Introduction to Physics Computing

LHC Grid Computing School

Sang Un Ahn
sahn@kisti.re.kr

About myself

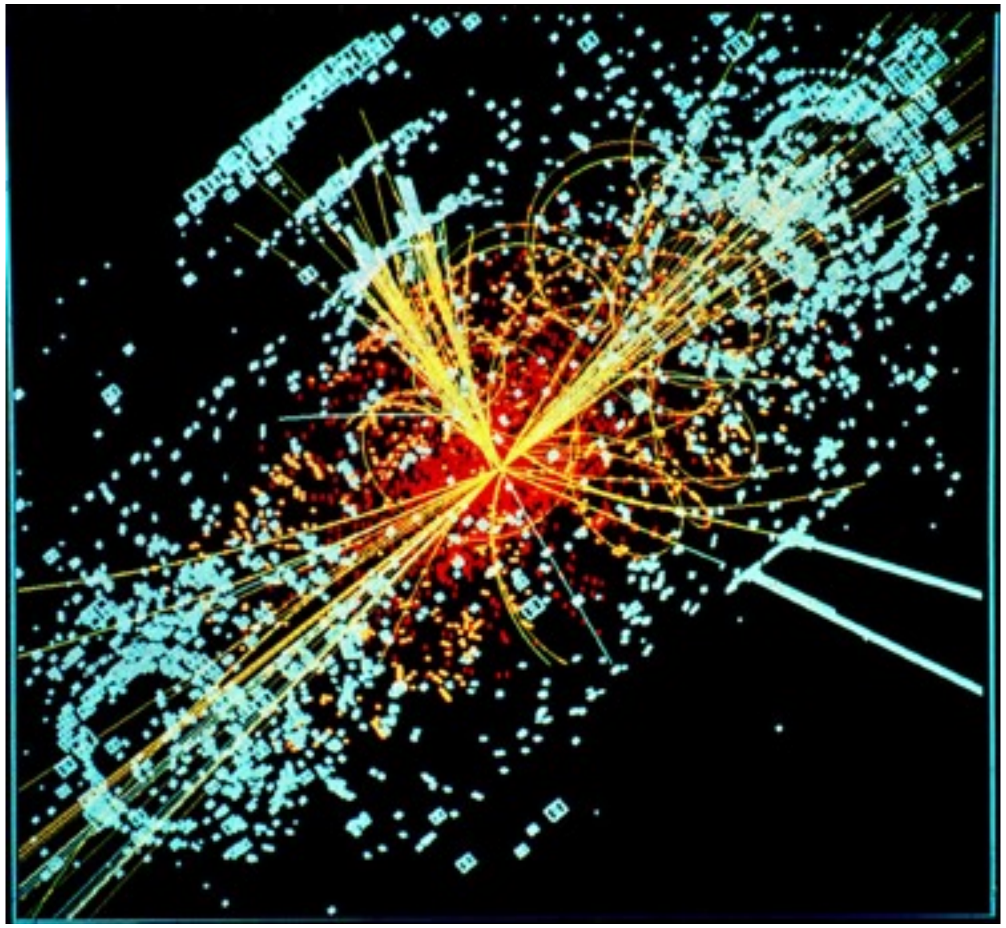
- I am an experimental particle physicist
  - PhD in ALICE experiment
  - Heavy Quarkonium (Upsilon) measurement via muons
  - Development of Data Quality Assurance for Muon Trigger system
- I am a system administrator
  - RHCSA/RHCE
  - KR-KISTI-GSDC (Tier-1) Site Administrator

Outline of the lecture

- Introduction
- Various aspects of Physics Computing:
  - Event Filtering
  - Calibration and alignment
  - Event Reconstruction
  - Event Simulation
  - Physics Analysis
  - Data Flow and Computing Resources

Technical Challenges at LHC

- Everything in LHC computing is connected to processing such data !!



Technical Challenges at LHC

- Very high (design) event rate: 40 MHz
- Large event size: O(1) MB
- Large background of uninteresting events
- Large background in each event
    - many interactions in each beam crossing
    - pile-up from adjacent beam crossing
    - many low-momentum particles

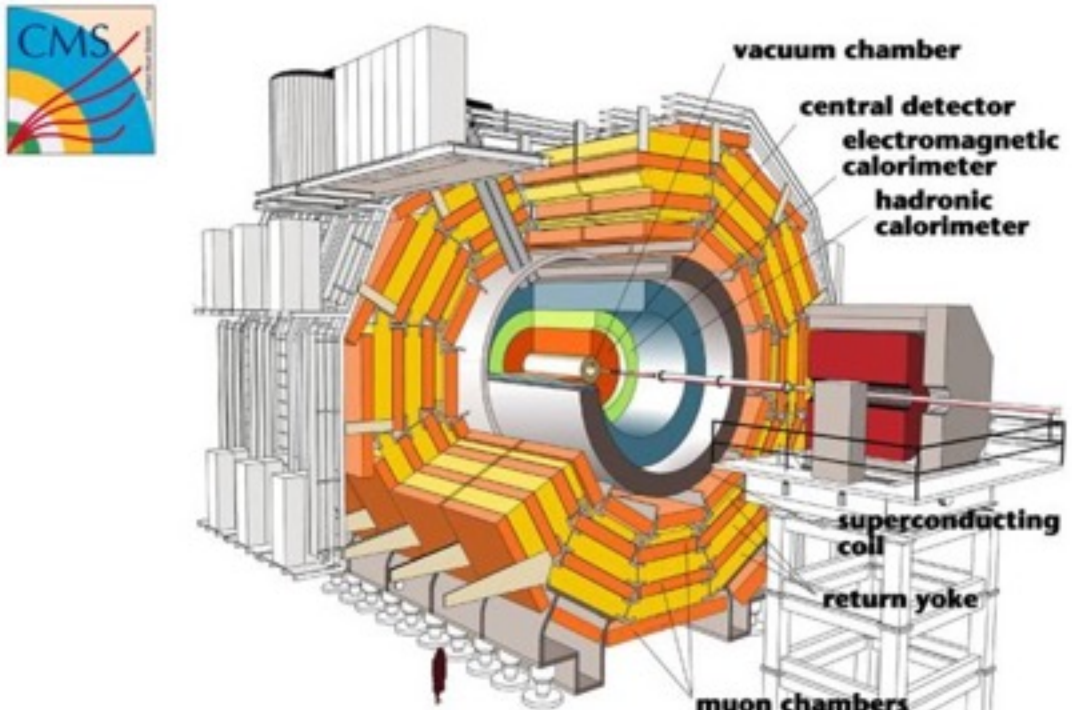
Technical Challenges at LHC

- Large number of physicists doing analysis
    - ATLAS and CMS experiments at the LHC: both consist of 170-180 institutes in about 40 countires
    - Distribution of data and programs
    - Bookkeeping is crucial
- High pressure, competitive spirit
    - Important discoveries to be (and have been) made
    - Computing has to be as fast as possible

What is Physics Computing?

- Yearly input: a few petabytes of data
- Yearly output: a few hundred physics papers
- Data reduction factor of 10^7 to 10^8 !!
- How is it done?
- Will try to answer this question in lecture

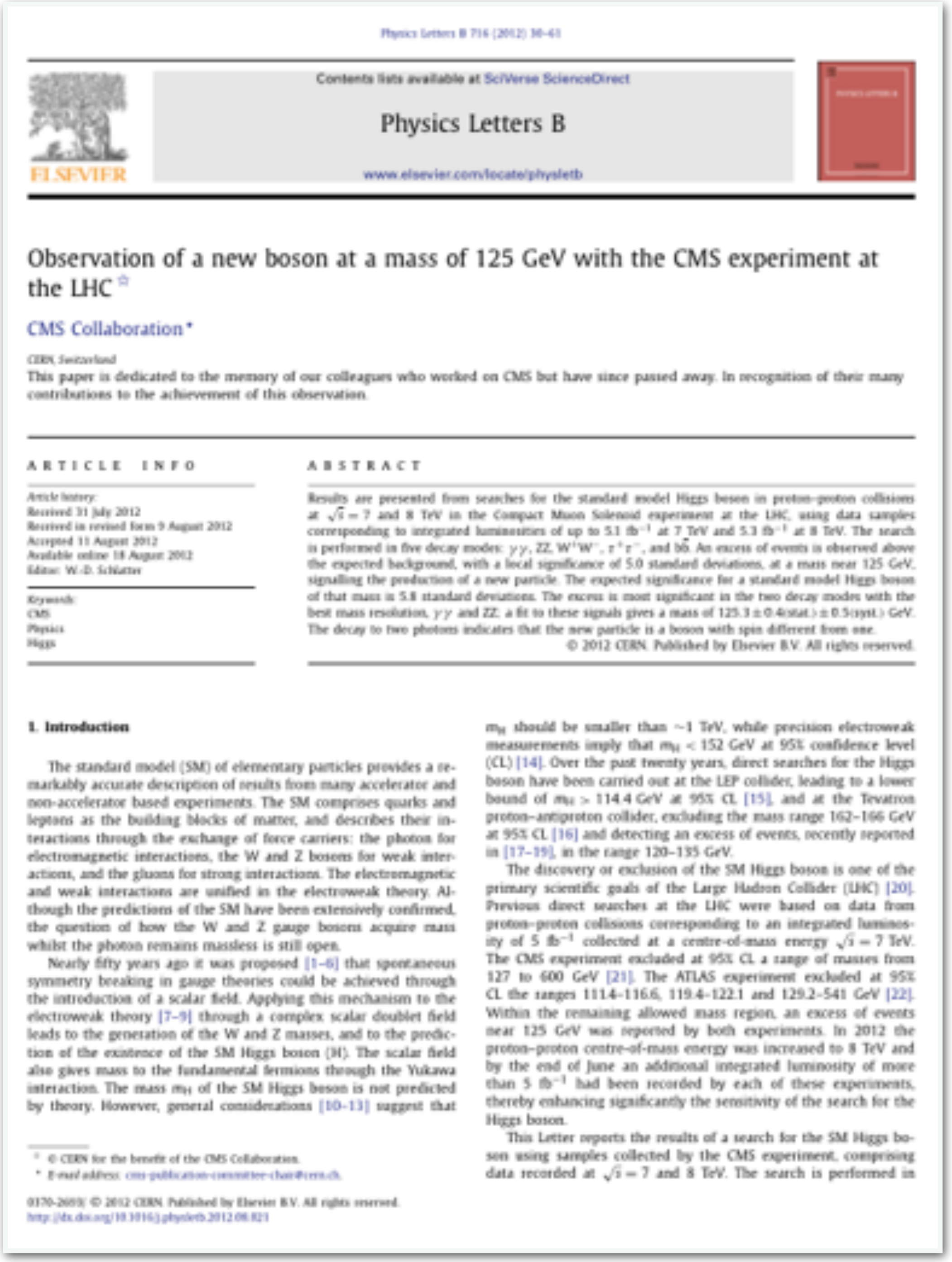
It’s simple ··· is it?



Detector characteristics

Width: 22m
Diameter: 15m
Weight: 14'500t

```
Paper paper15
Data higgsdata
…
paper15=make_paper(higgsdata)
…
```



Physics Letters B 716 (2012) 30–61

Contents lists available at SciVerse ScienceDirect

Physics Letters B

www.elsevier.com/locate/physletb

Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC ☆

CMS Collaboration *

CERN, Switzerland

This paper is dedicated to the memory of our colleagues who worked on CMS but have since passed away. In recognition of their many contributions to the achievement of this observation.

A R T I C L E I N F O

Article history:
Received 31 July 2012
Received in revised form 9 August 2012
Accepted 11 August 2012
Available online 18 August 2012
Editor: W.-D. Schlatter

Keywords:
CMS
Physics
Higgs

A B S T R A C T

Results are presented from searches for the standard model Higgs boson in proton–proton collisions at $\sqrt{s} = 7$ and 8 TeV in the Compact Muon Solenoid experiment at the LHC, using data samples corresponding to integrated luminosities of up to 5.1 fb^{-1} at 7 TeV and 5.3 fb^{-1} at 8 TeV. The search is performed in five decay modes: $\gamma\gamma$, ZZ, W^+W^-, $\tau^+\tau^-$, and $b\bar{b}$. An excess of events is observed above the expected background, with a local significance of 5.0 standard deviations, at a mass near 125 GeV, signalling the production of a new particle. The expected significance for a standard model Higgs boson of that mass is 5.8 standard deviations. The excess is most significant in the two decay modes with the best mass resolution, $\gamma\gamma$ and ZZ; a fit to these signals gives a mass of $125.3 \pm 0.4(\text{stat.}) \pm 0.5(\text{syst.})$ GeV. The decay to two photons indicates that the new particle is a boson with spin different from one.

1. Introduction

The standard model (SM) of elementary particles provides a remarkably accurate description of results from many accelerator and non-accelerator based experiments. The SM comprises quarks and leptons as the building blocks of matter, and describes their interactions through the exchange of force carriers: the photon for electromagnetic interactions, the W and Z bosons for weak interactions, and the gluons for strong interactions. The electromagnetic and weak interactions are unified in the electroweak theory. Although the predictions of the SM have been extensively confirmed, the question of how the W and Z gauge bosons acquire mass whilst the photon remains massless is still open.

Nearly fifty years ago it was proposed [1–6] that spontaneous symmetry breaking in gauge theories could be achieved through the introduction of a scalar field. Applying this mechanism to the electroweak theory [7–9] through a complex scalar doublet field leads to the generation of the W and Z masses, and to the prediction of the existence of the SM Higgs boson (H). The scalar field also gives mass to the fundamental fermions through the Yukawa interaction. The mass m_H of the SM Higgs boson is not predicted by theory. However, general considerations [10–13] suggest that m_H should be smaller than ~1 TeV, while precision electroweak measurements imply that $m_H < 152$ GeV at 95% confidence level (CL) [14]. Over the past twenty years, direct searches for the Higgs boson have been carried out at the LEP collider, leading to a lower bound of $m_H > 114.4$ GeV at 95% CL [15], and at the Tevatron proton–antiproton collider, excluding the mass range 162–166 GeV at 95% CL [16] and detecting an excess of events, recently reported in [17–19], in the range 120–135 GeV.

The discovery or exclusion of the SM Higgs boson is one of the primary scientific goals of the Large Hadron Collider (LHC) [20]. Previous direct searches at the LHC were based on data from proton–proton collisions corresponding to an integrated luminosity of 5 fb^{-1} collected at a centre-of-mass energy $\sqrt{s} = 7$ TeV. The CMS experiment excluded at 95% CL a range of masses from 127 to 600 GeV [21]. The ATLAS experiment excluded at 95% CL the ranges 111.4–116.6, 119.4–122.1 and 129.2–541 GeV [22]. Within the remaining allowed mass region, an excess of events near 125 GeV was reported by both experiments. In 2012 the proton–proton centre-of-mass energy was increased to 8 TeV and by the end of June an additional integrated luminosity of more than 5 fb^{-1} had been recorded by each of these experiments, thereby enhancing significantly the sensitivity of the search for the Higgs boson.

This Letter reports the results of a search for the SM Higgs boson using samples collected by the CMS experiment, comprising data recorded at $\sqrt{s} = 7$ and 8 TeV. The search is performed in

---

☆ © CERN for the benefit of the CMS Collaboration.

* *E-mail address:* cms-publication-committee-chair@cern.ch.

0370-2693/ © 2012 CERN. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.physletb.2012.08.021

Actually, at LHC we need…

- Millions of lines of code (C++, Python, Perl, …)
- Hundreds of neural networks
- Large infrastructure
    - Customized hardware
    - PC farms
    - Database and storage systems
    - Distributed analysis facilities
    - The grid

What happens to the data?

