QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

From Collisions to Physics:

LHC Data Processing

Dario Barberis
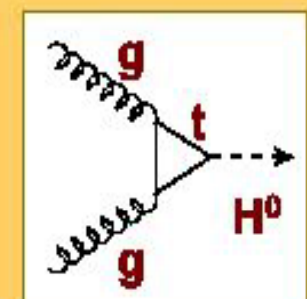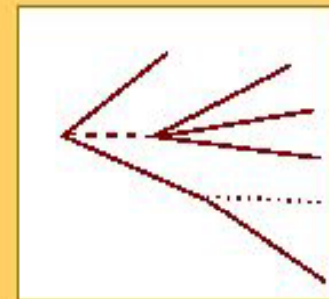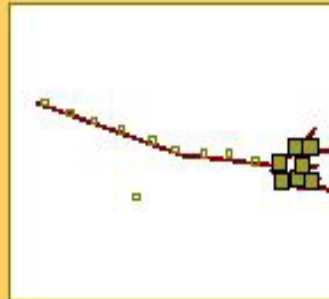
(CERN & Genoa University/INFN)

Events and Data Rates

- Particles produced in LHC interactions generate electronic signals in the detector elements they cross
    - The ensemble of all signals generated at the same time constitutes an "event"
- Every LHC experiment will record a few 100 events/second during active data-taking time
    - Corresponding to several 100 MB/s of data flow from each experiment to the CERN computer centre
- These events have to be stored, processed and distributed to physicists who will analyse them
    - Unprecedented (for HEP) computing resources are needed for this
        - In terms of CPU power, disk and tape storage capacity, and network performance
- The WLCG Grid was set up to support the distributed computing models of LHC experiments
    - In this way all physicists in each experiments have access to the same computing resources, independently of their geographical location

Dario Barberis: LHC Data Processing

2

From raw data to physics

Raw data	Detector response	Interaction with detector material	Fragmentation Decay	Basic physics
Convert to physics quantities	apply calibration, alignment	Pattern, recognition, Particle identification	Physics analysis	Results

Reconstruction

Analysis

Simulation (Monte-Carlo)

From raw data to physics

Raw data

Convert to physics quantities

| | | | |

2037 2446 1733 1699
0003 3611 952 1328
2132 1870 2093 3271
4732 1102 2491 3216
2421 1211 2319 2133
3451 1942 1121 3429
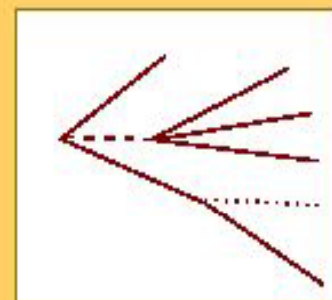3742 1288 2343 7142

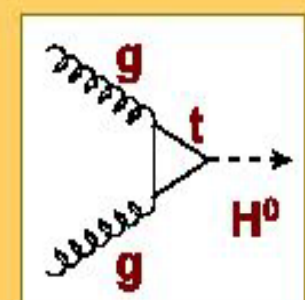**Detector response**

apply
calibration,
alignment

Interaction with detector material

Pattern,
recognition,
Particle
identification

Fragmentation Decay

Physics
analysis

Basic physics

Results

Reconstruction

Simulation (Monte-Carlo)

Analysis

Calibration

Calibration and Alignment

QuickTime™ and a
TIFF (Uncompressed) decompressor
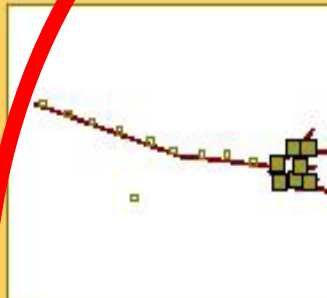are needed to see this picture.

- Calibration constants are the conversion factors between the recorded signal amplitudes and the energy deposited by particles in each detector element

- Alignment constants give the position in space of each detector element

- All these constants may vary with time
  - Depending on the local temperature of the detectors and atmospheric pressure
  - Also on LHC beam conditions (as higher interaction rates produce more activity in the electronics and therefore more heat) and on external factors like the humidity of the rocks surrounding the experiment cavern.

