Use of Grid at CMS: Software validation and detector alignment

Gregor Bruns, University of Leipzig

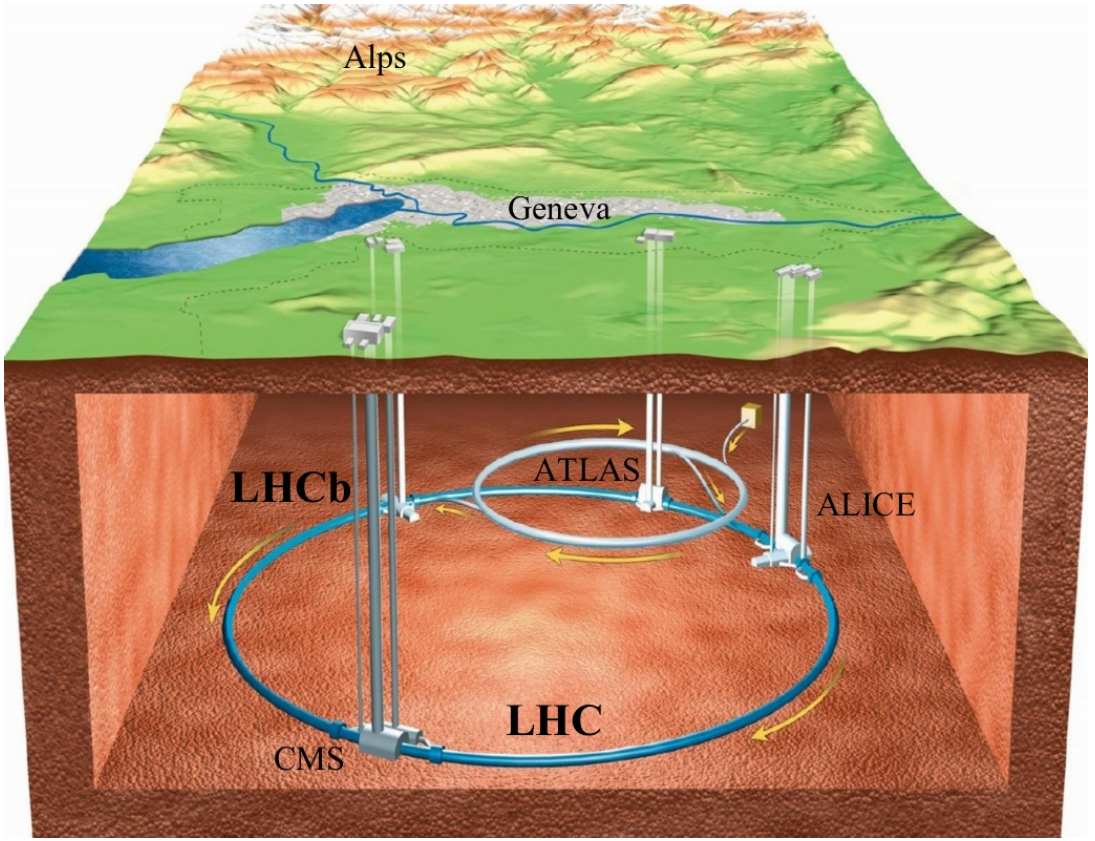
10.09.09 / Students Session

Outline