- Event filtering, tagging and storage
- Calibration, alignment
- Event reconstruction
- Storage
- Event simulation
- Physics analyses

Step by step

- Each step involves some data reduction
  - data are discarded (online)
  - data are compressed (offline)
- In each step the data get closer to be interpretable in physical terms
- Some steps are repeated many times until the output is satisfactory (offline processing)

Online vs Offline computing

- Online
  - In real time, fast!
  - Decisions are irreversible
  - Data cannot be recovered
- Offline
  - From almost real time to long delays
  - Decisions can be reconsidered
  - Data can be reprocessed

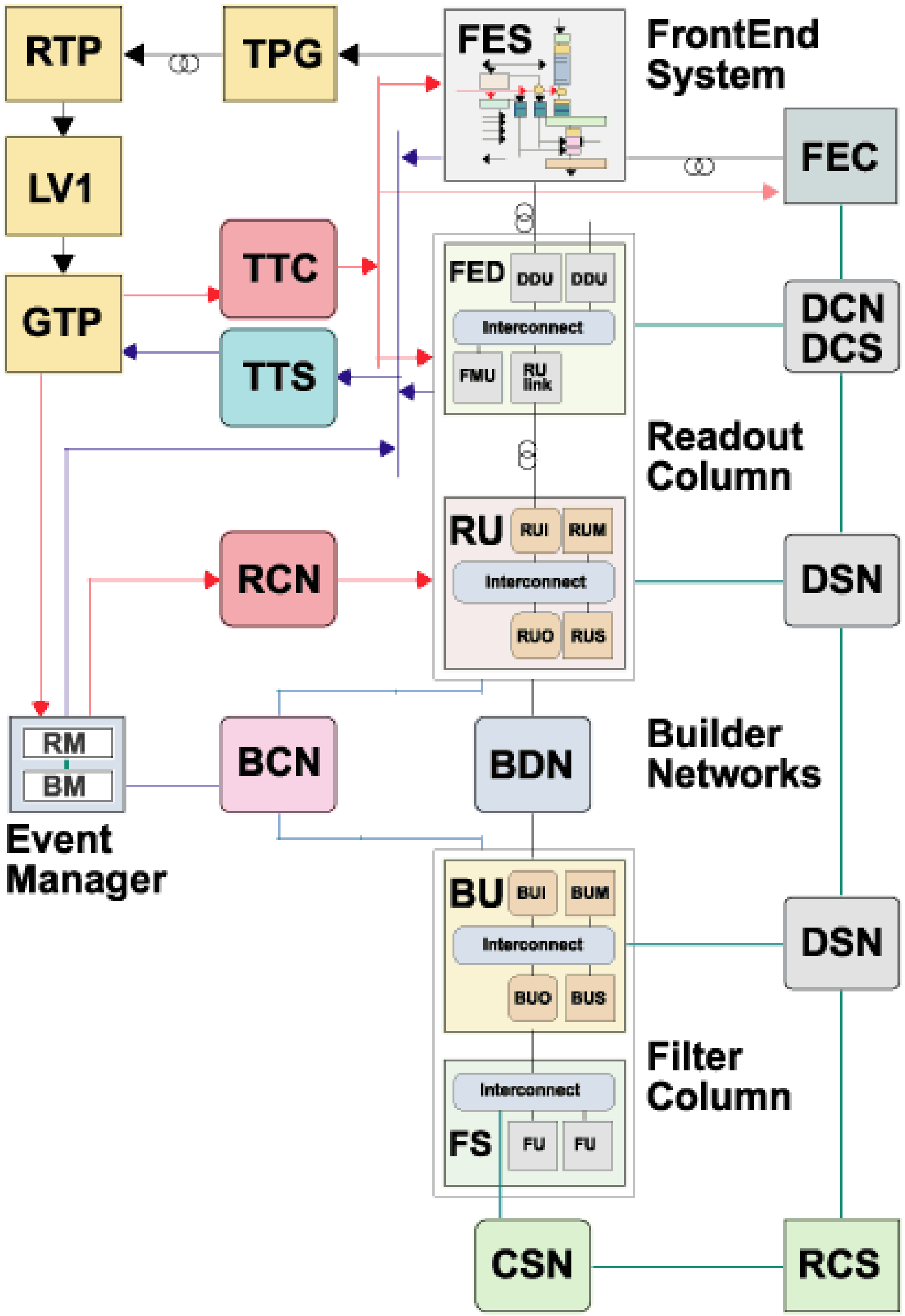
Online processing

- Trigger: event selection
  - Needs only a (small) subset of the detector data
  - Fast, as little dead-time (time period when triggering system is insensitive to new data) as possible
  - Gives “green” or “red” light to the data acquisition

Online processing

- Data acquisition
  - Interfaces to detector hardware
  - Builds complete events from fragments
  - Sends them to the higher level event filter(s)
  - Writes accepted events to mass storage
  - Very complex system

Complexity of Data acquisition



FLT	**FIRST LEVEL TRIGGER**
TPG	Trigger Primitive Generator
TDL	Trigger Data Link
RTP	Regional Trigger Processor
LV1	Calo & Trigger Processors
GTP	Global Trigger Processor
TTC	Timing, Trigger & Control
TTS	Trigger & Throttle System
EVB	**EVENT BUILDER**
BN	Builder Networks
BDN	Builder Data Network
RCN	Readout Control Network
BCN	Builder Control Network
EVM	Event Manager
RM	Readout Manager
BM	Builder Manager
FES	Front End System
DDL	Detector Data Link
RC	Readout Column
FED	FrontEnd Driver
FMU	Fast Monitoring Unit
DDU	Detector Dependent Unit
RDL	Readout Data Link
RU	Readout Unit
RUI	Readout Unit Input
RUM	Readout Unit Memory
RUO	Readout Unit Output
RUS	Readout Unit Supervisor
FC	Filter Column
BU	Builder Unit
BUI	Builder Unit Input
BUM	Builder Unit Memory
BUO	Builder Unit Output
BUS	Builder Unit Supervisor
FS	Filter Subfarm
FU	Filter Unit
HLT	High Level Trigger
RCS	**RUN CONTROL SYSTEM**
DCM	DAQ operation control and monitor
DDB	DAQ data bases
DCS	Detector Control System
DCN	Detector Control Network
FEC	FrontEnd Control
CDL	Control Data Link
CS	**COMPUTING SERVICES**
CSN	Computing and Services Network
DSN	DAQ Services Network
DA	Data archives

Computing and Communication main subsystems

Online Processing

- Monitoring
  - Detector status
  - Data acquisition performance
  - Trigger performance
  - Data quality check
- Control
  - Configure systems
  - Start/stop data taking
  - Initiate special runs (calibration, alignment)
  - Upload trigger tables, calibration constants, …

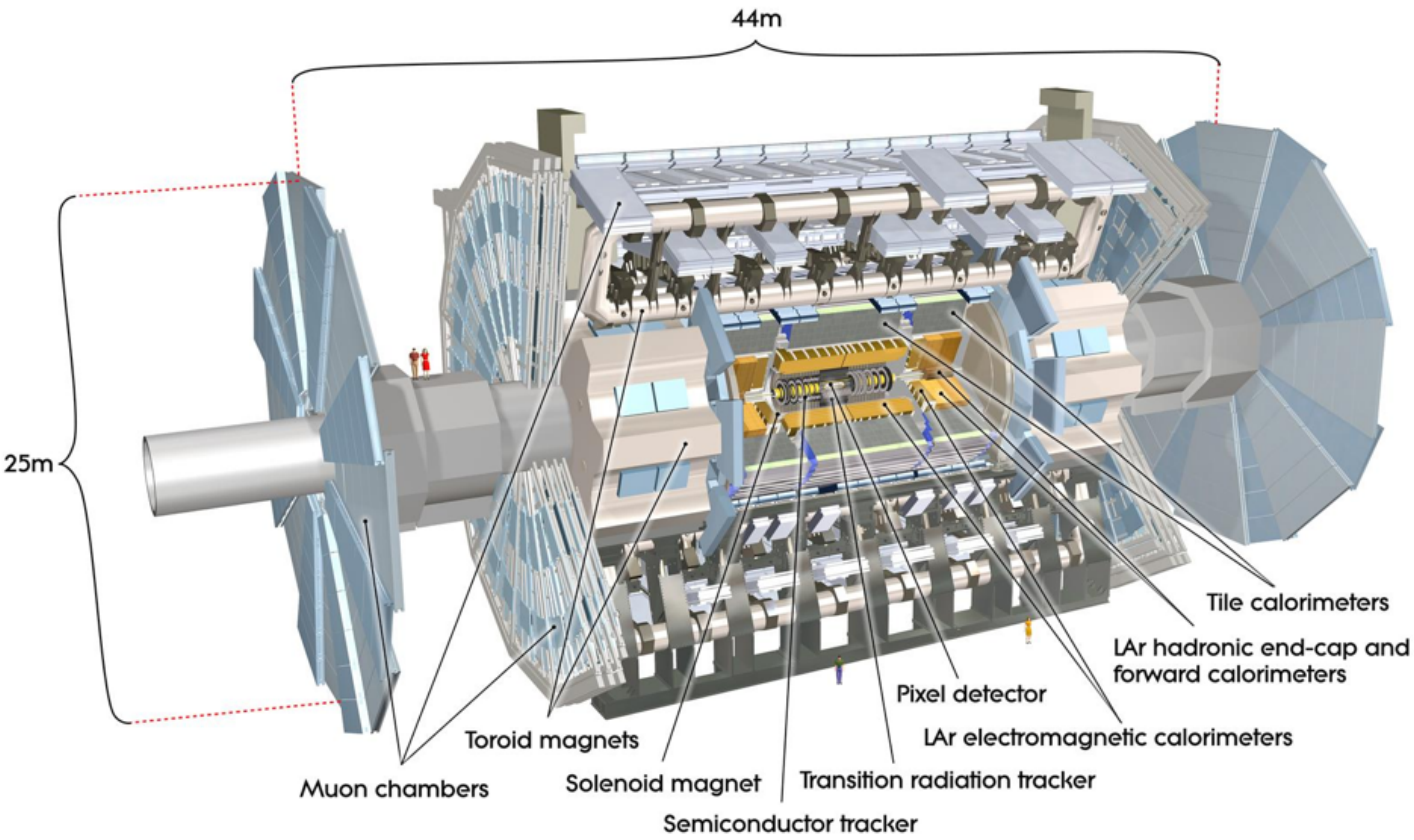
Event selection

- Primary (design) collision rate: 40 MHz
- Recording rate: a few hundred Hz
- How is this achieved?
  - Multi-level trigger - chain of yes/no decisions
  - Very fast first level: (Programmable) hardware
  - Slower higher level(s): Software on specialized or commodity processors

Event selection

- Has to be reliable
- Rejected data are lost forever
- Continuous monitoring
- Do not lose new physics
- Must therefore be open to many different signatures of potentially new physics in the detector system

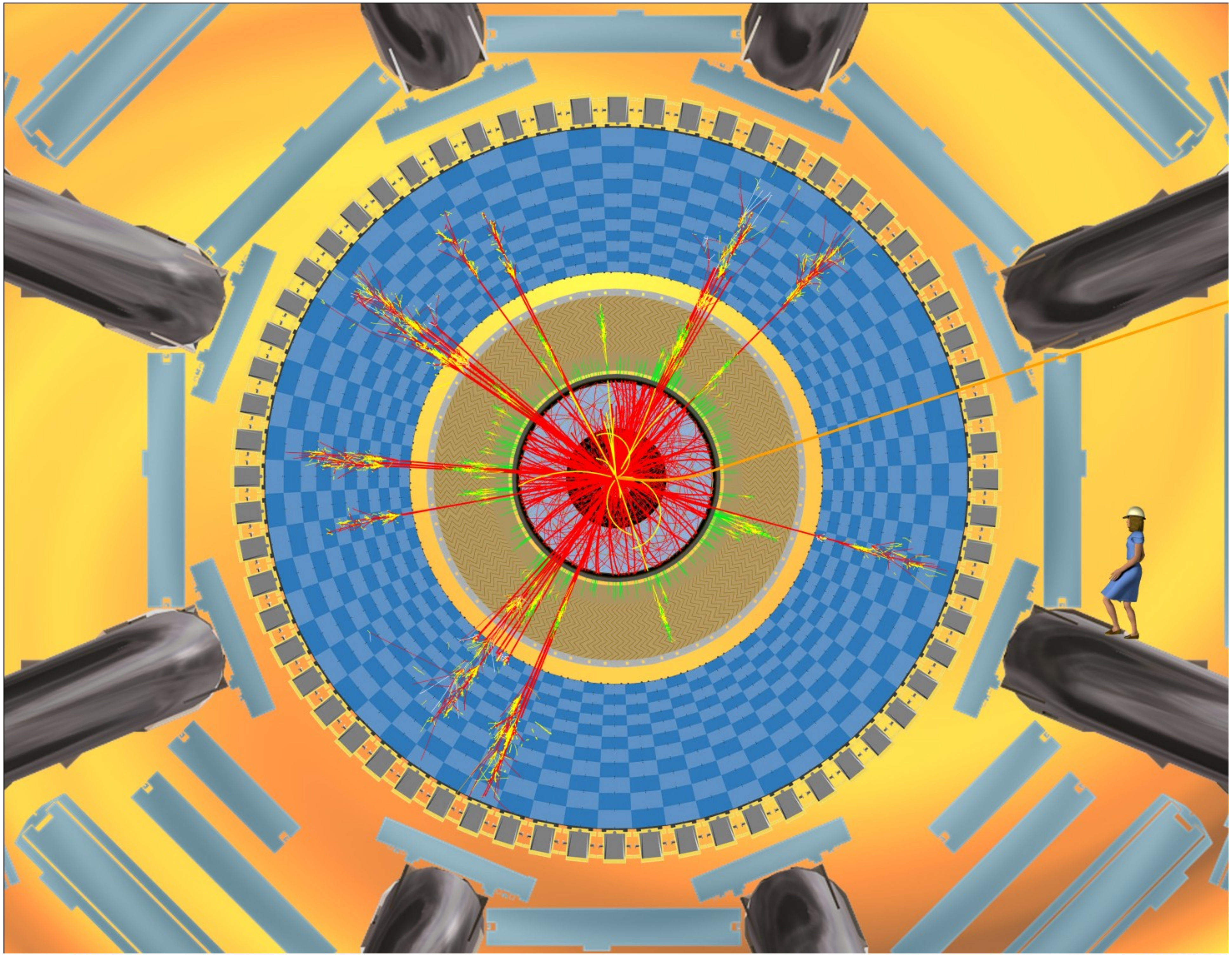
Example: ATLAS



What ATLAS sub-detectors measure

- Inner detector
    - Momentum and position of charged particles
- Electromagnetic calorimeter
    - Energy of photons, electrons and positrons
- Hadron calorimeter
    - Energy of charged and neutral hadrons
- Muon system
    - Momentum and position of muons

ATLAS detector



Event selection

- Overall guideline in designing trigger system: what are the essential features of interesting physics in the detectors?
  - Typically high-energy particles moving transversely to the beam direction
  - Results in large energy deposits in the calorimetric systems, high-energy muons in the muon system, etc.
- Multi-level trigger explores such features in various degrees of detail

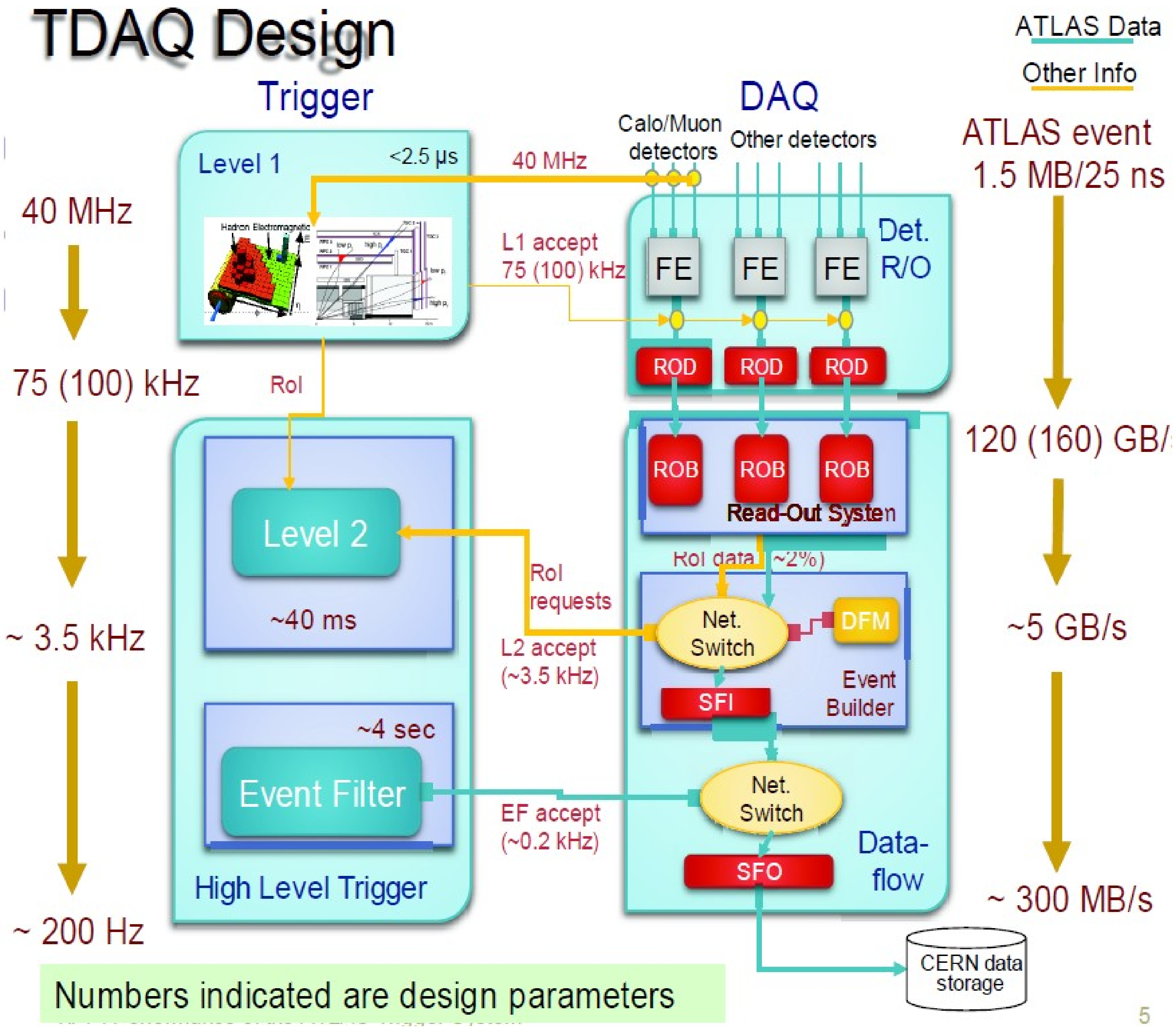
Multi level selection

- Many events can be discarded very quickly - fast level-1 trigger
- Only the surviving ones are scrutinized more carefully - high-level filter(s)
- Triggers are tailored to specific physics channels (Higgs, top, WW, ZZ, …)
  - Many such hypotheses are investigated in parallel

ATLAS triggering system

- ATLAS has three-level trigger system
    - Level 1 purely hardware-based (ASICs and FPGAs)
    - High-level trigger (level 2 and Event Filter (EF)) software-based
- Level 1 uses information mainly from calorimeters and muon system
- Level 2 also includes information from Inner Detector, uses data from Regions of Interest (RoI) identified by level 1
- EF has access to complete set of data and uses same algorithms as offline event reconstruction

