LHCb Computing Model and some thoughts about long term data preservation

Thomas Ruf

- The Experiment
- Data Flow
- Software Organization
- Data and Analysis Preservation, Some Thoughts

Disclaimer

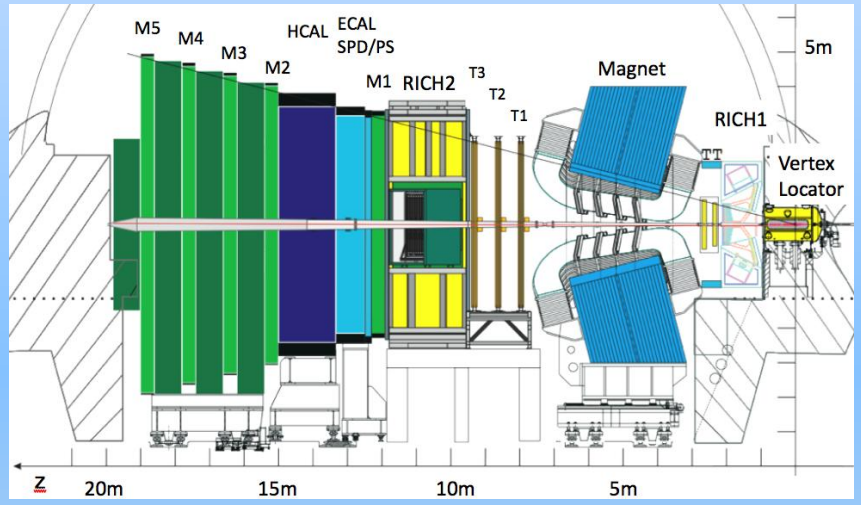
- This is not a presentation on a LHCb data preservation model.
- Discussion in LHCb just started, mainly triggered by this workshop.
- Hope to learn from this workshop and profit from the experience of other experiments.

LHCb Experiment

- **LHCb is a precision experiment to measure CP violation in beauty and charm sector and search for rare decays of heavy mesons exploring a forward spectrometer at the LHC using pp-collisions.**

- **Data taking started December 2009.**
  - Phase I until ~2017
  - Phase II (upgrade) up to ~2025

- **Collaboration, ~750 people from 53 institutes in 15 countries**

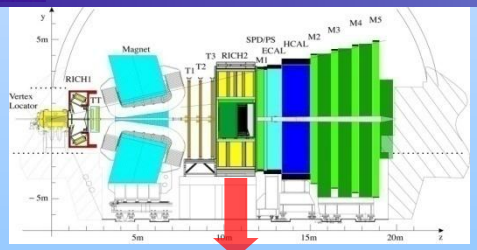


http://LHCb.cern.ch

LHCb Dataflow and Data Reduction

- 40 MHz input rate to Level 0, hardware trigger
  - Select events with large p_T, efficiencies between 30% (hadrons, photons) and 95%(muons), also trigger lines for calibration, prescaled, noBias, lowMult, …
- 1 MHz input rate to Level 1, High Level software Trigger
  - ~25000 processes running in parallel
  - Event reconstruction software almost identical to offline version, some simplifications to fit in CPU budget
- 3 kHz output rate to mass storage, ~50kB / event, ~1PB / year (for one copy)
  - Offline reconstruction using 7 Tier1 centers
- Instantaneous offline processing, + 1 major reprocessing / year
  - Offline reconstruction including creation of analysis objects at 7 Tier1 sites
  - Further reduction of data, either by reducing events, 200Hz full information, ~150kB/event or ~1kHz of μDST ~10kB/event, **permanently on disk at Tier1s**
  - Individual users cannot access the complete set of raw data, stored on tapes
  - Offline selection can be re-done 2-3 times / year with improved or new analysis code, takes about 1 month, requires team of people to follow production at Tier1 sites and to react to problems.
  - (maybe in 10 years from now, working with PBs of data will become much easier)

Data Flow



40 MHz

Level -0

e, γ | had | μ

hardware

1 MHz

High-Level Trigger

Close to offline reconstruction, inclusive selections

25k parallel processes

3 kHz

Storage: Event size ~50kB

Offline reconstruction 7 Tier1

~1-2 days latency

Selection of events for different analysis streams

Copies at different Tier1

~10%

Input to analysis:
Full DST 150kb/event
Micro DST 10kB/event
Ntuples

- **End of year re-processing for coherent set of data**

Software Organization

- **Software management based on cmt and svn**
  - Core framework: Gaudi (developed inside LHCb, adopted by Atlas and other non-LHC experiments), data persistency ROOT
    - SLOC = ~110k, ~350k
  - use of external packages: Boost, GSL, Oracle, ...
    - SLOC = ~2M
  - interface to the Grid, bookkeeping, ... core: DIRAC, user entry: Ganga
    - SLOC = ~310k
  - core LHCb packages for reconstruction, HLT selections, Stripping selections, analysis tools, detector description
    - SLOC = ~350k
  - Simulation: several event generators, interface to Geant4, emulation of digitized raw data
    - SLOC = ~210k
- **Databases: Oracle, sqlite**

SLOC = Source lines of code

Software Organization, cont.