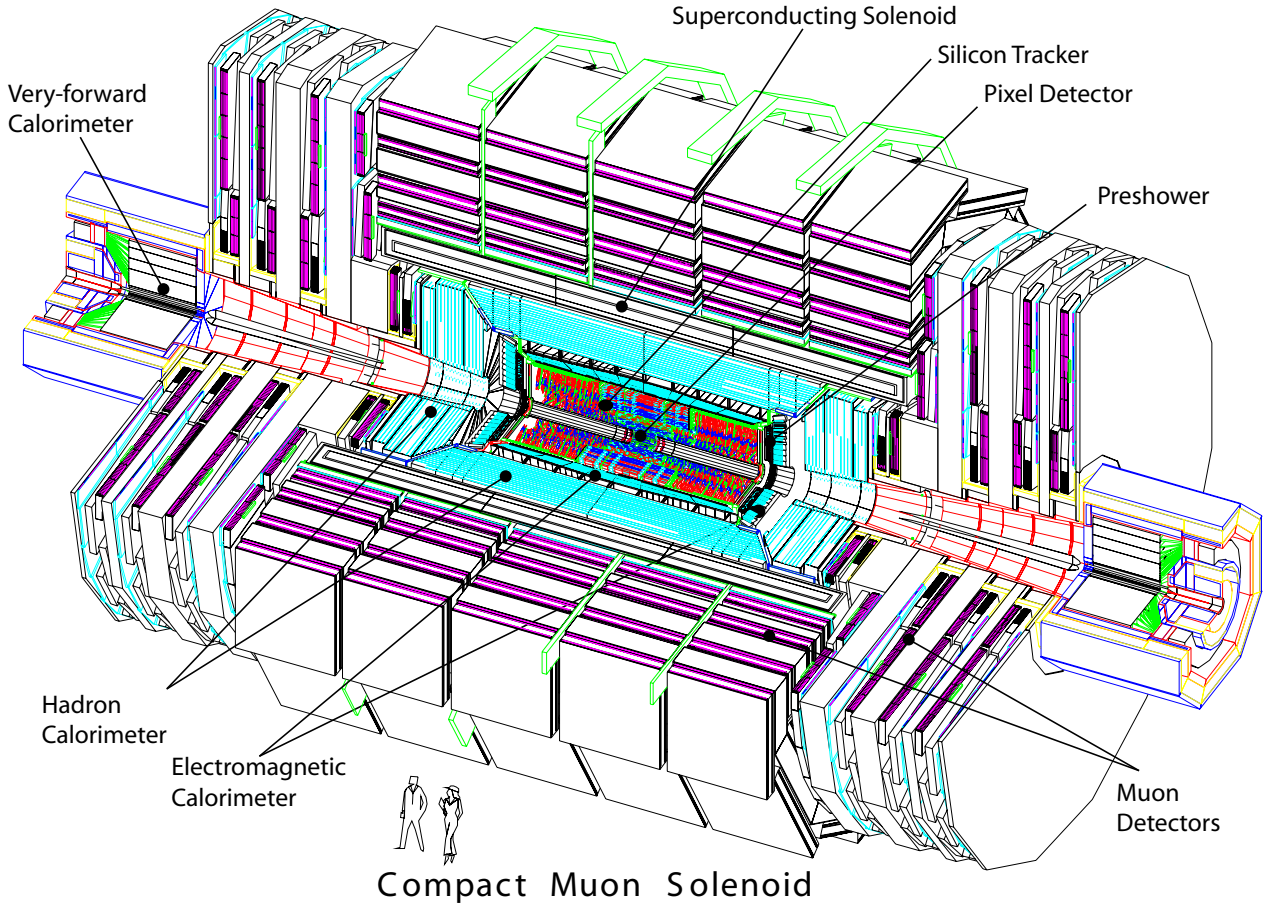
Outline

The CERN LHC

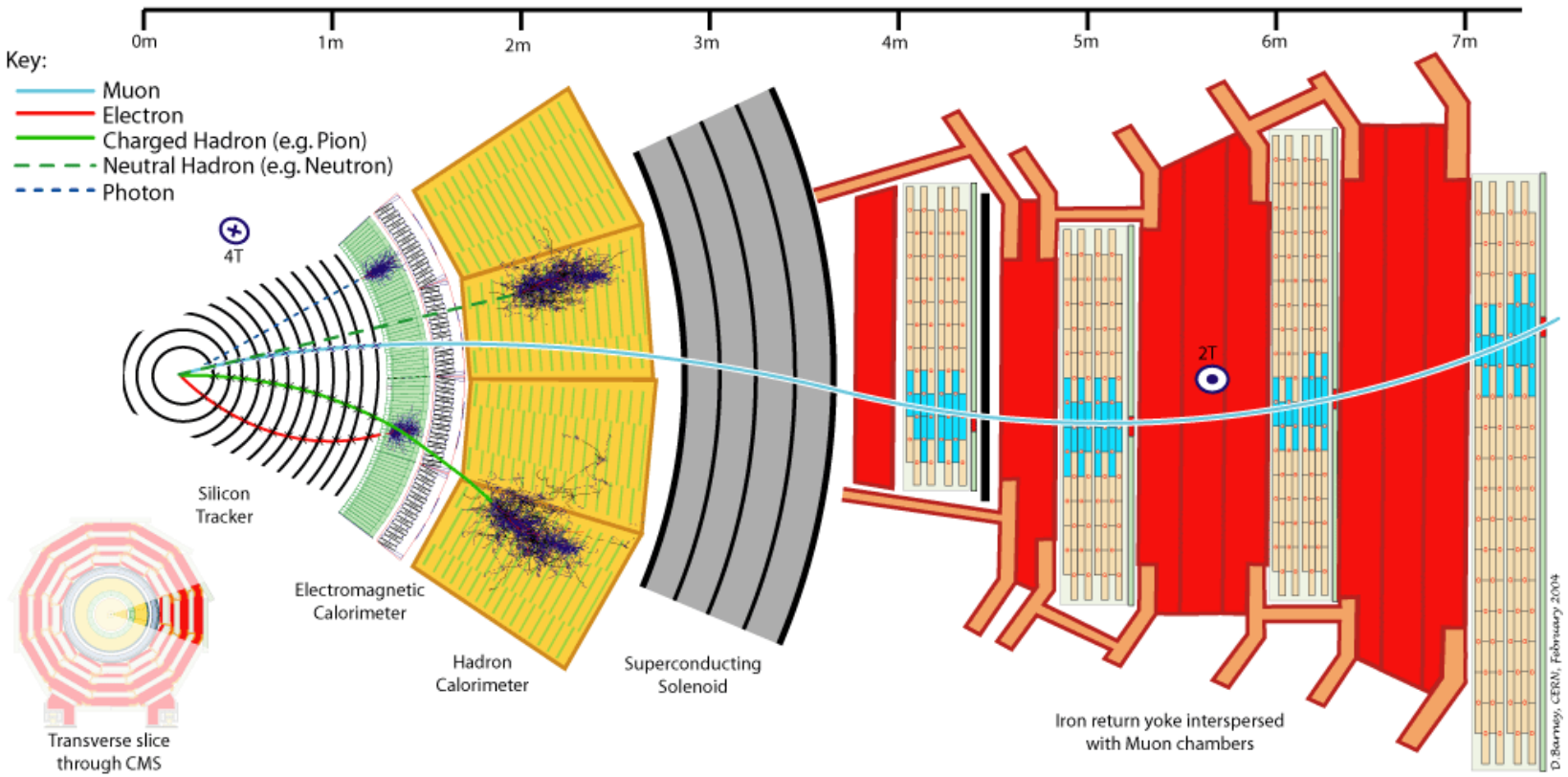
- 27 *km* tunnel
- *pp* and lead ion collider
- Superconducting magnets up to 8.6 *T*
- Max. CoM energy $E = 14\ TeV$ (pp)
- Inst. luminosity $L = 2 \cdot 10^{33}\ cm^{-2}s^{-1}$



Detector Design



A Slice of Detector



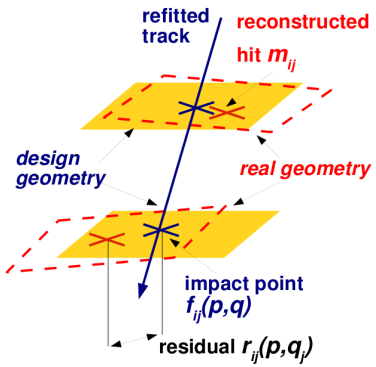
Why Alignment?