ATLAS Trigger & DAQ



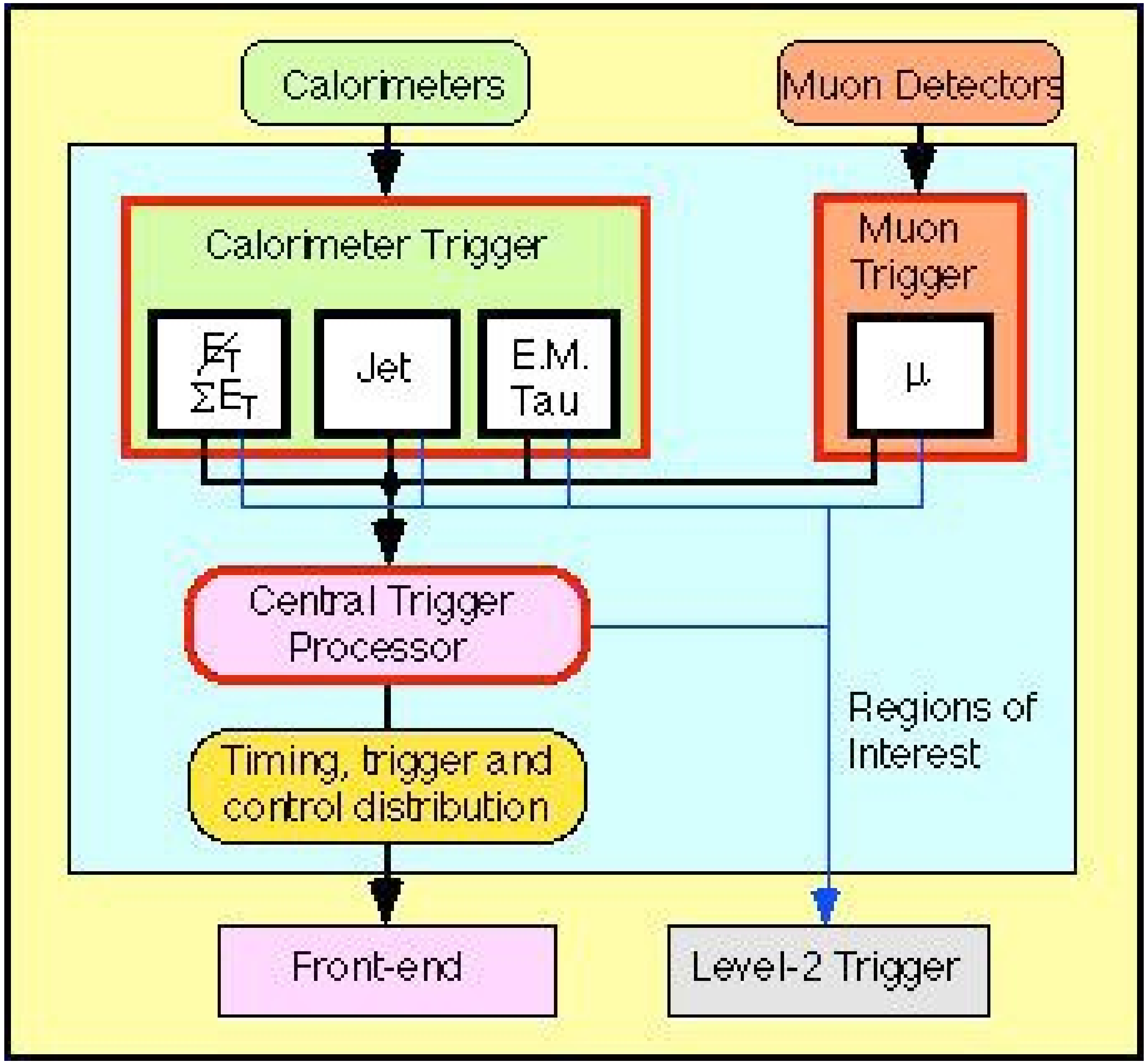
TDAQ Design

Numbers indicated are design parameters

ATLAS L1 Trigger

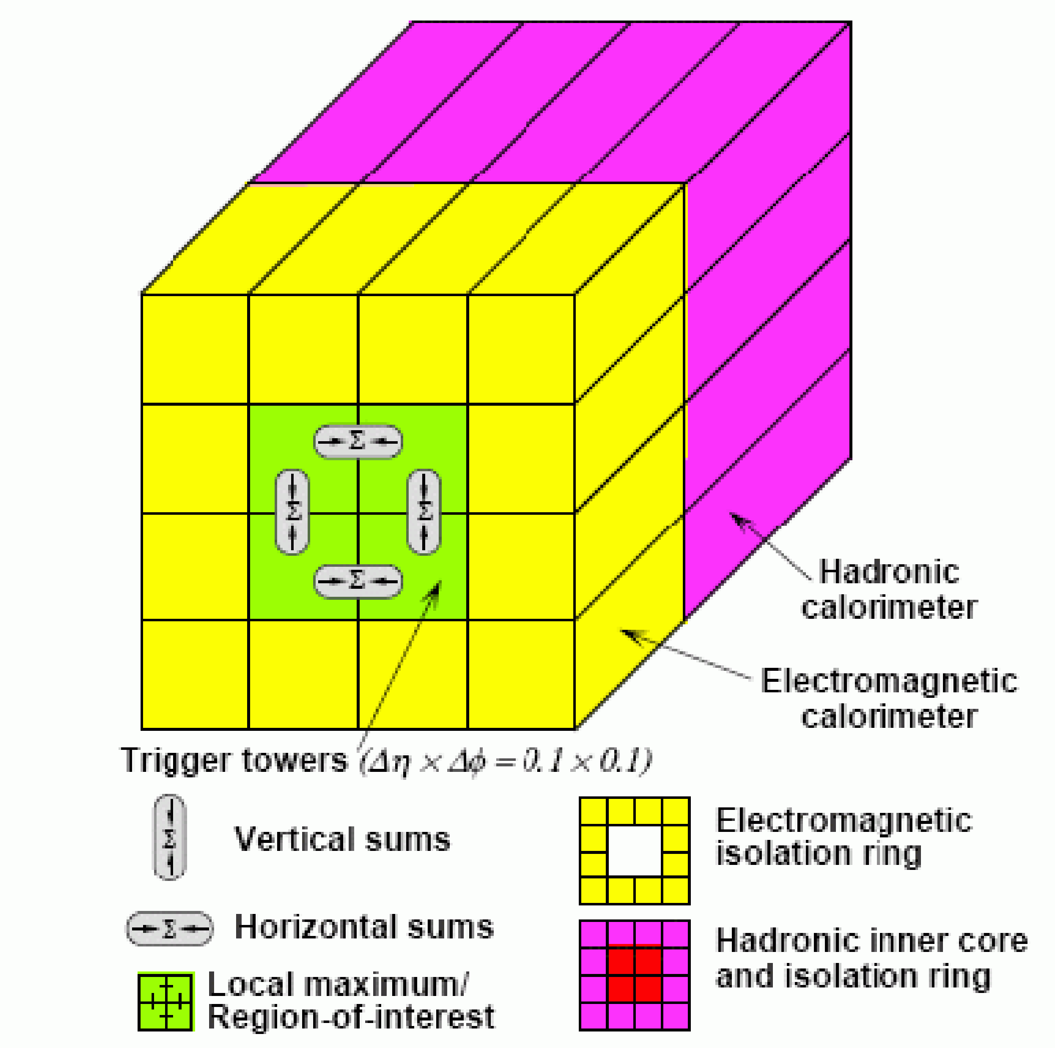
- Input (design) rate: 40 MHz
- Output rate: up to 100 kHz
- Latency (time to reach trigger decision): O(1μs)
- Data pipelined until trigger decision can be made
- Mainly 2 detector systems:
  - muons/calorimeters

ATLAS L1 Trigger



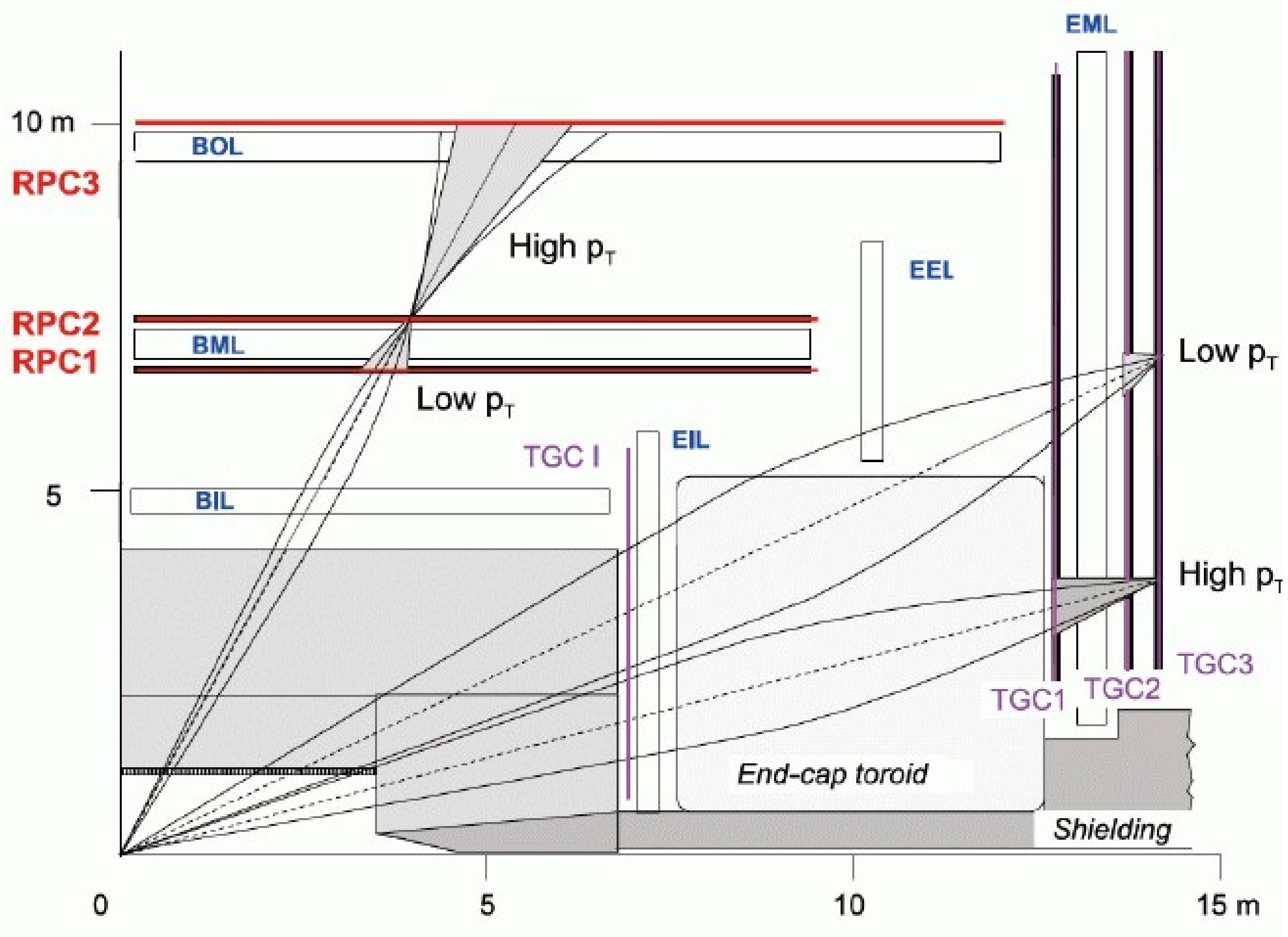
ATLAS L1 calorimeter trigger

- High-energy objects in an event:
  - Electrons/photons
  - Hadronic decays of tau lepton
  - Jet candidates
- Global event properties:
  - Total transverse energy (ET)
  - Missing ET
  - Jet sum ET
- Sends to Central trigger:
  - Multiplicity of electrons/photons and jets passing thresholds
  - Thresholds passed by total and missing ET



ATLAS L1 muon trigger

- Dedicated muon trigger chambers with good time resolution:
  - RPCs (barrel region)
  - TGCs (endcap regions)
  - Search for patterns of measurements consistent with high momentum muons coming from collision point



ATLAS L1 CTP

- Central Trigger Processor
- L1 inputs are combined to form L1 items
  - e.g. an input EM10 (electromagnetic cluster above 10 GeV) can be used in the generation of several L1 items:
    - L1_EM10: at least one EM cluster above 10 GeV
    - L1_2EM10: at least two EM cluster, each above 10 GeV
    - L1_EM10_MU6: an EM cluster above 10 and a muon above 6 GeV
- A L1 Accept is generated and sent to the detector readout electronics only if at least one L1 item survives

High-Level Filter

- Further data selection:
  - Up to 100 kHz input rate
  - A few hundred Hz output rate

- Event tagging:
  - Reconstruct physics objects
  - Mark events having interesting features facilitates quick access later

High-Level Filter

- More detailed analysis of event and underlying physics
- Runs on standard processors (commodity PCs)
- CMS: 1 stage (in contrast to ATLAS two-stage solution)

CMS high-level trigger

- Has to keep pace with the L1 output (up to 100 kHz)
- Solution: massive parallelism
- Filter farm
    - O(10000) cores
    - Decision time: O(100) ms

CMS high-level trigger

- Same software framework as in offline reconstruction
- Transparent exchange of algorithms with offline code
- Regional reconstruction
  - Concentrates on region(s) found by Level 1
- Partial reconstruction
  - Stop as soon specific questions are answered

Output of CMS high-level trigger

- Raw data are sent to Tier-0 farm (at CERN)
  - Detector data (zero compressed)
  - Trigger information + some physics objects
  - O(50) physics datasets, depending on trigger history, O(10) online streams (calibration/monitoring/alignment)
- Physics: O(1) MB @ a few hundred Hz = a few hundred MB/sec
- Alignment/Calibration: O(50) MB/sec

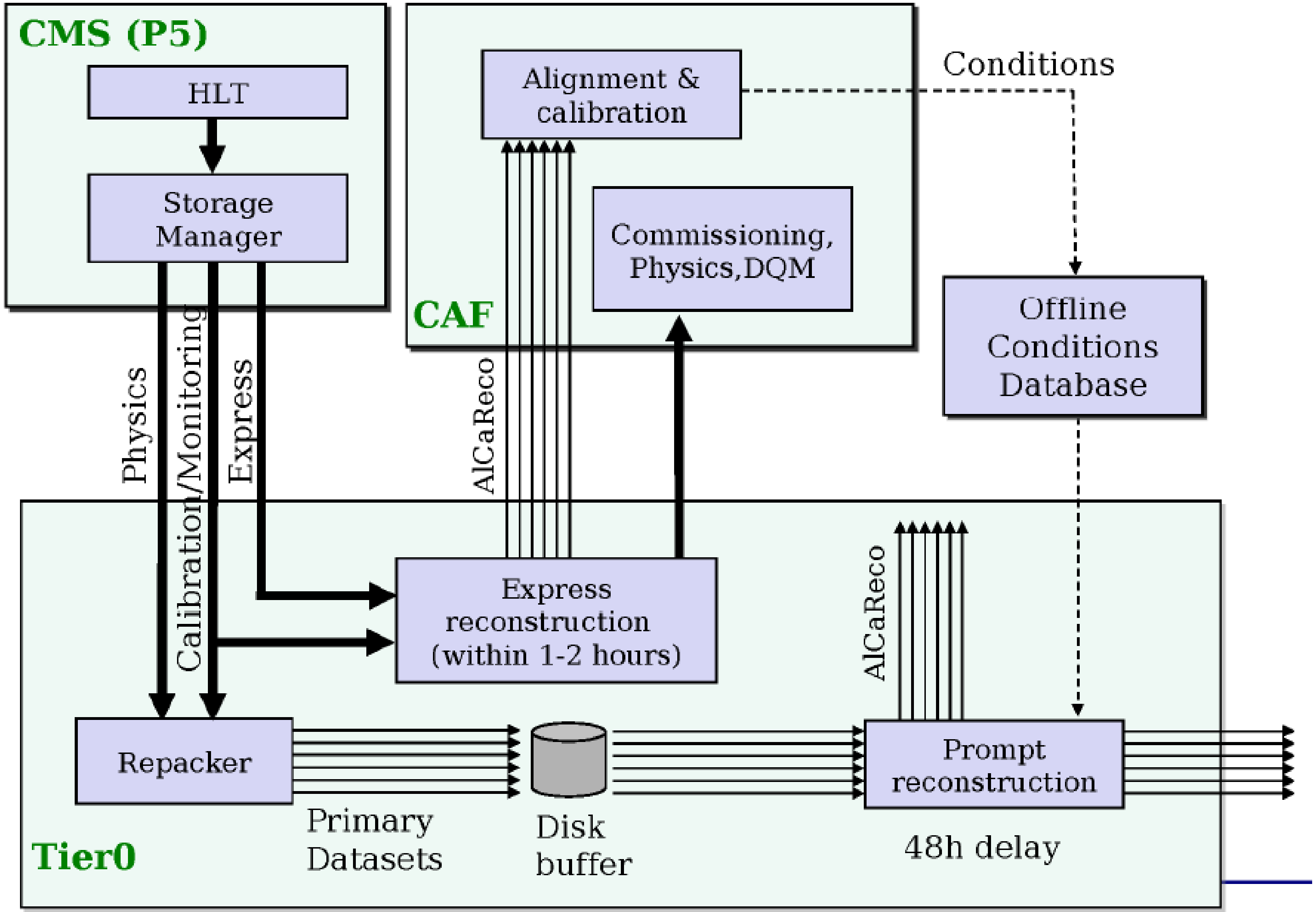
Output of CMS high-level trigger

- LHC runs for ~ 10^7 sec/year
- A few PB per year at design luminosity

Tier-0 processing

- Archive raw data on mass storage
- First event reconstruction without or with a small delay
- Archive reconstructed data on mass storage
    - A few hundred kByte/event, depending on physics
    - Reconstructed objects (hits/clusters, tracks, vertices, jets, electrons, muons)
- Send raw and processed data to Tier-1

