*
  - ▪ *28% gluon*
  - ▪ *54% light-quark jets*
  - ▪ *12% c-quark jets*
  - ▪ *6% b-quark jets (between 3000 and 5000 jets after selection depending on |η|)*
    (decreasing to around 500 for high p_T jets above 120 GeV)
- ♦ *calculate the jet corrections as a function of the b-tag probability*
  (using the same b-tag method as in the physics case analysis !!)
- ♦ *momentum range : ~20 GeV < p_T < 200 GeV*
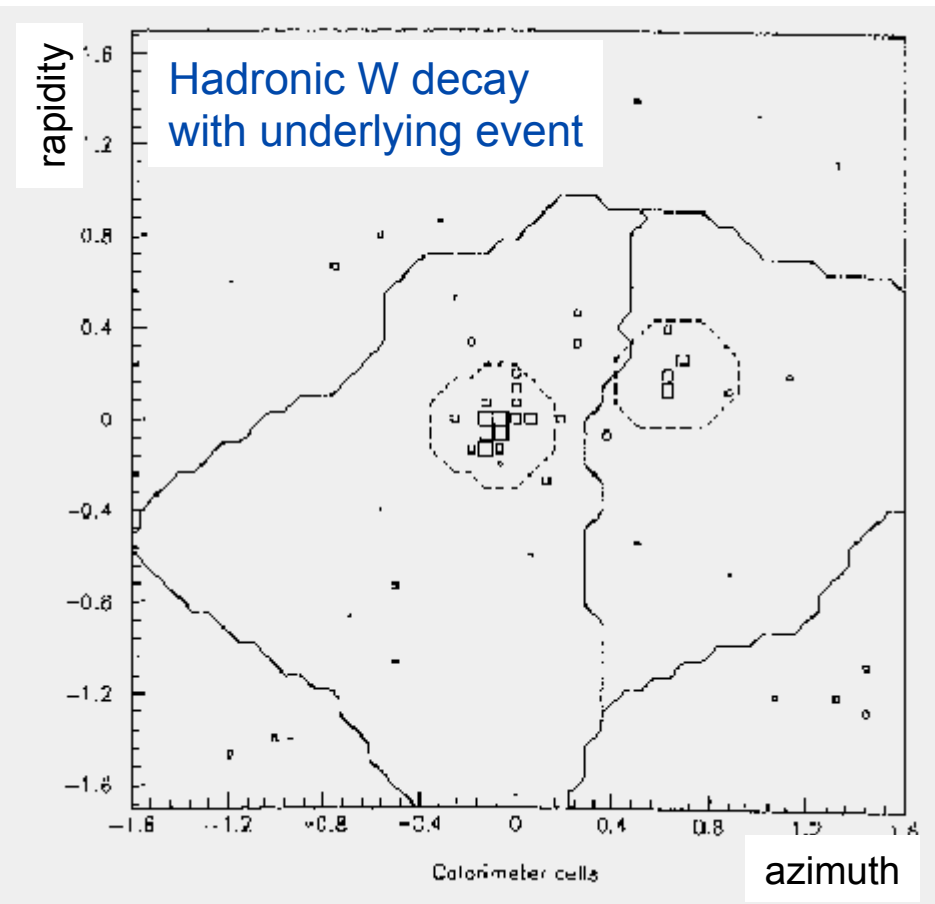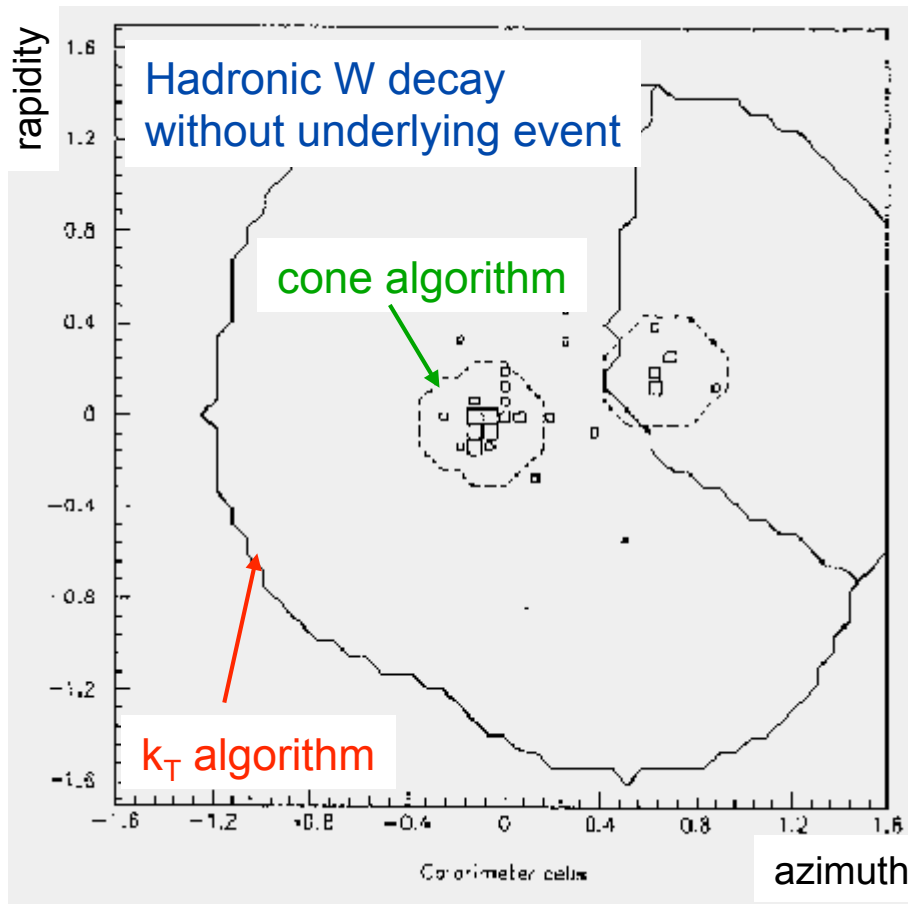
Need both data samples to reduce possible systematic effects !!