- Tracker precision up to a few μm
- Hardware installation of parts only with precision $\approx 100\ \mu m$
- Correction of component positions required

- Detect displacements
  - due to hardware installation
  - while detector is working (temperature shift etc.)
- Different methods in use:
  - Detector positioning survey (direct measurement)
  - Laser alignment system (dedicated hardware)
  - Track based alignment (use analysis data)

Alignment Procedure for Track Based Alignment

- High number of particle trajectories used
- Reconstruct track from detector hits

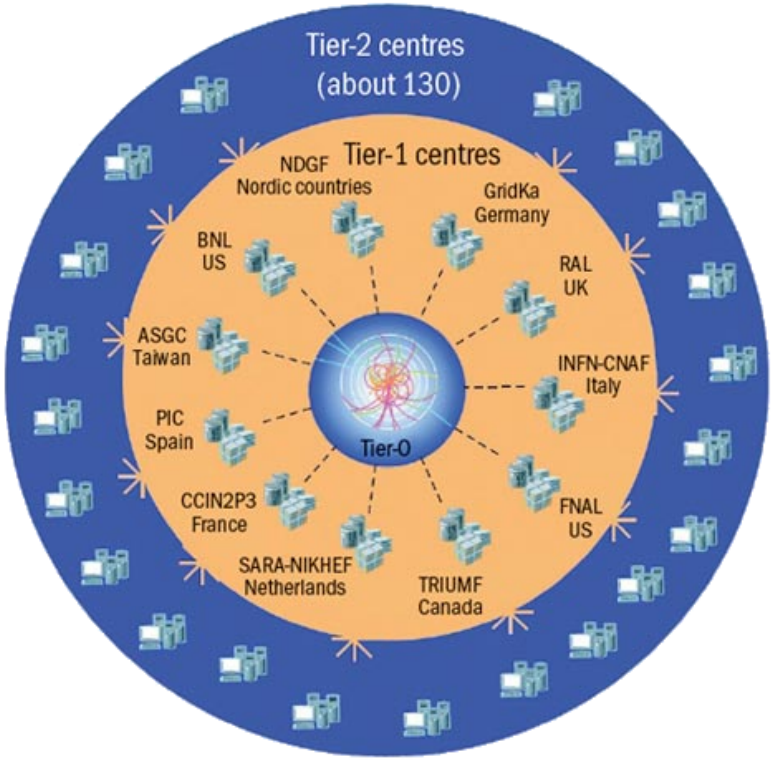


- Compare measured and expected hit positions
- Determine shift of detector modules
- Validation of the procedure on simulated data
- Perform real alignment on cosmic muon data

The Grid Network

- Why do we need a global computing network?
  - CMS will produce $\geq$ 3 PB of data per year
  - MC simulations 10 times the amount of measured data
  - Local CERN facilities cannot provide necessary storage and CPU power
- How are these problems solved?
  - Mass storages at each facility
  - Scheduling systems for submitting jobs to computing farms
  - Global access possible
    - Distributed computing
    - Easy data access

The Grid Network



- Tier-0: CERN
- Tier-1: National centers
- Tier-2: Regional groups
- Tier-3: Institutes
- Tier-4: Workstations

CMS Remote Analysis Builder (CRAB)

- Part of the CMS SoftWare package (CMSSW)
- Gets location of data
- Fetches information on the job
- Manages distribution of jobs on the Grid
- Handles storage of files for the user

Outline

MillePede Production System (MPS)

- MillePede: Algorithm for track based alignment
- MPS: Automatization of alignment steps
  - Millions of traces are required
  - Splitting and parallelizing of huge jobs
  - User-friendly interface
  - Runs on local CERN queues

MPS with CRAB

- CRAB submits jobs to whole Grid
- Information on jobs obtained automatically
  - Automatic splitting of jobs
  - Automatic locating of input files
- But: Datasets not available everywhere
  - Request transfer of data needed
  - This will change when LHC is running
- Configuration now:
  - Run only on local CERN queues
  - Supply information where data is stored
- Can be changed easily in the future