Tier-0 processing



Summary, event selection

- Selecting a small subset of all collision events for offline analyses
    - Reducing from 40 MHz collision rate to recording rate of a few hundred Hz
- Multi-level triggering system
    - Looking for signatures of potentially interesting physics in detectors
    - First level purely hardware-based with pipelined data
    - Higher level(s) software-based, massively parallelized on filter farms

Offline processing

- Calibration
  - Convert raw data to physical quantities
- Alignment
  - Find out precise detector positions
- Event reconstruction
  - Reconstruct particle tracks and vertices (interaction points)
  - Identify particle types and decays
  - Impose physics constraints (energy and momentum conservation)

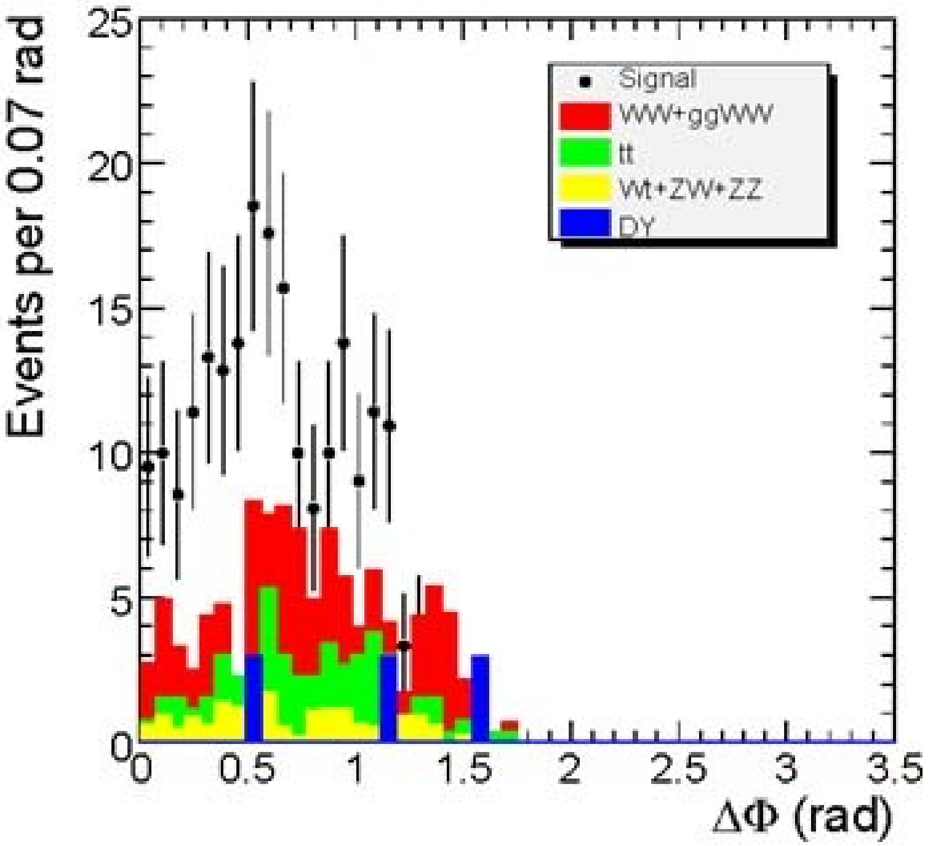
Offline Processing

- Simulation
  - Generate artificial events resembling real data as closely as possible
  - Needed for background studies, corrections, error estimation, …

Monte Carlo Method

Offline Processing

- Physics analysis
  - Extract physics signals from background
  - Compute masses, cross-sections, branching ratios, discovery limits, …
- Requires sophisticated multivariate techniques

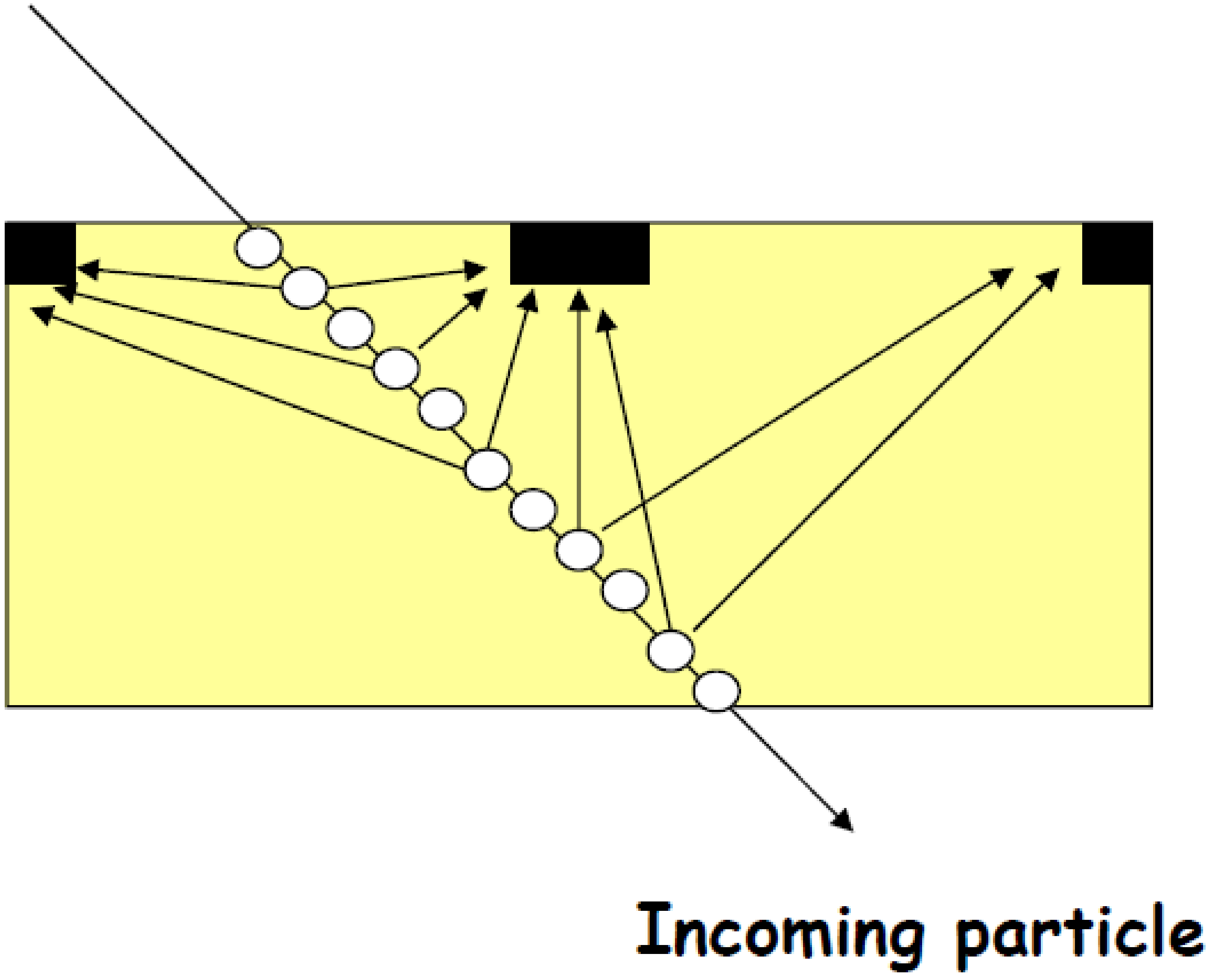


Calibration: from bits to GeV and cm

- Raw data are mostly ADC or TDC counts
- They have to be converted to physical quantities such as energy or position
- Very detector dependent
- Every detector needs calibration
- Calibration constants need to be updated and stored in a database

Silicon Tracker calibration

- Incoming particle creates electric charge in strips or pixels



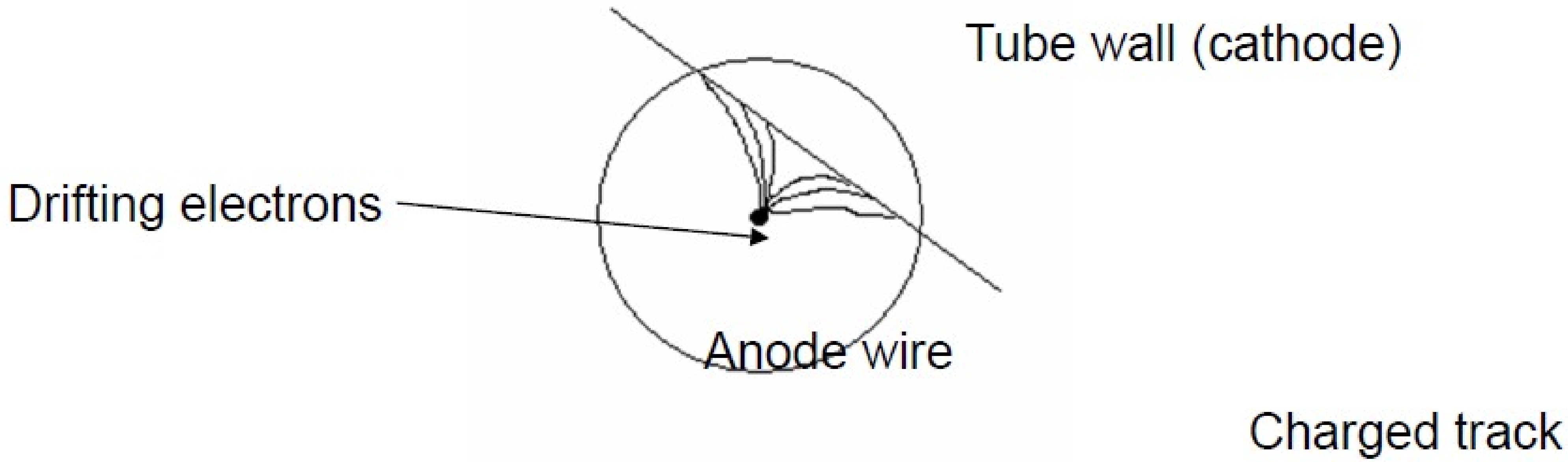
Silicon Tracker calibration

- Charge distribution depends on location of crossing point and crossing angle

- Solve inverse problem: reconstruct crossing point from charge distribution and crossing angle

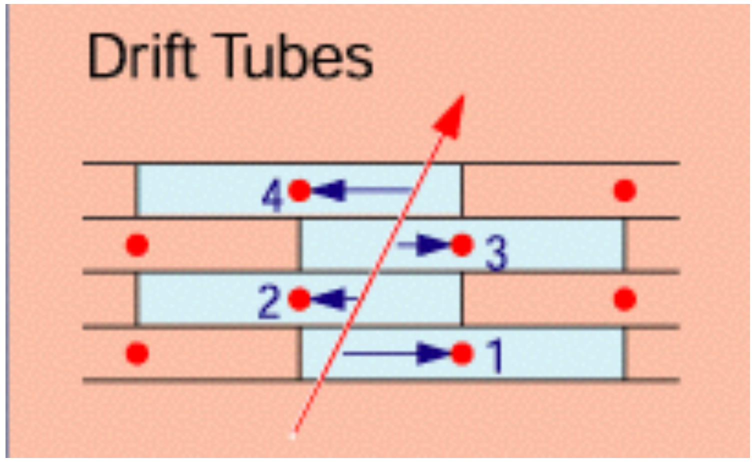
- Test beam, real data

Drift tube calibration



Drift tube calibration

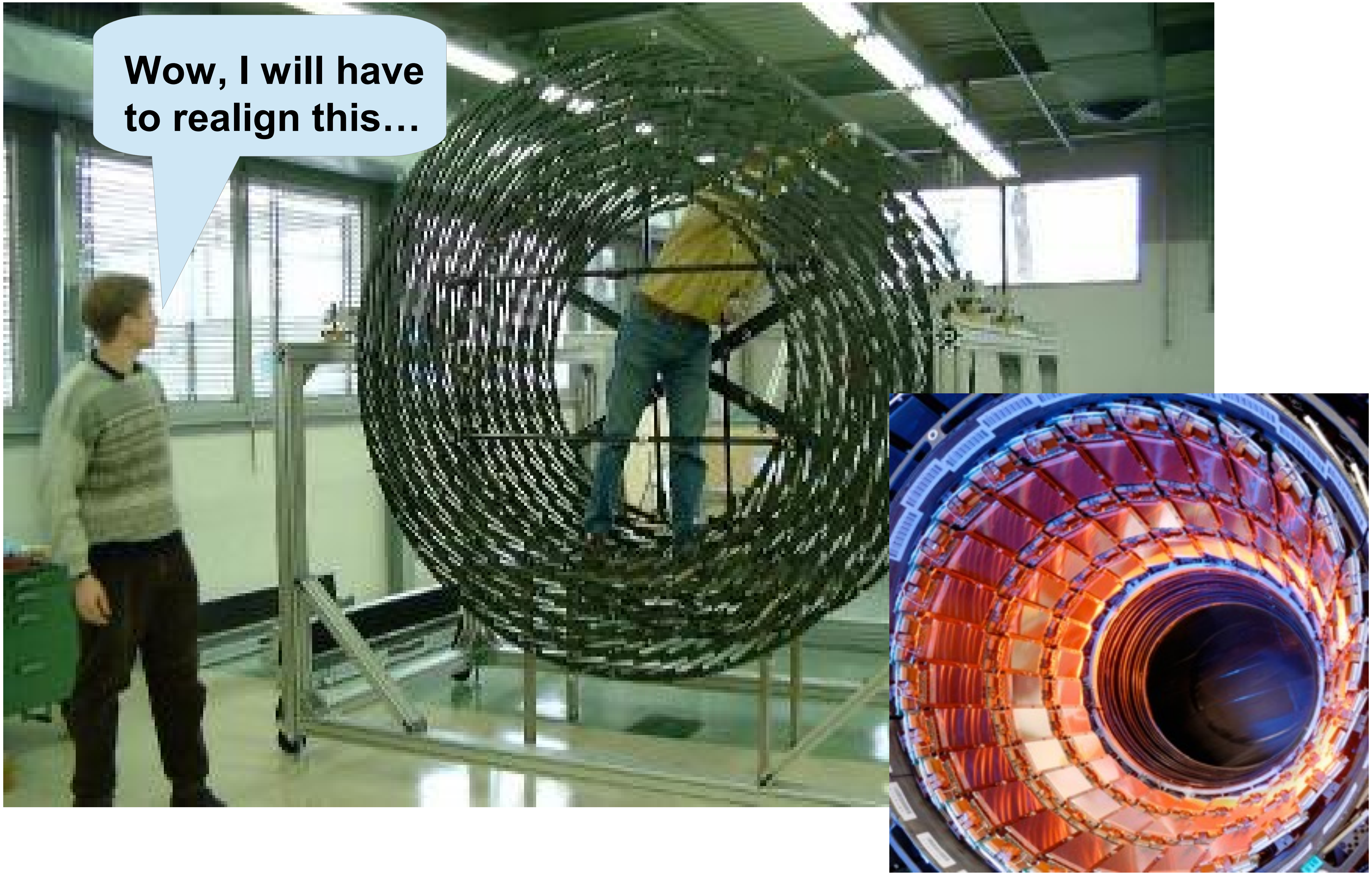
- Incoming particle ionizes gas in tube
- Electrons/ions drift to anode/cathode
- Drift time is measured
- Must be converted to drift distance
- Time/distance relation must be determined (not always linear)
- Test beam, real data



Alignment: Where are the detectors?