Influence of Pile-up

Some jet algorithm will be less sensitive to the underlying event
⇨ one has to optimize the algorithm parameters
⇨ study the behaviour of the algorithms (try to differentiate !!)

rapidity

Hadronic W decay
without underlying event

cone algorithm

k_T algorithm

Calorimeter cells

azimuth

rapidity

Hadronic W decay
with underlying event

Calorimeter cells

azimuth

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 129

Influence of Pile-up

Some jet algorithm will be less sensitive to the pile-up events
- one has to optimize the algorithm parameters
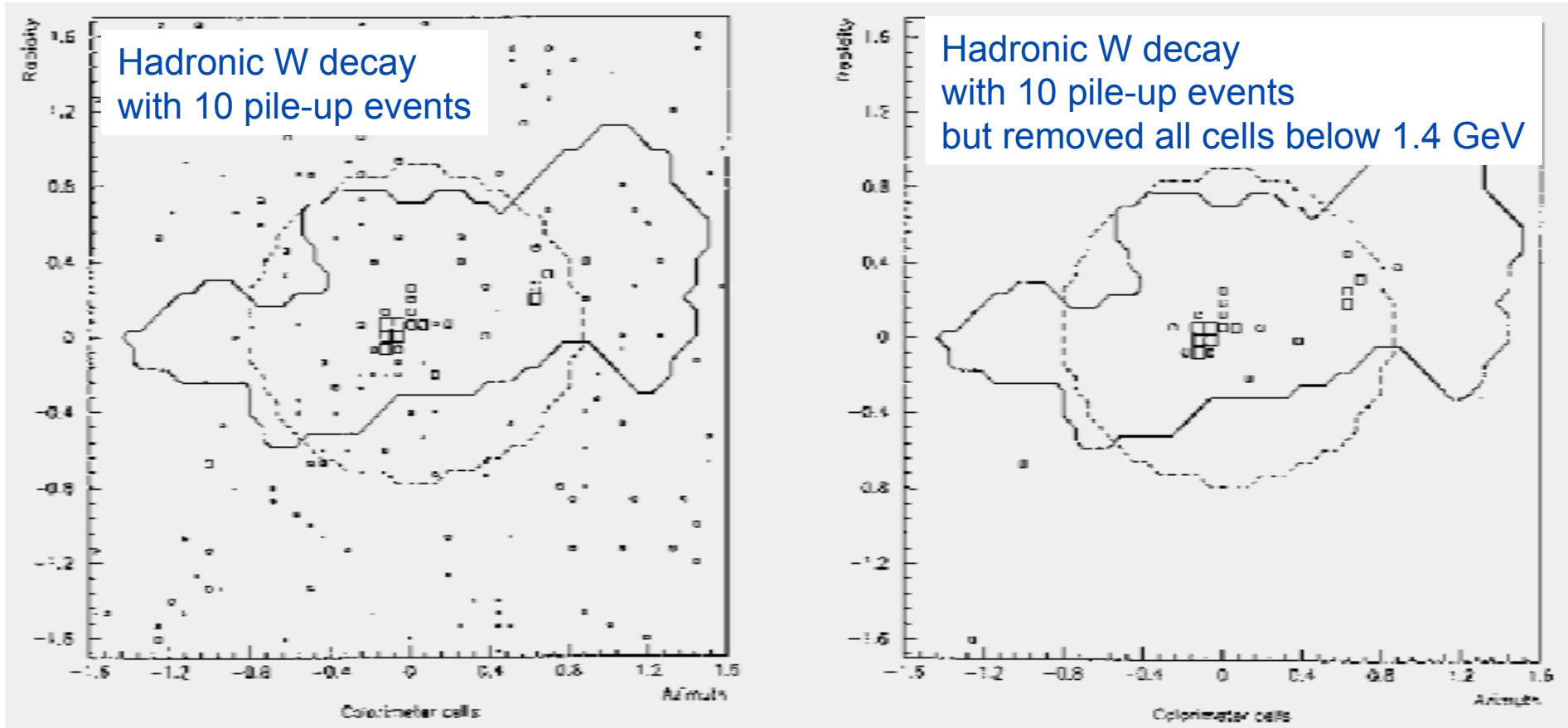- study the behaviour of the algorithms (try to differentiate !!)