Outline

CMSSW Releases

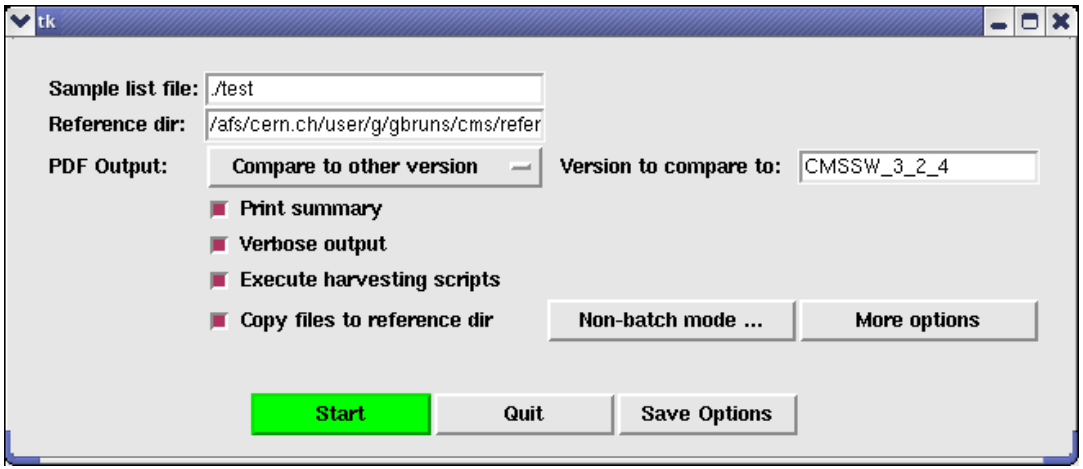
- Consider new CMSSW release
  - Errors may be introduced
  - Analysis may be more accurate
- Check results by comparison of results from physical processes
- Standardized samples are evaluated
- Compare standard plots from previous releases

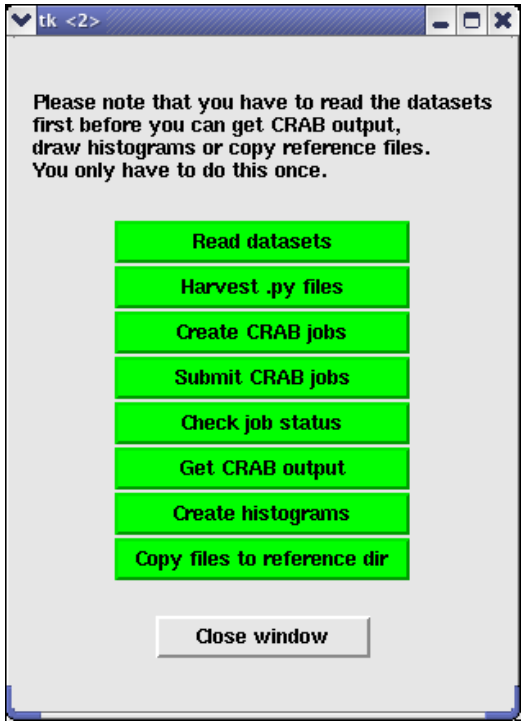
Automatization of the Validation Procedure

- Python script for automatization
  - Input: List of data samples
  - Output: Plots comparing CMSSW versions
- Usage of CRAB for retrieving samples
- Automatic search for histograms in ROOT files
- Little dependent on file name conventions