- Tracking detectors are very precise instruments
- Silicon strip detector: ~ 50 μm
- Pixel detector: ~ 10 μm
- Drift tube: ~ 100 μm
- Positions of detector elements need to be known to a similar or better precision

Example: CMS tracker



Alignment

- Mechanical alignment
- Measurement taken before assembly
- Switching on the magnetic field
- Laser alignment
- Alignment with charged tracks from collisions, beam halo and cosmic rays

Alignment

- Difficult because of huge number of parameters to be estimated (~ 100000)

- Continuous process

- Alignment constants need to be updated and stored in a database

Event reconstruction

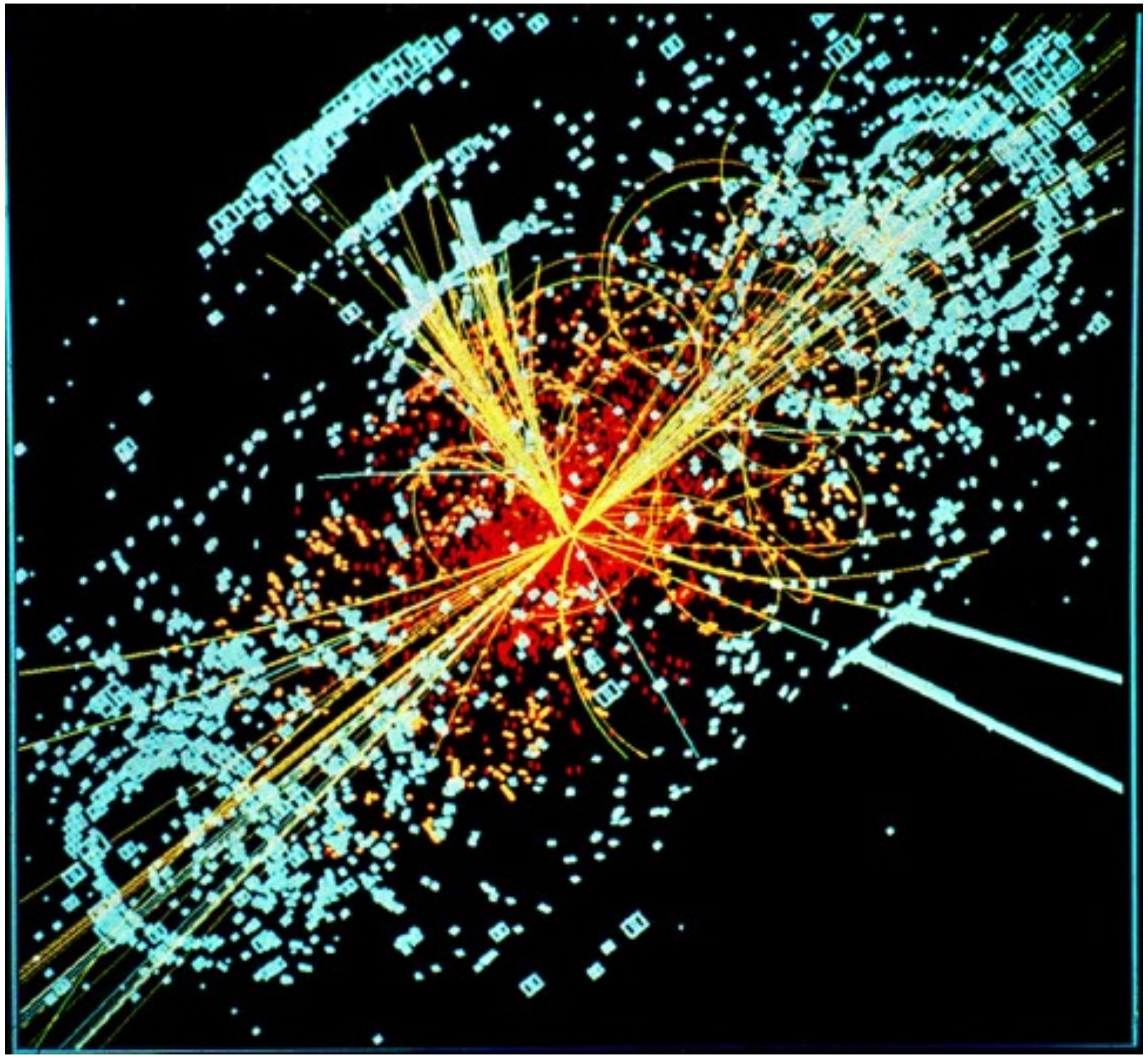
- Find out which particles have been created where and with which momentum
- Many can be observed directly
- Some are short-lived and have to be reconstructed from their decay products
- Some (neutrinos) escape without leaving any trace

Event reconstruction

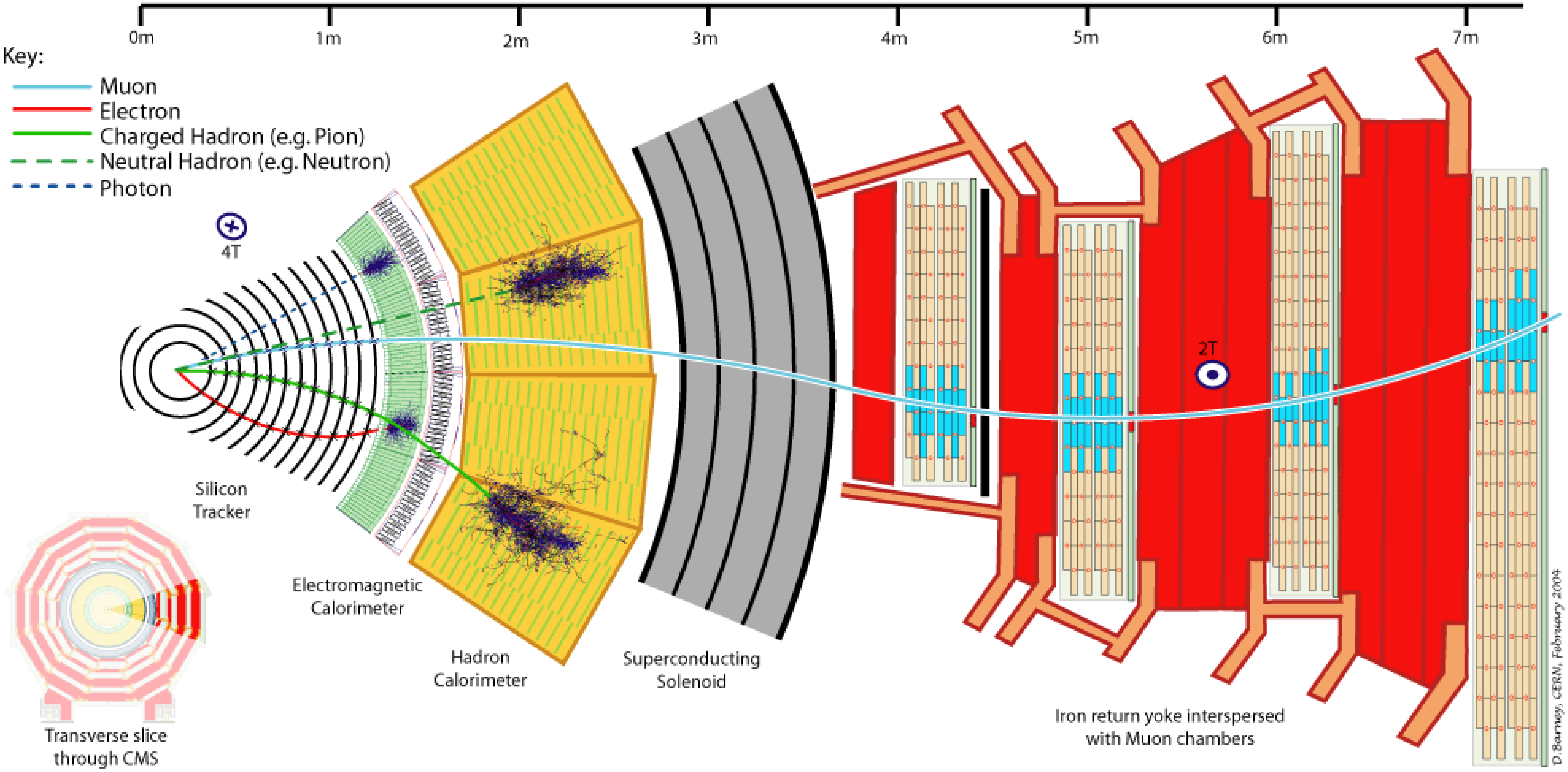
- Reconstruct charged particles
- Reconstruct neutral particles
- Identify type of particles
- Reconstruct vertices (interaction points)
- Reconstruct kinematics of the interaction
- Not trivial, very time-consuming …

Event reconstruction

CMS: Higgs decay into two jets



What CMS sub-detectors measure

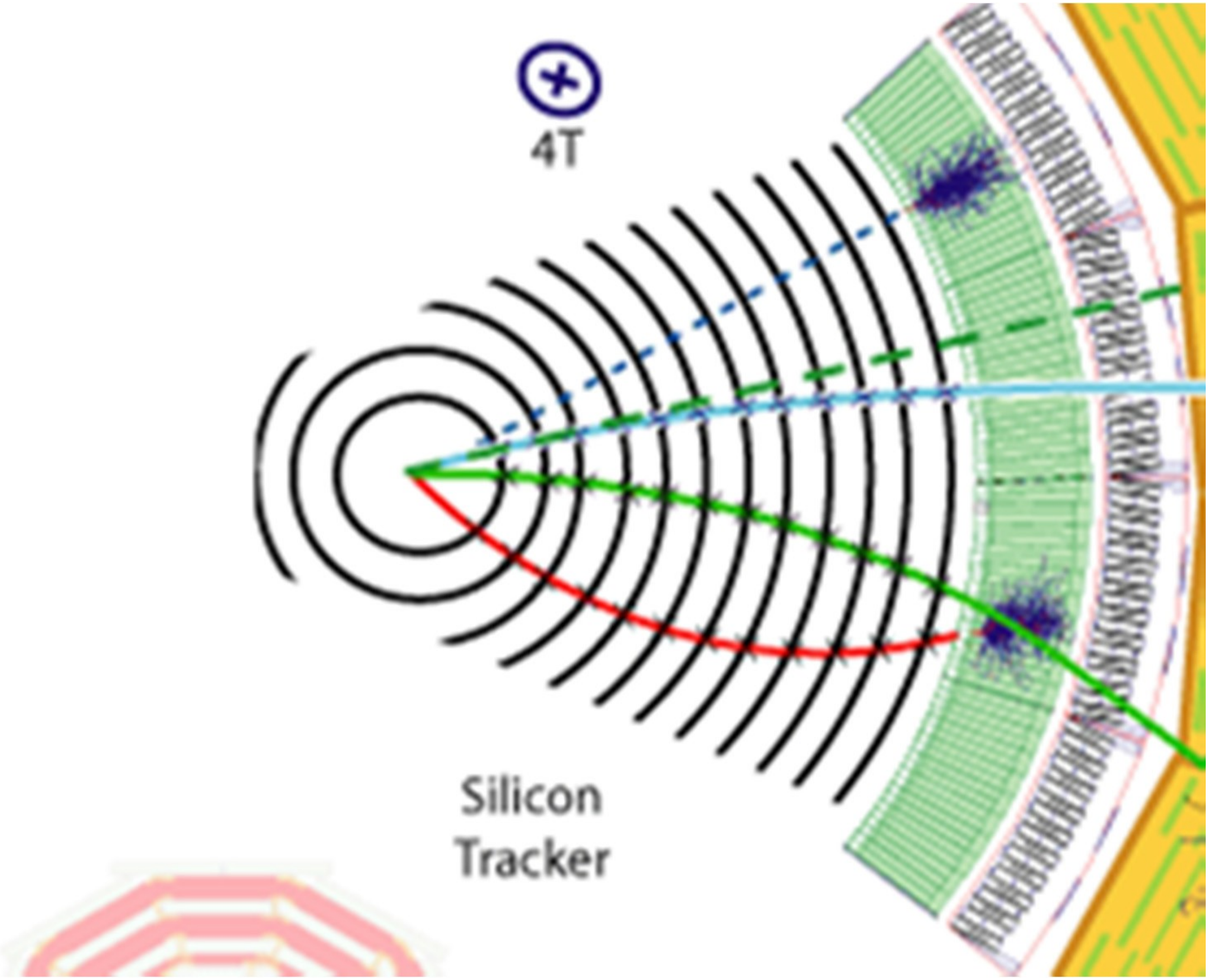


D.Barney, CERN, February 2004

Charged particles

- Charged particles are detected by tracker and calorimeters
- Muons also reach the muon system
- Very high number of low-momentum charged particles
- Select by threshold on transverse momentum

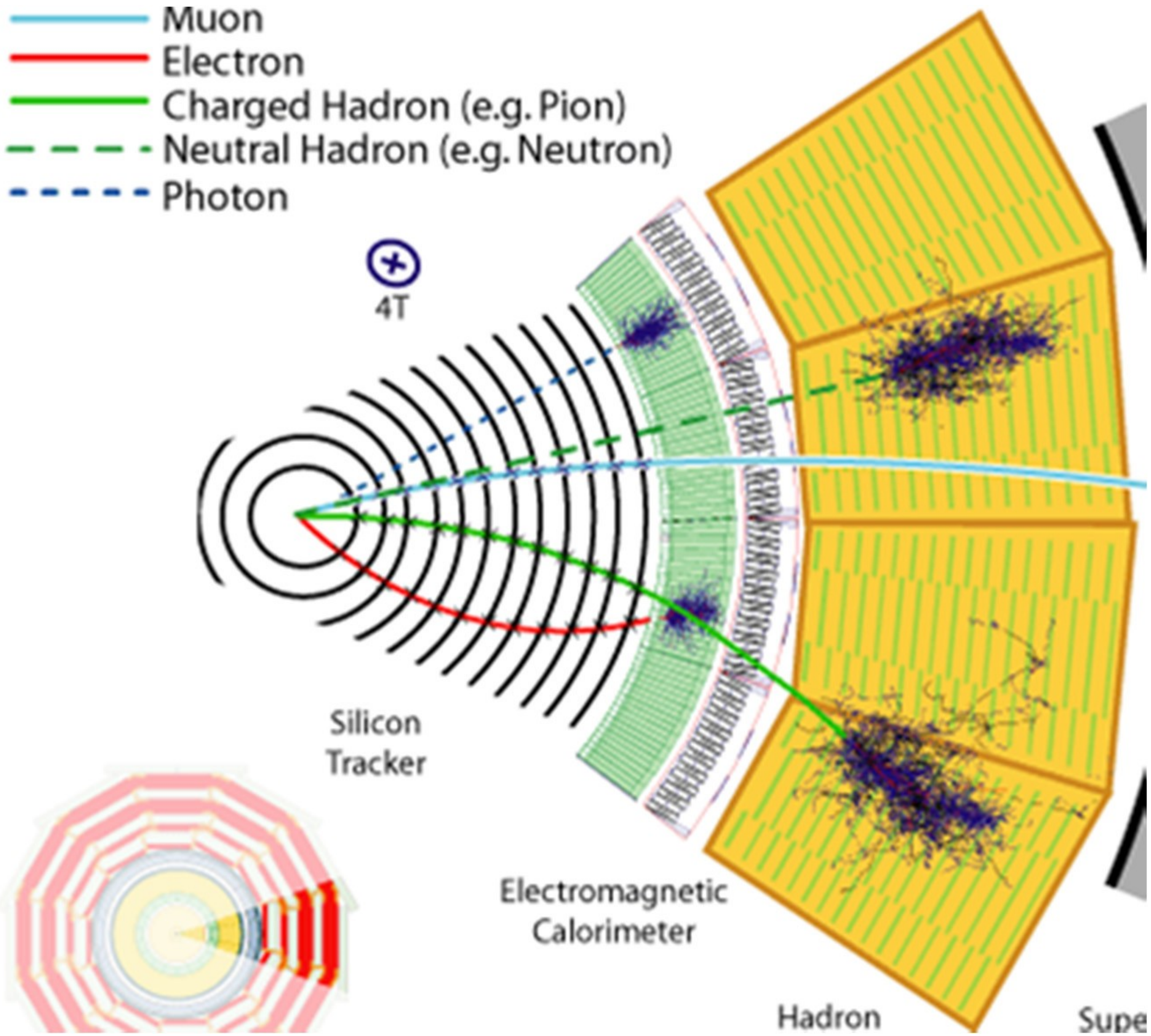
Charged particles



Neutral particles

- Neutral particles are detected mainly by calorimeters (e.g. photons, neutrons)
- They should deposit their entire energy
- Some of them decay into charged particles which are detected by the tracker (e.g. K^0)
- Neutrinos escape without leaving a trace (missing energy)

Neutral particles



Reconstruction of charged particles

- Trajectory is curved because of the magnetic field
- Position is measured in a number of places – “hits”
- Determine track parameters (location, direction, momentum) plus their estimated uncertainties from the position measurements

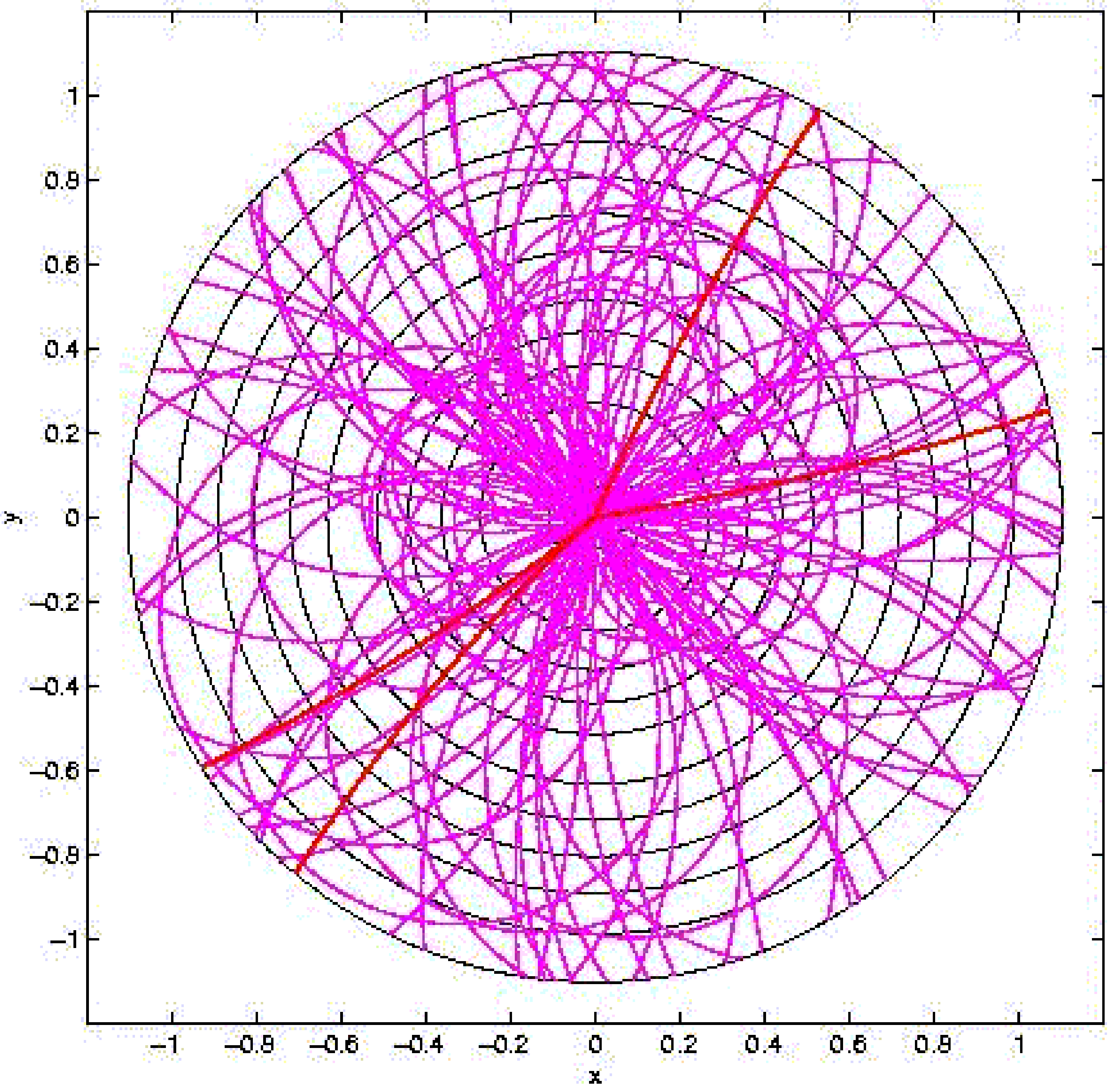
The difficulties

- Assignment of hits to particle is unknown
- Huge background from low-momentum tracks
- Additional background from other interactions in the same beam crossing, from adjacent beam crossings and from noise in the electronics