- **Almost all external software packages are shared with other LHC experiments, like ROOT, GEANT4, PYTHIA, EVTGEN, …**
- **Assume maintenance of these packages handled outside LHCb.**

Data and Analysis Preservation, Some Thoughts

- **LHCb is a precision experiment, at the end of lifetime, will be limited in many topics by systematic errors,**
  - understanding of detector effects, differences of reconstruction efficiencies for positive and negative charged particles down < 10^{-4}, biases due to tight selections, ...
- **LHCb lifetime:**
  - Phase I, 2009 - 2017 (data taking)
  - Phase II (upgrade to 40MHz readout, new detectors), 2019 - 2025 (data taking)
- **Three types of data used for any analysis: signal, calibration channels, MC simulation**
  - all analysis are done in groups of people, from 3-4 to 20-30 to cover different aspects, like, efficiency corrections, particle ID calibrations, detector asymmetries, detector resolution, …
  - same data used by many different analysis

Re-running old software

- **Old Question discussed related to the online High Level Trigger reconstruction and selection package**
  - Can we re-run the same version of the software as used during data-taking in the life of the experiment ?
  - Answer: Would require enormous efforts and manpower
  - Solution: Provide HLT summary banks which allow unique identification what caused the event to pass the trigger together with a tool to match the online and offline objects.

  - In general, no support for re-running old software versions. Processed data should be self consistent, should contain all information needed without the necessity to rerun the software which produced it.

Data and Analysis Preservation, cont.

- **Any new analysis would require to keep the LHCb computing machinery running**
  - Most straightforward case
    - All cross checks and efficiency corrections can be done with calibration channels: would require to keep the latest offline selection, analysis tools, bookkeeping, grid tools running.
  - Most demanding case
    - A novel analysis which requires new MC simulations. Would also require to run older versions of the high level trigger selections, the offline reconstruction and stripping selections, in order to match the recorded data.
- **Some first thought: From operational point of view, at end of Phase II, produce one set of data with as much as possible abstract physics objects, like identified particles, 4-vector + particle ID, trigger and selection summary.**
  - Keeping the LHCb grid-computing machinery alive, reconstruction / stripping / bookeeping, seems to be extremely difficult more than ~2 years after end of data taking
  - Complex system, needs continous running to stay robust. Expertise and manpower will disappear after end of data taking.
  - Possibility: Invest in a generic distributed computing framework which allows to run also the LHCb workflows ?

Data and Analysis Preservation, cont.

- Keeping computing operational up to 5 years after data taking would require 2-10 FTE/year. Current operation, ~10 FTE/year including dedicated people at Tier1 sites.
- For Phase I probably not so big problem, since large overlap in computing with Phase II. However, detector changes, raw data format changes, …
- Starting new analysis from raw data will be a major work, only possible with very experienced people.
- Which future analysis are made impossible by being only able to work on already reconstructed objects ? Difficult to say.