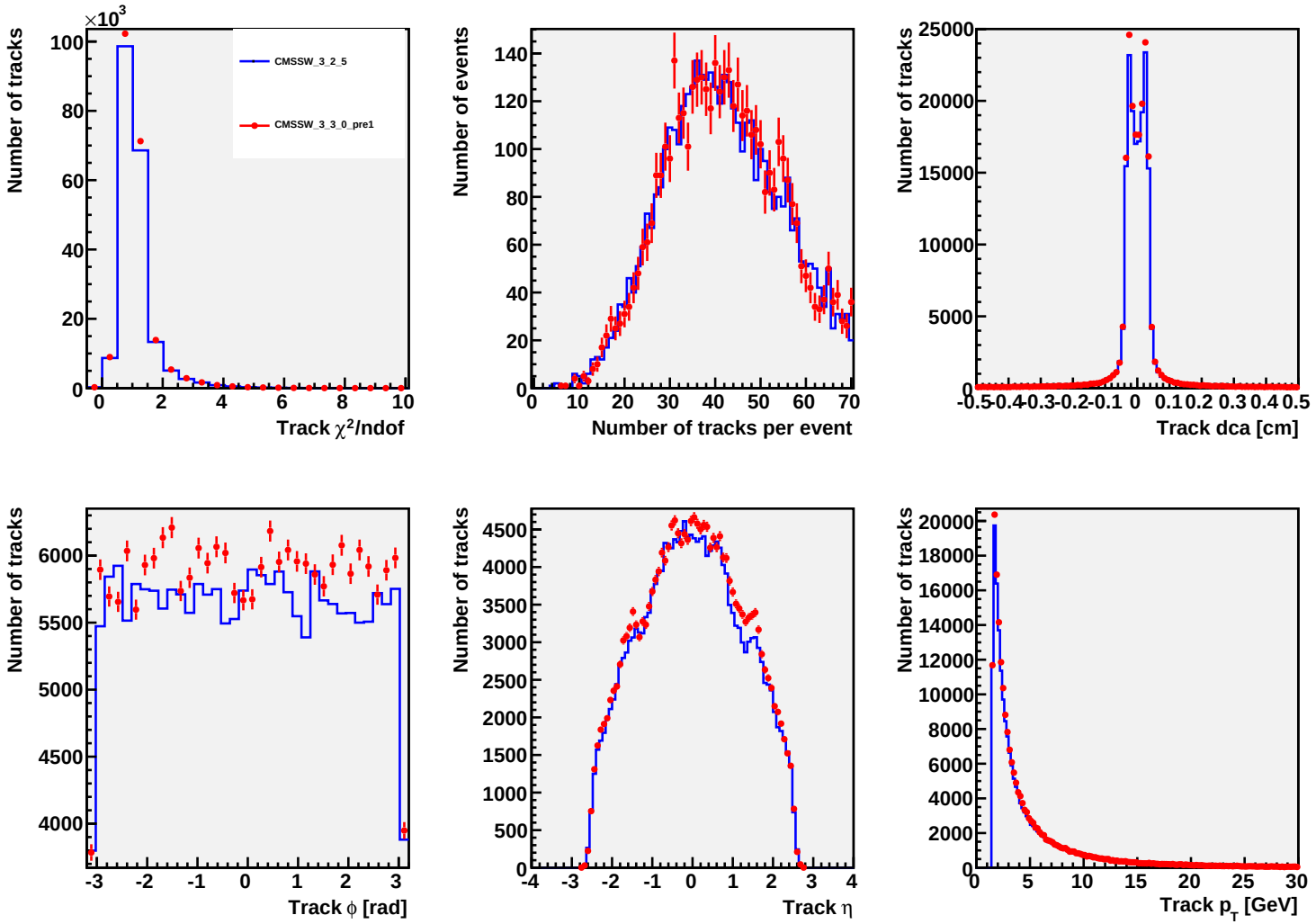
A GUI for more Usability

- Written in Tkinter (Python's Tk wrapper)
- User enters options and executes scripts
- Configuration is stored in a file
- Batch mode or individual steps





Comparison of Plots



Summary

- Results:
  - Release validation has been simplified (GUI)
  - MPS scripts can use CRAB commands
- Outlook:
  - Validation script might be used for other purposes
  - Maybe CRAB can take over some MPS tasks

Acknowledgement

I would like to thank

- My supervisors, Andrea Parenti and Justyna Tomaszewska
- Gero Flucke for providing explanations
- All people who have made this stay possible