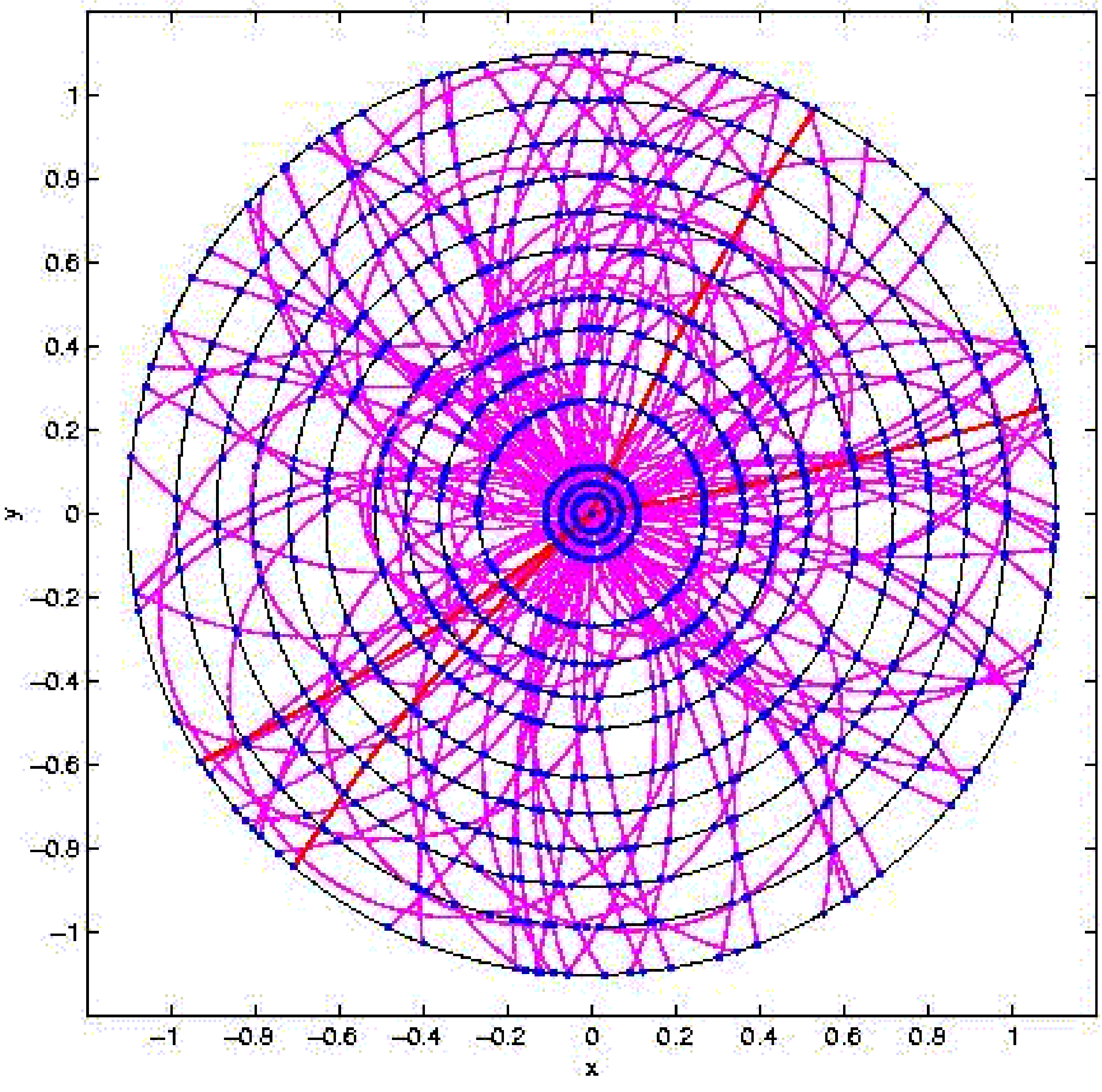
More difficulties

- Charged particles interact with all the material, not only the sensitive parts
  - Multiple Coulomb scattering
    - Changes direction, but not momentum
  - Energy loss by ionization
    - All charged particles, changes momentum
  - Energy loss by bremsstrahlung
    - Electrons and positrons, changes momentum

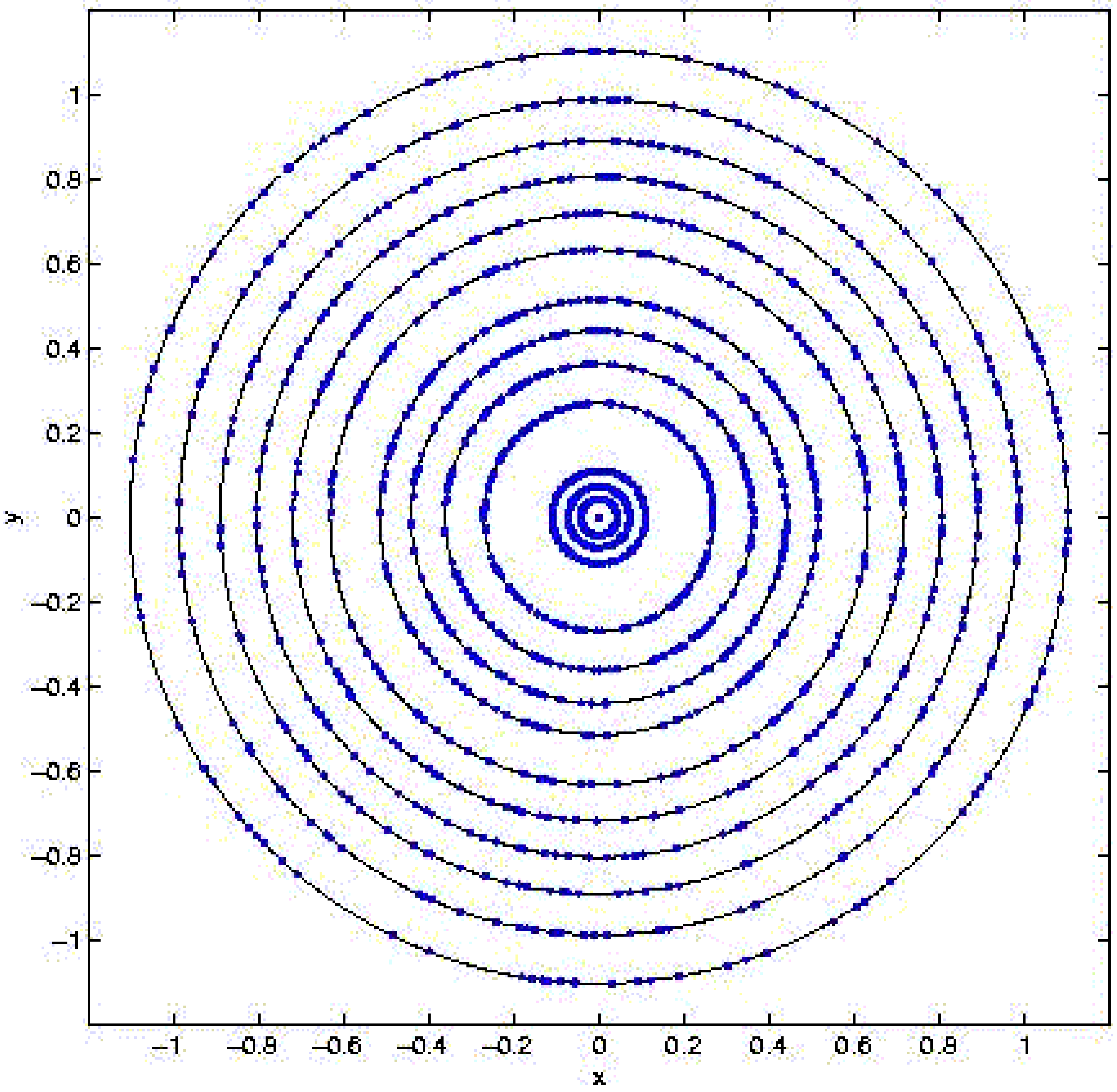
Tracks only



Tracks with hits



Hits only



?

Decomposition of the problem

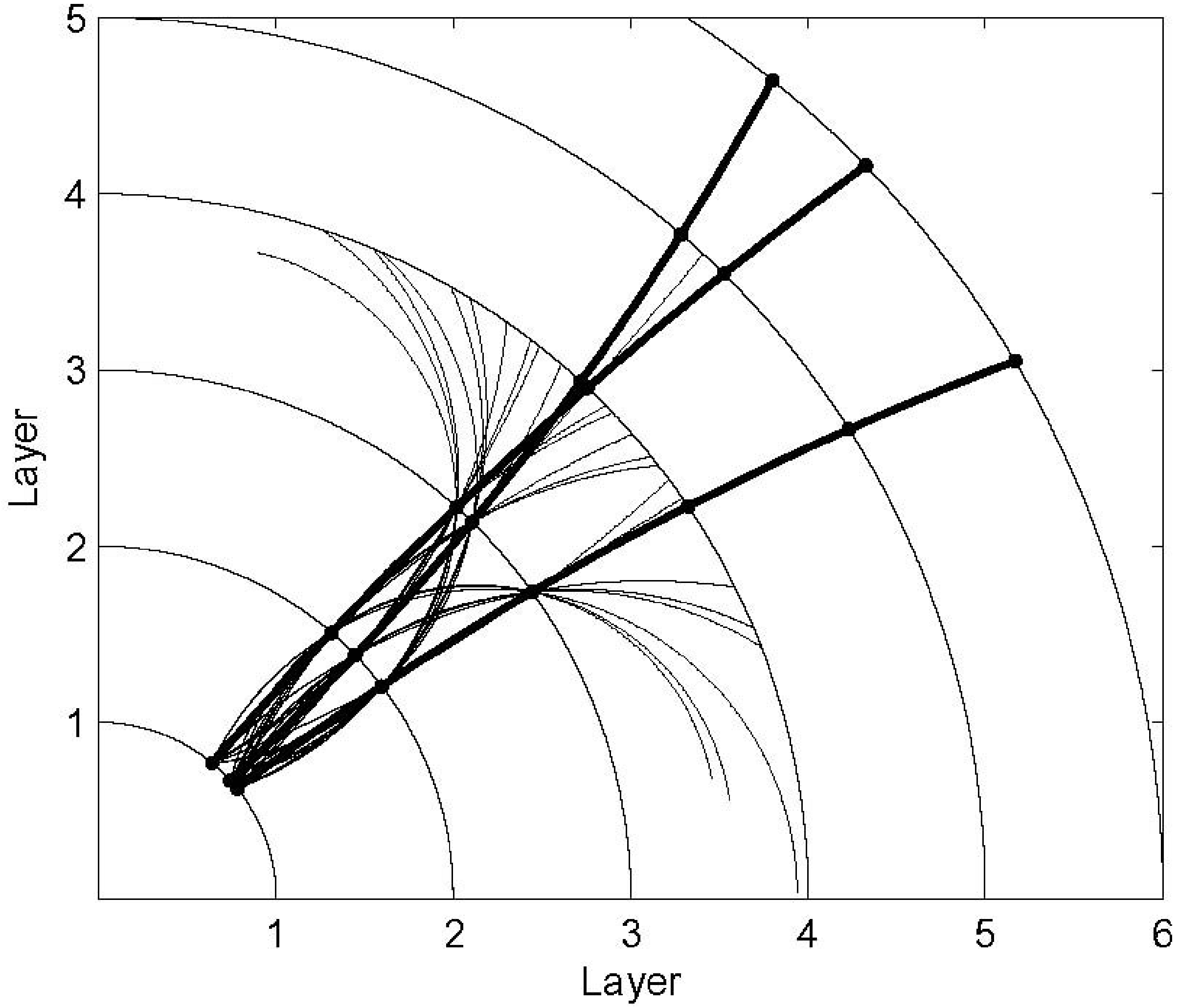
- Pattern recognition or track finding
  - Assign detector hits to track candidates (collection of hits all believed to be created by the same particle)
- Parameter estimation or track fit
  - Determine track parameters + their estimated uncertainties (covariance matrix)
- Test of the track hypothesis
  - Is the track candidate the trace of a real particle?

Track finding

- Depends a lot on the properties of the detector:
  - Geometry, configuration
  - Magnetic field
  - Precision
  - Occupancy
- Many solutions available
- No general recipe

A few track finding algorithms

- Track following ▶
- Kalman filter
- Combinatorial
- Hough transform
- Artificial neural network



Track fit

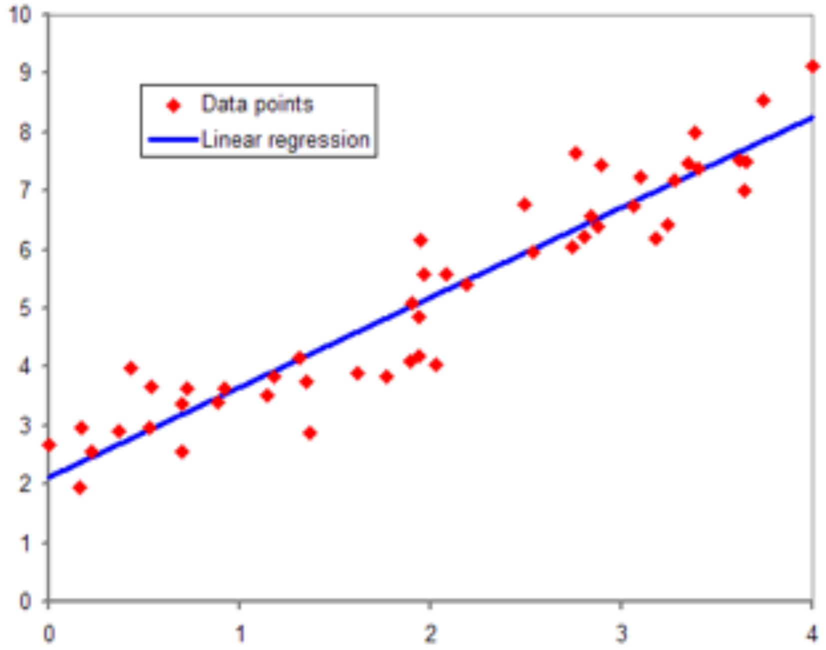
- Determine (estimate) track parameters
- Determine uncertainties of estimated track parameters (covariance matrix)
- Test track hypothesis
- Reject outliers
  - Distorted hits
  - Extraneous hits
  - Electronic noise hits

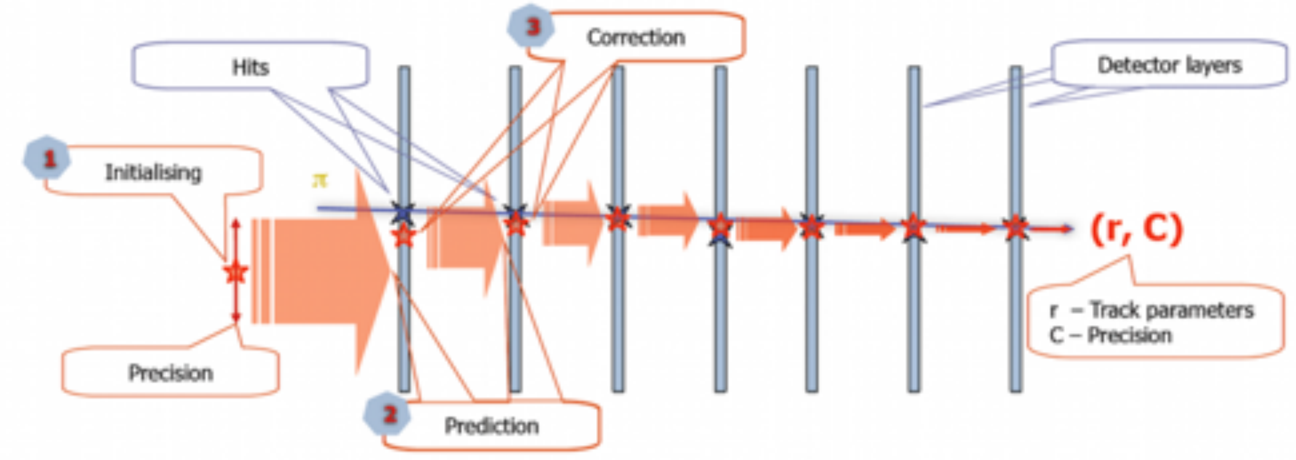
Ingredients

- Magnetic field
  - Constant or variable
- Track model
  - Solution of the equation of motion
  - Analytic (explicit) or numerical
- Error model
  - Observation errors
  - Process noise

Estimation of track parameters

- Most estimators minimize a least-squares objective function
  - Linear regression
  - Kalman filter
- Robust estimation
  - Adaptive filter
  - Automatic suppression of outlying hits



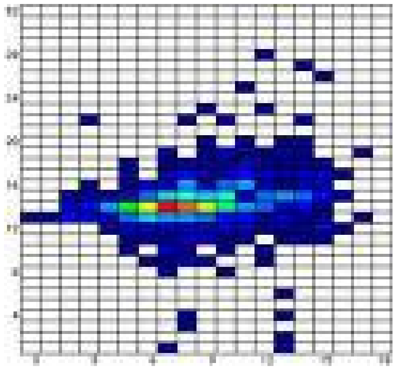


Reconstruction of neutral particles

- Neutral particles are only seen by the calorimeters
- Photons are absorbed in the electromagnetic calorimeter
- Neutral hadrons are absorbed in the hadronic calorimeter
- Neutrinos are not detected directly