- External and internal measurements are used to derive calibrations and alignments as a function of time
  - Measurements from detector control system sensors
  - Partial reconstruction of samples of events (particle trajectories in magnetic fields, energy matching between different detector sub-systems etc)

- This information is stored in databases that are then accessed during reconstruction and analysis

From raw data to physics

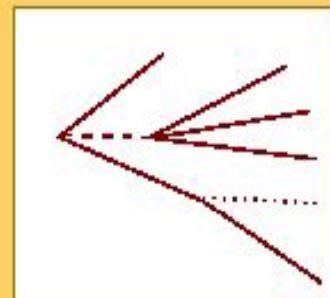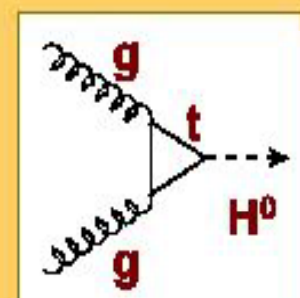|				
2037 2446 1733 1699 4003 3611 952 1328 2132 1870 2093 3271 4732 1102 2491 3216 2421 1211 2319 2133 3451 1942 1121 3429 3742 1288 2343 7142				

Raw data

Convert to physics quantities

Detector response

apply calibration, alignment

Interaction with detector material

Pattern, recognition, Particle identification

Fragmentation Decay

Physics analysis

Basic physics

Results

Reconstruction

Analysis

Simulation (Monte-Carlo)

Reconstruction

Event Reconstruction

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

- First level:
  - Central tracking and muons:
    - Reconstruction of charged particle trajectories
    - Identification of the primary collision position and of interaction and decay vertices
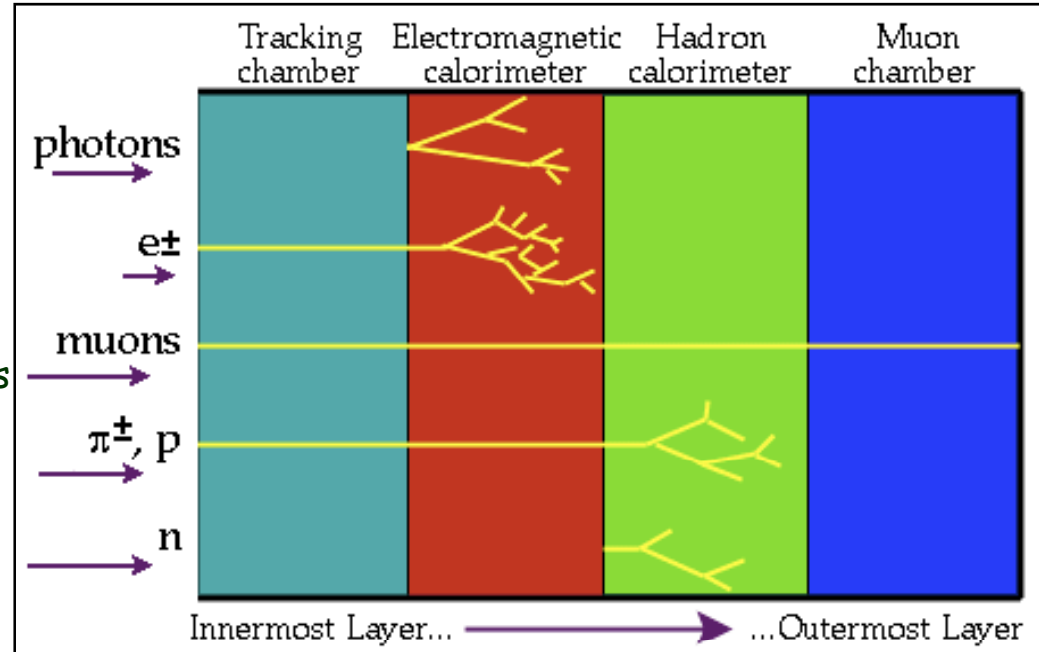 - Electromagnetic calorimetry:
    - Reconstruction of electromagnetic showers
  - Hadronic calorimetry:
    - Reconstruction of particle jets
    - Identification of isolated hadrons
- Second level:
  - Identification of electrons, photons, muons
  - Determination of the origin of jets: quarks u/d/s, c, b, gluons, τ
  - Measurement of transverse energy balance and other global event quantities