Hadronic W decay
with 10 pile-up events

Hadronic W decay
with 10 pile-up events
but removed all cells below 1.4 GeV

- removing the calorimeter cells below a fixed energy helps (to be checked for CMS)
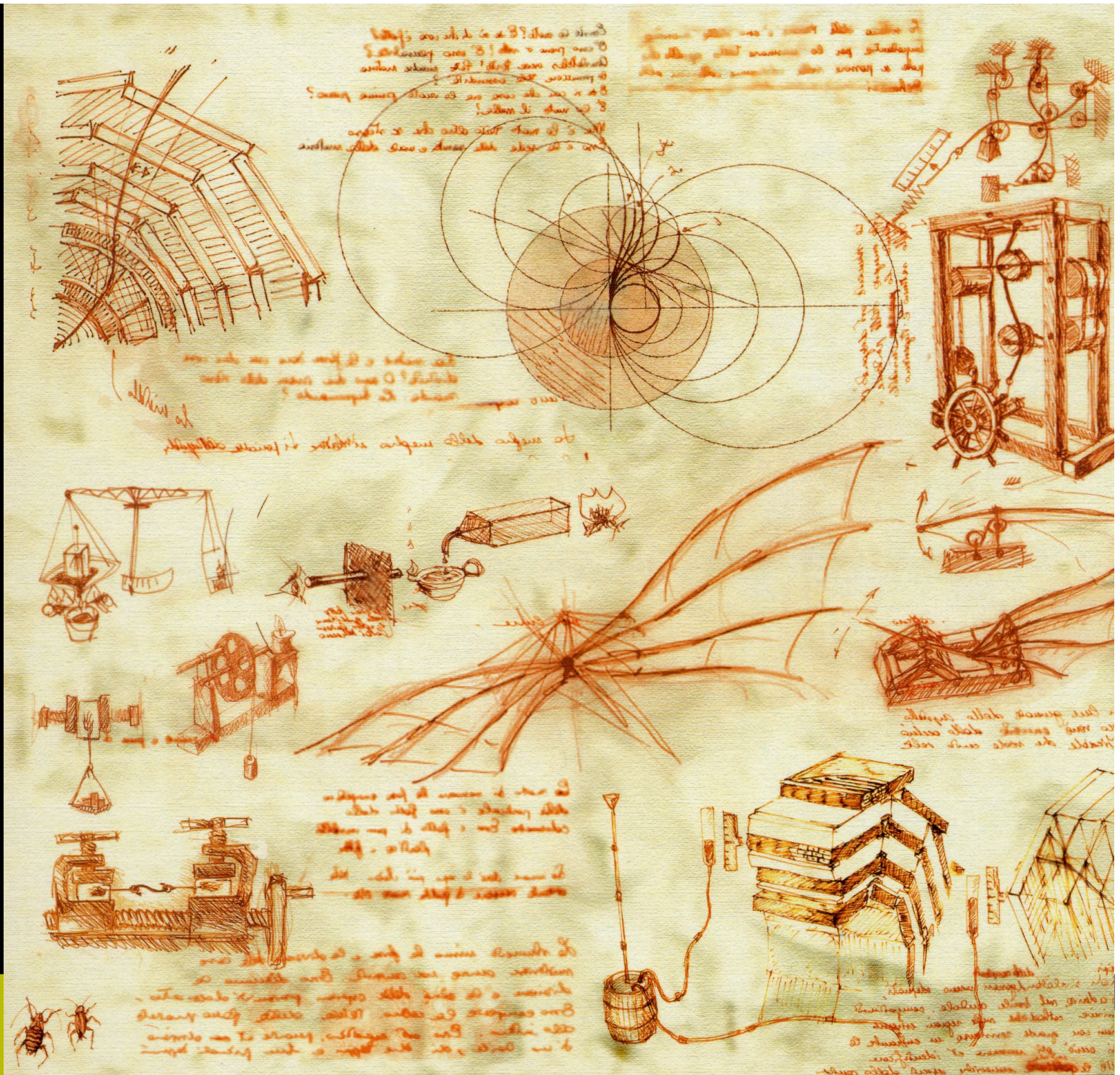- this removal of energy can be used when calculating the jet momentum

Properties of b-quark jets

B-tagging algorithms

Summary of second part

- Two important concepts of reconstruction tools to be used to analyse the collision data at for example the LHC and CMS

End
of
Part II

The third part

Advanced analysis techniques
- General

Summary of third part

- Outreach

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 136

Grand summary

- Outreach

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 137

Final remarks

- Outreach

Data-analysis at colliders - Prof. J. D'Hondt (VUB)
Master Level Course 2007-2008

Pag. 138