Shower finding

- An incident particle produces a shower in the calorimeter
- A shower is a cluster of cells with energy deposit above threshold



Shower finding

- Overlapping clusters must be separated

- Various clustering techniques are used to find showers

- The algorithms depend on various characteristics of the calorimeter

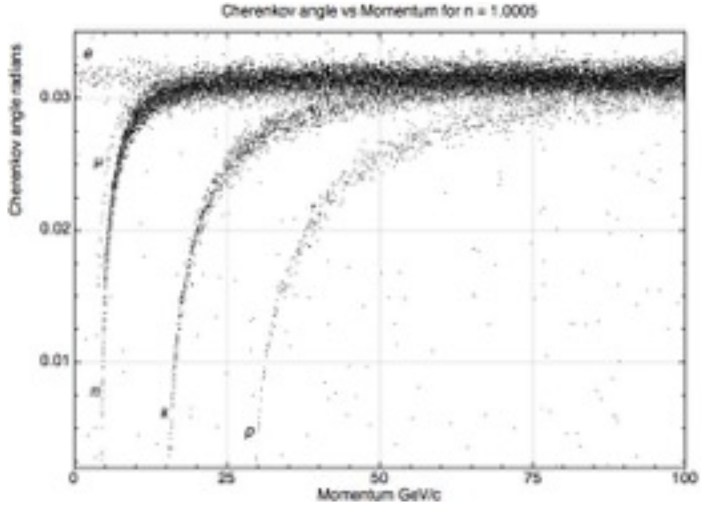
  - Type (electromagnetic or hadronic)

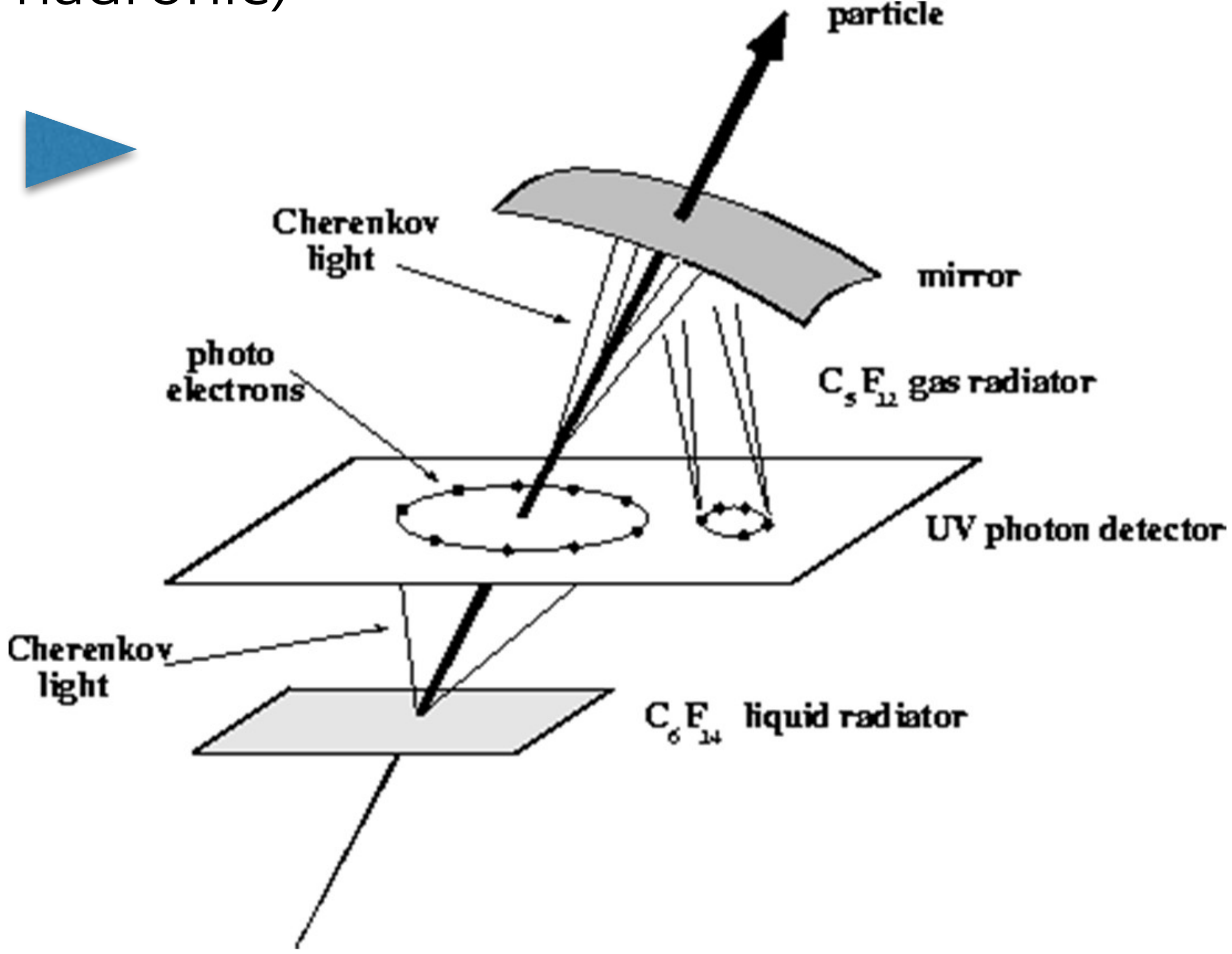
  - Technology (homogeneous or sampling)

  - Cell geometry, granularity

Particle identification

- Determining the type of a particle
- Dedicated detectors
  - Calorimeter (electromagnetic or hadronic)
  - Ring imaging Cherenkov (RICH)
  - Transition radiation detector
  - Ionization measurements





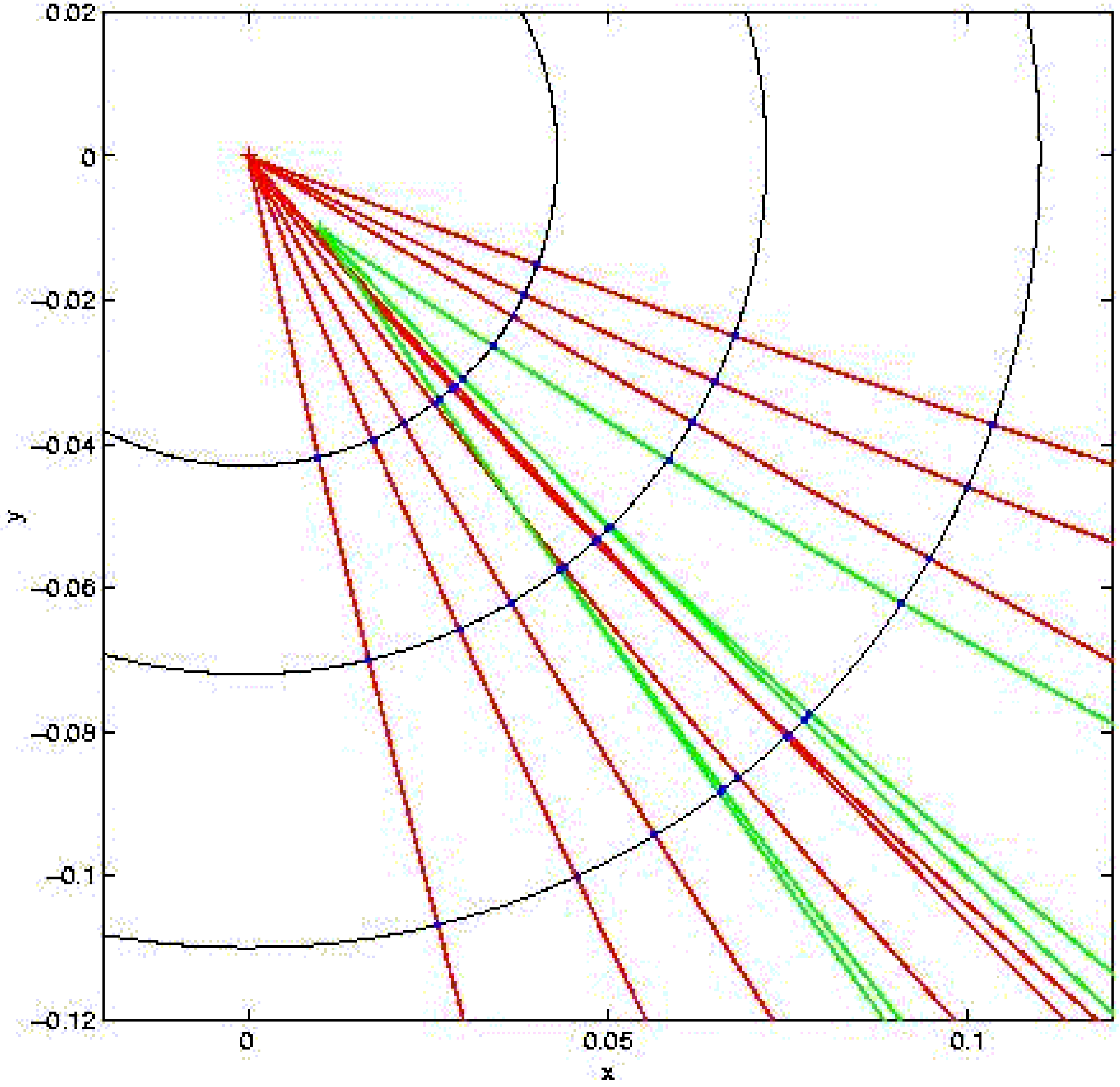
Particle identification

- Combining information from several detectors
  - Shower in electromagnetic calorimeter + no matching track in tracker -> photon
  - Shower in electromagnetic calorimeter + matching track in tracker -> electron/positron
  - Shower in hadronic calorimeter + matching track in tracker -> charged hadron
  - Track in muon system + matching track in tracker -> muon

Vertex reconstruction

- Primary vertex: interaction of the two beam particles – easy
- Secondary vertices: decay vertices of unstable particles – difficult
- Emphasis on short-lived unstable particles which decay before reaching the tracker

Primary and secondary tracks



Primary tracks

Secondary tracks

The difficulties

- Association of tracks to vertices is unknown
- Secondary tracks may pass very close to the primary vertex (and vice versa)
    - Especially if decay length is small
- Track reconstruction may be less than perfect
    - Outliers, distortions, incorrect errors

Decomposition of the problem

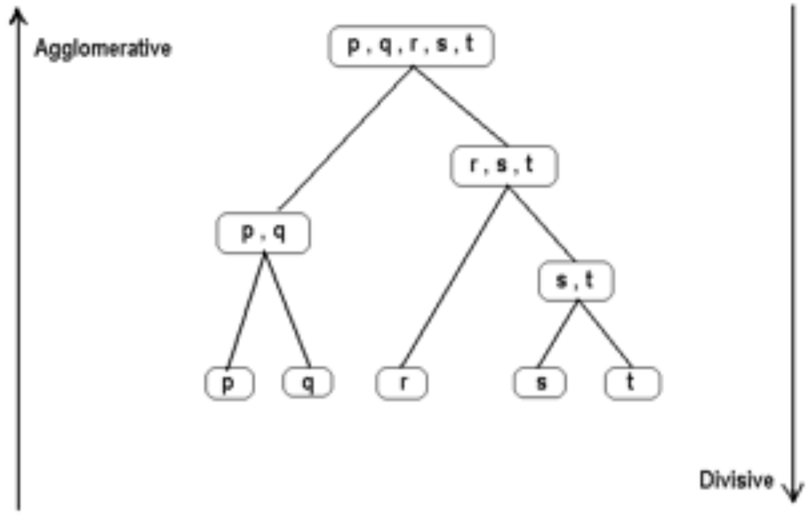
- Pattern recognition or vertex finding
    - Assign tracks to vertex candidates
- Parameter estimation or vertex fit
    - Determine vertex location + covariance matrix, update track parameters
- Test of the vertex hypothesis
    - Is the vertex candidate a real vertex?

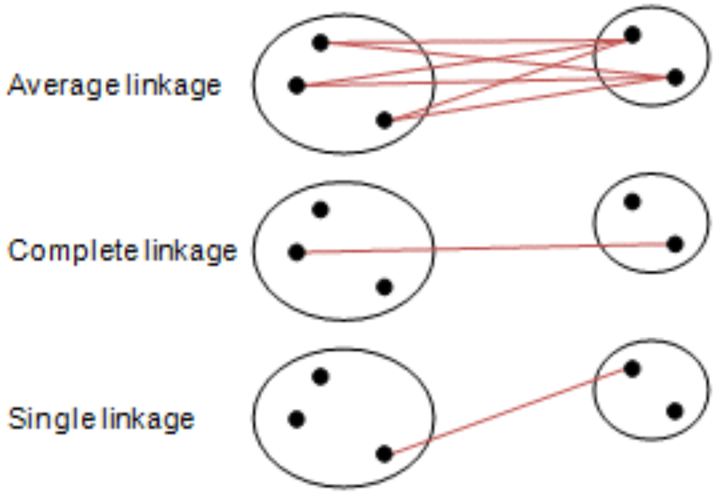
Vertex finding

- Almost independent of the detector geometry
- Secondary vertex finding may depend on the physics channel under investigation
- Essentially a clustering problem
- Many solutions available

A few vertex finding algorithms

- Hierarchical clustering
  - Single linkage, complete linkage, …
- Machine learning
  - k-means, competitive learning, deterministic annealing, …
- Estimation based
  - robust location estimation, iterated vertex fit





Vertex fitting

- Most estimators minimize a least-squares objective function
  - Linear regression
  - Kalman filter
- Robust estimation
  - Adaptive filter
  - Automatic suppression of outlying hits

Kinematical fitting

- Impose physical constraints
    - Momentum conservation
    - Energy conservation
- Test mass hypotheses
    - See whether kinematics are compatible with the decay of a certain particle
- Reconstruct invisible particles

Storage

- Event reconstruction produces physics objects
    - Tracks
    - Vertices
    - Identified particles
    - Jets
    - Tags
- Need to be stored

Storage

- Preferred tool for event data: ROOT
- Physics objects depend on
  - Alignment
  - Calibration
  - Version of the reconstruction program
  - Algorithm parameters
- Must be stored as well (database)

Summary, event reconstruction

- Track reconstruction
  - Charged: determine track parameters from hits
  - Neutral: find showers in calorimeters
- Particle identification
- Vertex reconstruction
  - Determine number of production points and their positions from the set of reconstructed tracks
- Kinematic fitting
  - Refine estimates by e.g. imposing physical constrain

Simulation

- Why do we need simulation?
  - Optimization of detector in design phase
  - Testing, validation and optimization of trigger and reconstruction algorithms
  - Computation of trigger and reconstruction efficiency
  - Computation of geometrical acceptance corrections
  - Background studies
  - Systematic error studies

Simulation steps

- Physics generation
  - Generate particles according to physics of the collision
  - General-purpose and specialized generators
- Event simulation
  - Track particles through the detector, using detector geometry and magnetic field
  - Simulate interaction of particles with matter
  - Generate signals in sensitive volumes
  - Simulate digitization process (ADC or TDC)
  - Simulate trigger response

Simulation steps

- Reconstruction
  - Treat simulated events exactly as real events
  - Keep (some) truth information: association of hits to tracks, association of tracks to vertices, true track parameters, true vertex parameters, …
  - Store everything

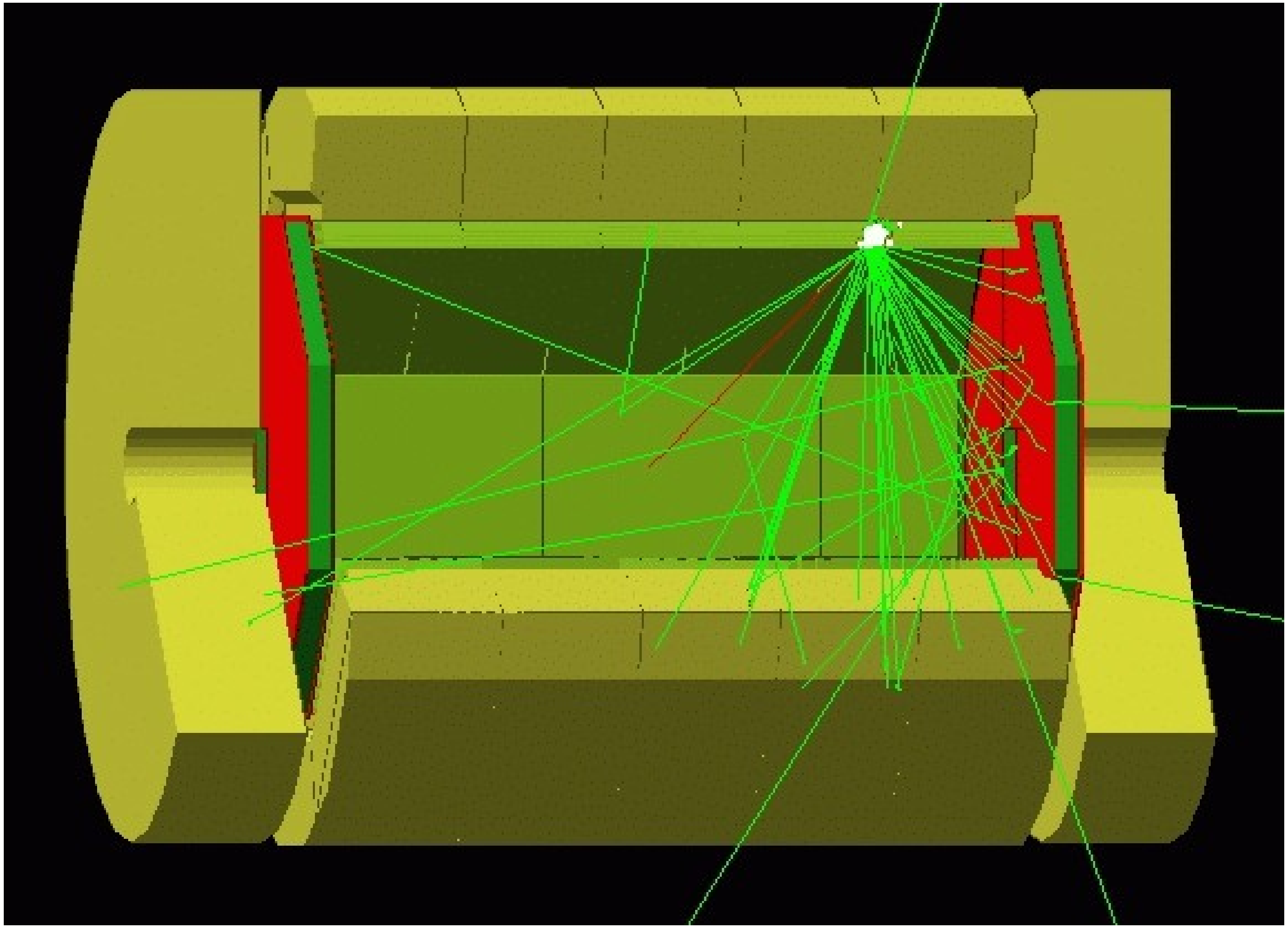
Event simulation

- Was frequently (and still sometimes is) experiment-specific
- Now there is a widely used standard:
  - GEANT4
    - Object oriented, C++
    - Extremely general and versatile