Tracking chamber | Electromagnetic calorimeter | Hadron calorimeter | Muon chamber

photons

e±

muons

$\pi^\pm$, p

n

Innermost Layer... ⟶ ...Outermost Layer
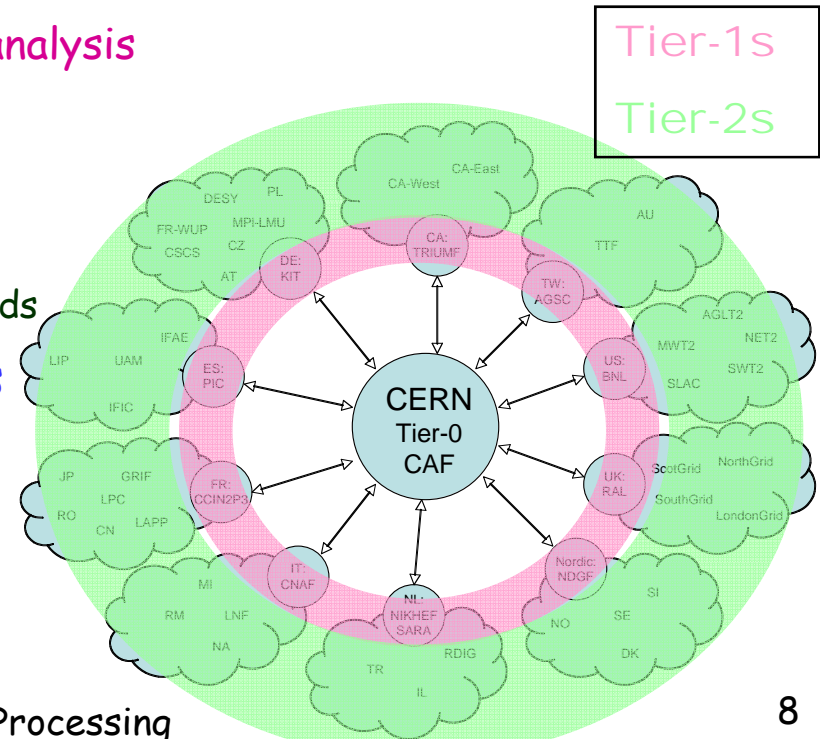
Dario Barberis: LHC Data Processing

7

Data Distribution

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

- We all use the WLCG grid infrastructure to distribute events world-wide

    - Raw and full reconstruction records go to larger computer centres ("Tier-1 centres")

        - If/when better calibration and alignment constants are available, reconstruction processing can be re-run there and output data distributed further for physics analysis

    - Summaries of event reconstruction are distributed also to smaller computer centres ("Tier-2 centres") for analysis

    - Peak export data rates from CERN for the 4 experiments amount to almost 2 GB/s during data-taking periods

        - Equivalent to one full DVD every 2.5 seconds

- Databases containing detector conditions information are also distributed to the Tier-1 centres using WLCG tools
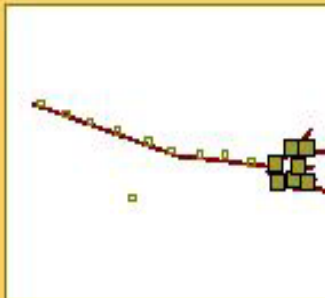
Tier-1s
Tier-2s

From raw data to physics

Raw data

Convert to physics quantities
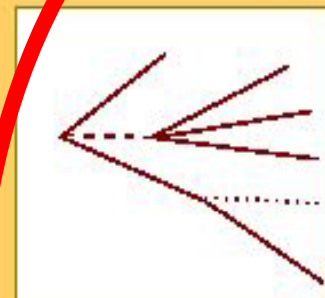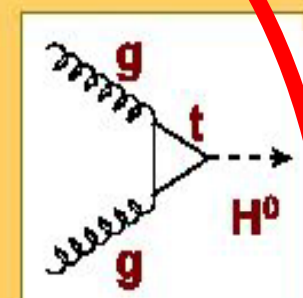
Detector response

apply calibration, alignment

Interaction with detector material

Pattern, recognition, Particle identification

Fragmentation Decay

Physics analysis

Basic physics

Results

Reconstruction

Analysis

Simulation (Monte-Carlo)

Analysis

Physics Analysis

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

- We have set up a highly automated system to process and distribute data around the world
  - The WLCG Grid infrastructure is one of the building blocks of this system
    - Together with the experiment software and operation tools
- Individual physicists and organised analysis groups use routinely Grid tools to access data and run jobs world-wide for their analysis tasks
  - Next talk
- This system would not have come together without the dedication and hard work of several hundred people who helped building, deploying and operating it over the last 6 years
  - Our thanks go mostly to them!