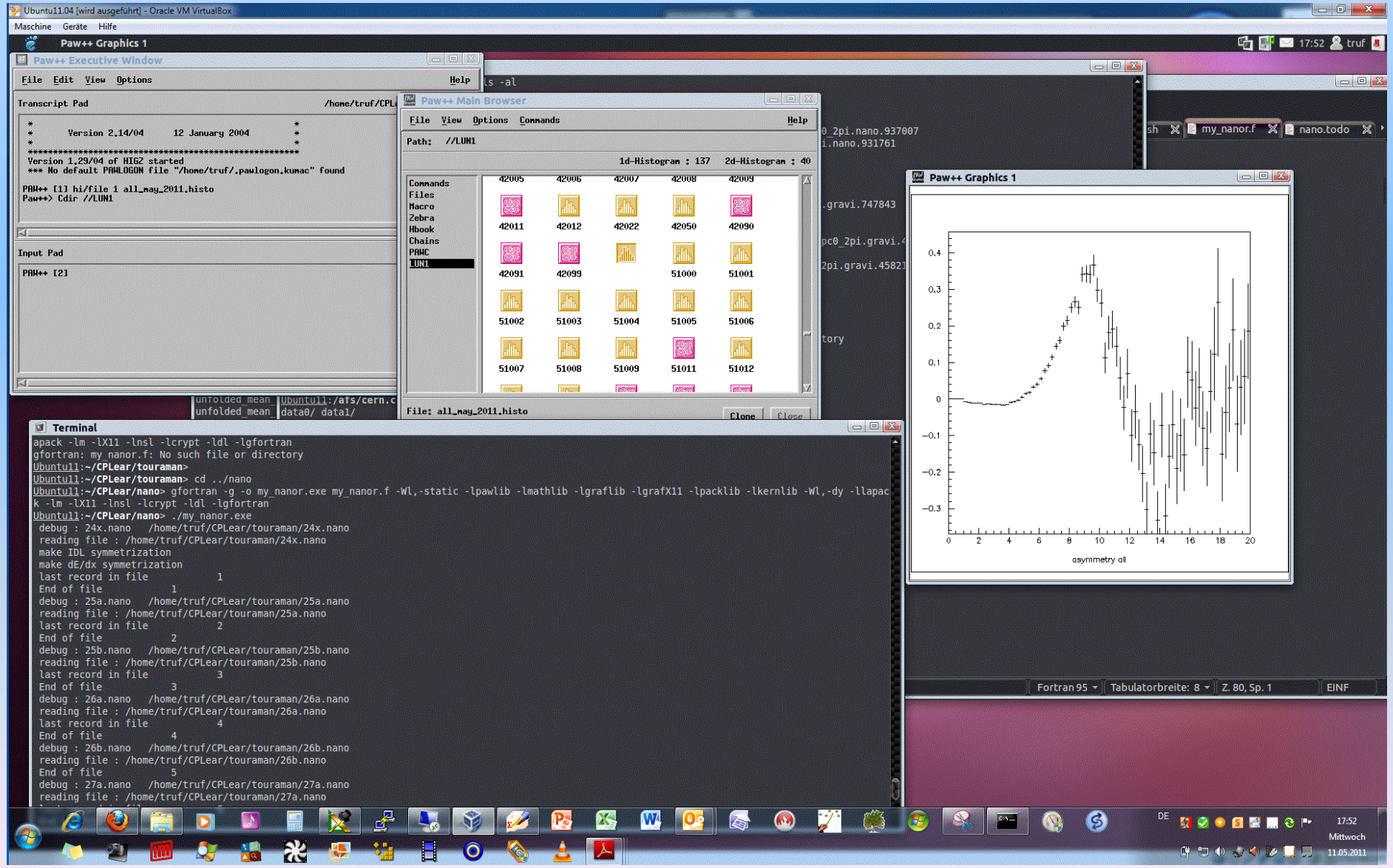
Already a use-case ?
For discussion in LHCb

- **LHCb phase I is supposed to collect 5fb^{-1}, recorded in 2010: 0.04fb^{-1}. One might think, no need to make big efforts to preserve 2010 data.**
- **HOWEVER:**
  - Large statistics of quasi unbiased data with no pile-up, due to low luminosity at startup
  - Unique samples of data taken at 0.9TeV, 2.76TeV and 7TeV center of mass energy
  - However, not mainstream physics, soft-QCD, underlying event, diffraction, …
- **Need to make sure that this data can be analyzed in the future:**
  - Despite possible changes in data formats
  - Preserve detector response calibration
  - Explanation of trigger setttings, prescales, etc.
  - Machine information, bunch filling scheme, bunch currents, vacuum conditions, …

Past Experience: CPLEAR

- **Precision and unique measurements of T and CPT violation in the neutral Kaon system**
  - Textbook plots of CP violation, ~10^9 events, $p\bar{p} \rightarrow K^0\pi^+K^-$ and $p\bar{p} \rightarrow \bar{K}^0\pi^-K^+$
- **Data taking: 1992 - 1996**
  - All raw data at CERN destroyed in 2001, space problem, 100k 3480 cartridges, order of 10 TB data
    10 years later, nobody discusses about problems storing 10TB
  - Leftover: histograms which entered into publications and thesis, some nano-dst files.
  - Log books from data taking, minutes of analysis meetings. Not in electronic format. What should be done with these ?
  - Interesting to note that, reading nano-dst with a fortran program and producing histograms still possible. PAW/HBOOK, available for example in standard Ubuntu distribution.

Unfreezing of CPLEAR Data



Summary

- **Discussion in LHCb about data preservation just started. Needs to be done together with preservation of analysis tools.**
- **New and good ideas needed. Hope to learn from other experiments**
- **Unlike in the past, when an experiment had a successor producing&analyzing much more data, the forward pp data of LHCb might be unique for a long time. Statistics in B_s decays will never be exceeded by any e^+e^- B-factory.**
- **Costs for data and analysis preservation are largely due to manpower. Should be seen in the view of building or restarting a machine like LHC together with a detector. Needs to be recognized by the funding agencies&institutues and included in their manpower planning.**

Acknowledgment: Very useful discussion with
Marco Cattaneo, Peter Clarke and Pere Mato

BACKUP

Data Persistency

- **Raw data:**
  - “sequential” primitive file of Integers and Doubles, no C++ classes
- **Reconstructed data:**
  - Based on ROOT. Dictionary of C++ classes.