Detector description

- Geometry
  - Partition the detector into a hierarchy of volumes
  - Describe their shape and their position relative to a mother volume
  - Use possible symmetries
- Material
  - Chemical composition, density
  - Physical properties: radiation length, interaction length

An example detector model



Physics analysis

- Event selection
  - Multidimensional criteria
  - Statistics, neural networks, generic algorithms, ⋯
- Signal extraction
  - Study background
  - Determine significance of signal
- Corrections
  - Detector acceptance, reconstruction efficiency, ⋯
  - From simulated and from real data

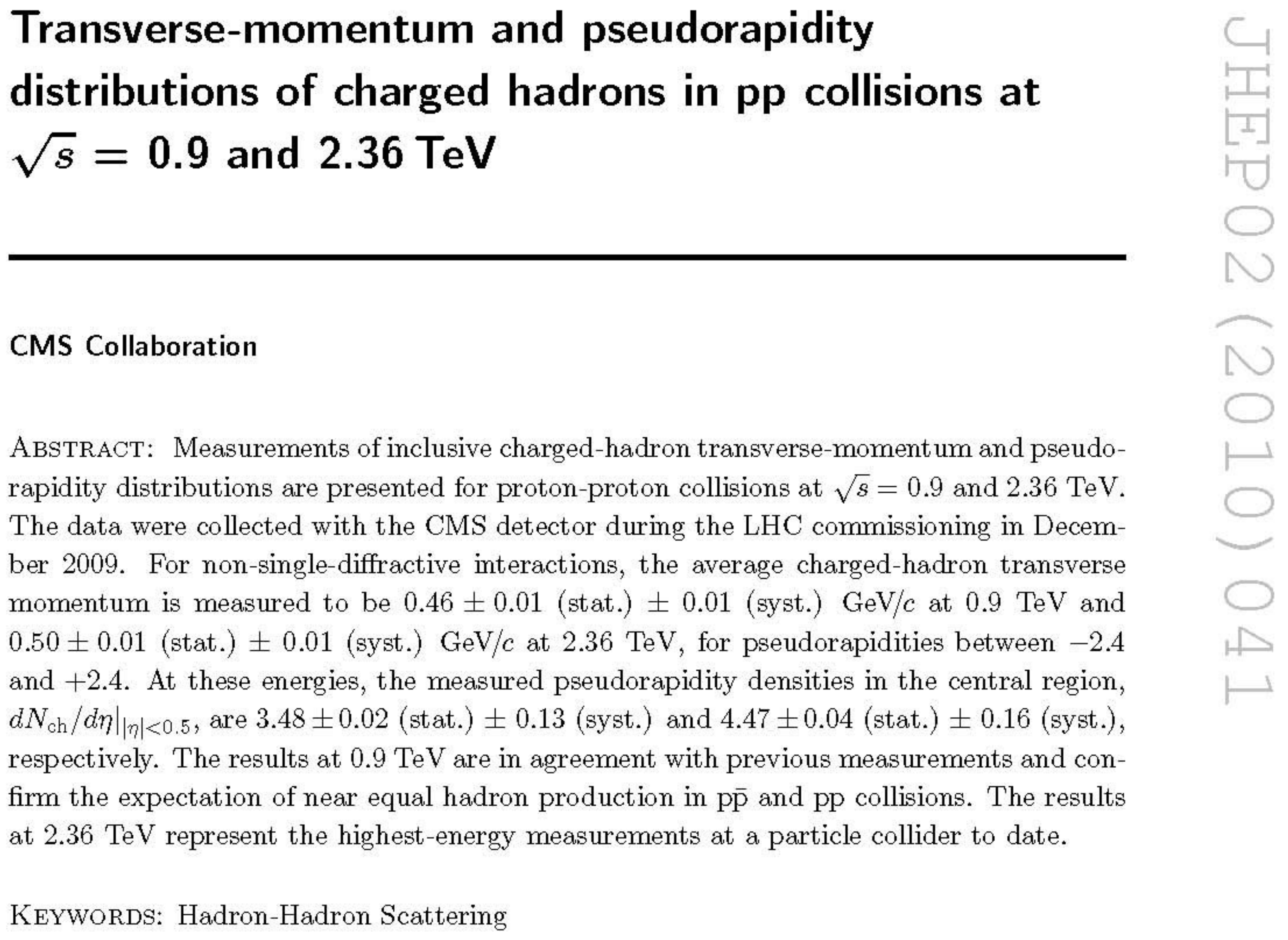
Physics analysis

- Computation of physical quantities ⋯
  - Cross section, branching ratios, masses, lifetimes, ⋯
- ⋯ and of their errors
  - Statistical errors: uncertainty because of limited number of observations
  - Systematic errors: uncertainty because of limited knowledge of key assumptions (beam energy, calibration, alignment, magnetic field, theoretical values, background channels, ⋯)

Analysis tools

- Need versatile tools for
  - Multidimensional selection, event display and interactive reprocessing
  - Histogramming, plotting, fitting of curves and models
  - Point estimation, confidence intervals, limits
- Main tool currently used: ROOT
  - Data analysis and storage, but also detector description, simulation, data acquisition, …

And finally ···



Transverse-momentum and pseudorapidity distributions of charged hadrons in pp collisions at $\sqrt{s} = 0.9$ and 2.36 TeV

CMS Collaboration

Abstract: Measurements of inclusive charged-hadron transverse-momentum and pseudorapidity distributions are presented for proton-proton collisions at $\sqrt{s} = 0.9$ and 2.36 TeV. The data were collected with the CMS detector during the LHC commissioning in December 2009. For non-single-diffractive interactions, the average charged-hadron transverse momentum is measured to be 0.46 ± 0.01 (stat.) $\pm$ 0.01 (syst.) GeV/c at 0.9 TeV and 0.50 ± 0.01 (stat.) $\pm$ 0.01 (syst.) GeV/c at 2.36 TeV, for pseudorapidities between -2.4 and $+2.4$. At these energies, the measured pseudorapidity densities in the central region, $dN_{\mathrm{ch}}/d\eta|_{|\eta|<0.5}$, are 3.48 ± 0.02 (stat.) $\pm$ 0.13 (syst.) and 4.47 ± 0.04 (stat.) $\pm$ 0.16 (syst.), respectively. The results at 0.9 TeV are in agreement with previous measurements and confirm the expectation of near equal hadron production in p$\bar{\mathrm{p}}$ and pp collisions. The results at 2.36 TeV represent the highest-energy measurements at a particle collider to date.

Keywords: Hadron-Hadron Scattering

JHEP02(2010)041

Distributed analysis

- Physics analysis tasks place in many labs all over the world
- Physicists need fast access to event data and corresponding calibration, alignment and bookkeeping data … and to simulated data
- We need the grid!

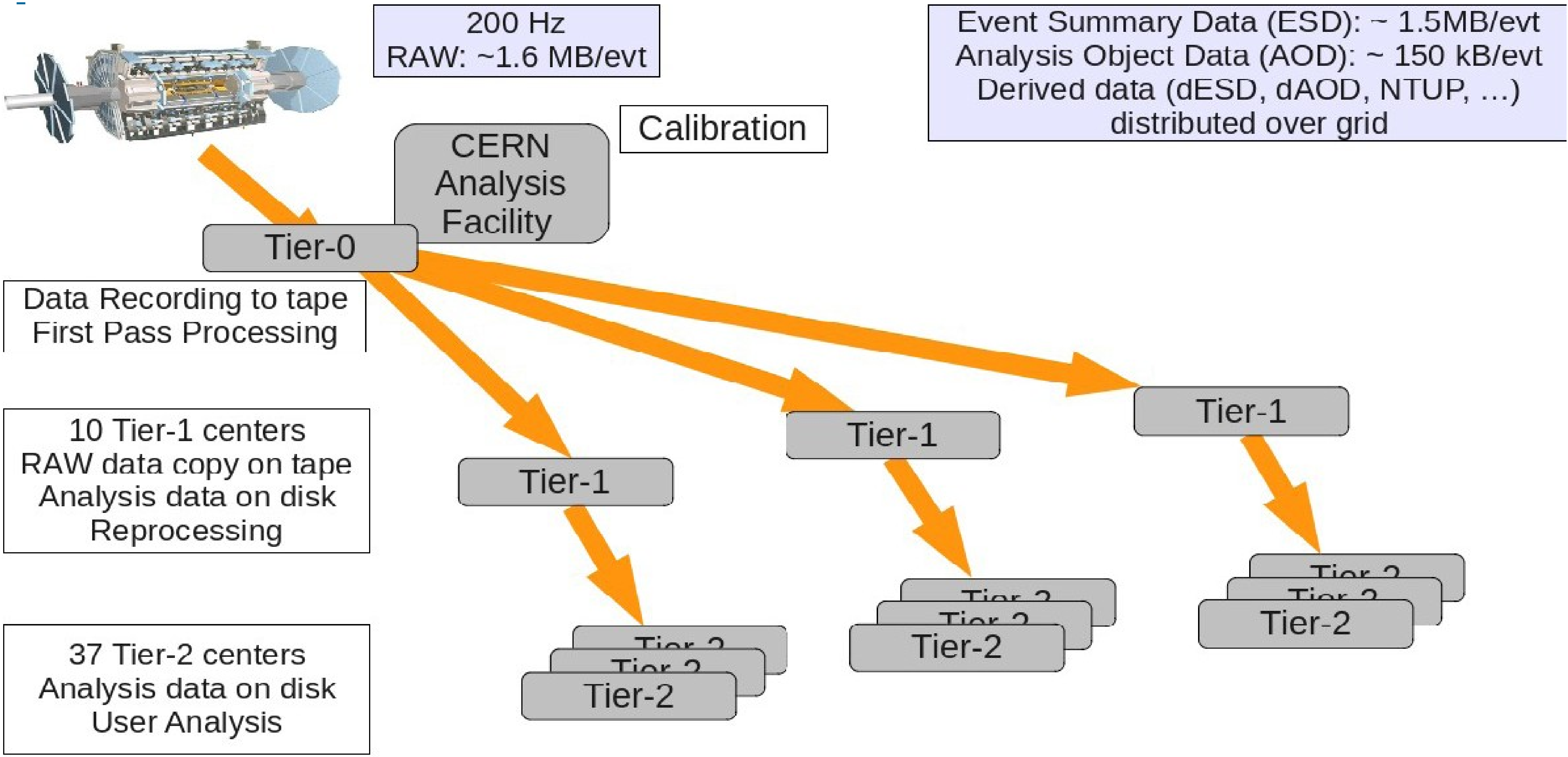
The LHC Computing Grid

- Global collaboration of more than 170 computing centers in 36 countries
- Four-tiered model
- Data storage and analysis infrastructure
- O(10^5) CPUs
- O(100) PByte disk storage (tiers 0 and 1)

Data management

- Dataset bookkeeping
    - Which data exist?
- Dataset locations service
    - Where are the data?
- Data placement and transfer system
    - Tier-0 -> Tier-1 -> Tier-2
- Data access and storage
    - Long-term storage, direct access

Data flow in ATLAS



200 Hz
RAW: ~1.6 MB/evt

Event Summary Data (ESD): ~ 1.5MB/evt
Analysis Object Data (AOD): ~ 150 kB/evt
Derived data (dESD, dAOD, NTUP, ...)
distributed over grid

Calibration

CERN
Analysis
Facility

Tier-0

Data Recording to tape
First Pass Processing

Tier-1

Tier-1

Tier-1

10 Tier-1 centers
RAW data copy on tape
Analysis data on disk
Reprocessing

Tier-2

Tier-2

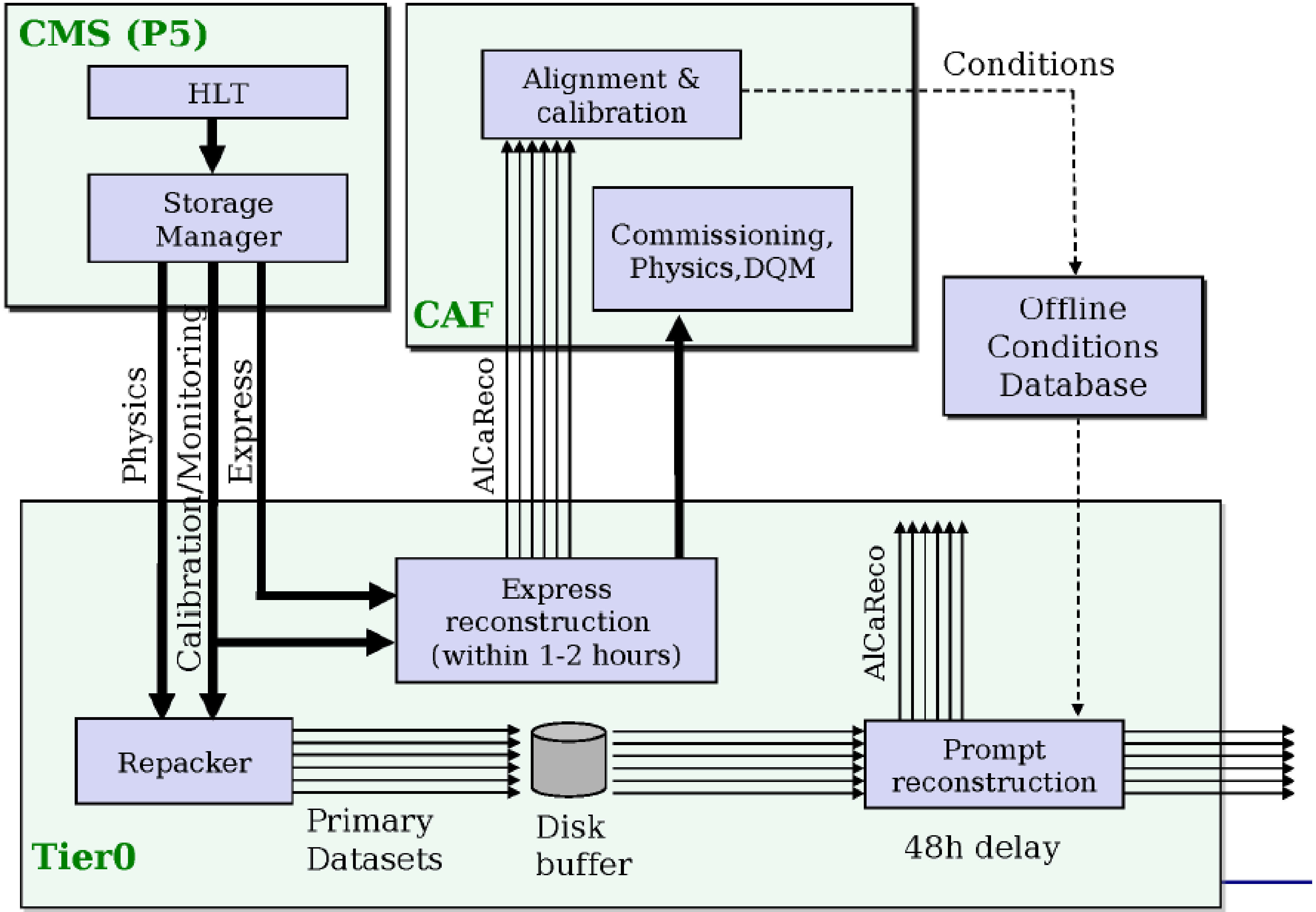
Tier-2

37 Tier-2 centers
Analysis data on disk
User Analysis

Additional resources

- CAF (CERN Analysis Facility)
  - O(100) worker nodes, O(1000) cores (CMS)
  - Ready access to calibration and express streams
  - Fast turnaround
  - Operation critical tasks
    - trigger and detector diagnostics
    - alignment and calibration
    - monitoring and performance analysis
  - Physics data quality monitoring

Data flow in CMS-CAF



Summary

- Physics computing involves:
    - Event filtering with multi-level trigger
    - Storage of raw data
    - Calibration and alignment
    - Storage of calibration and alignment data
    - Event reconstruction
    - Storage of reconstruction objects and metadata

Summary

- Physics computing involves:
  - Simulation of many million events
  - Storage of simulated raw data and truth information
  - Reconstruction of simulated events
  - Storage of reconstruction objects and truth information
  - Distributed physics analysis and event viewing
  - Storage of high-level physics objects

Acknowledgement

- All illustrations in this lecture from the ATLAS and CMS experiments are subject to © 2014 CERN (CERN Copyright)
- The original lecture was given @ CSC 2014 by Prof. Arnulf Quadt
- This lecture was given to you by courtesy of him