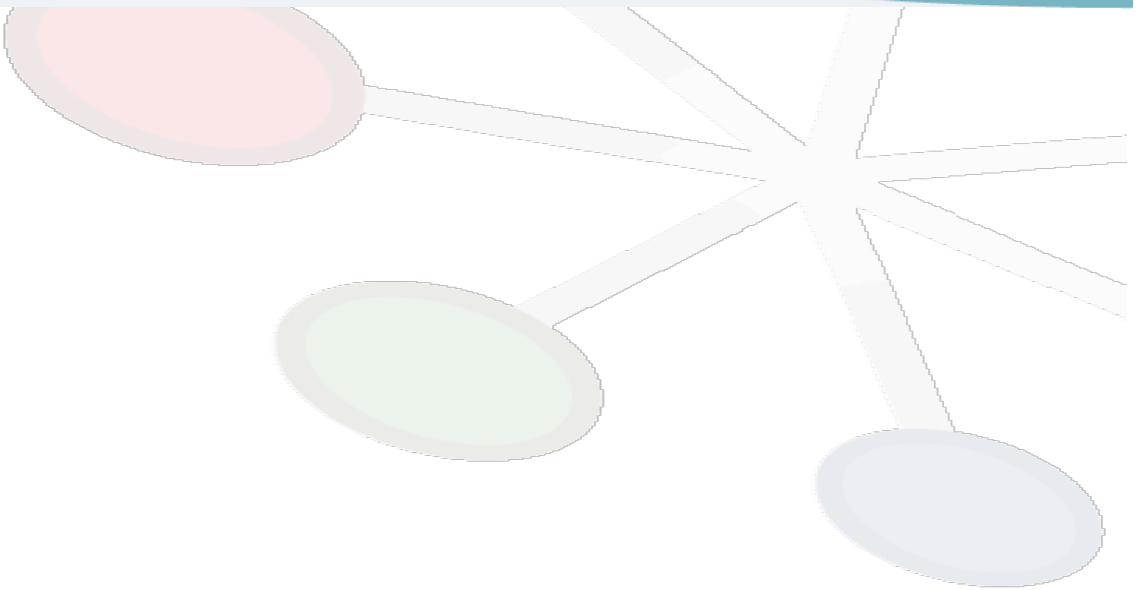
LHCb 2009-Q4 report

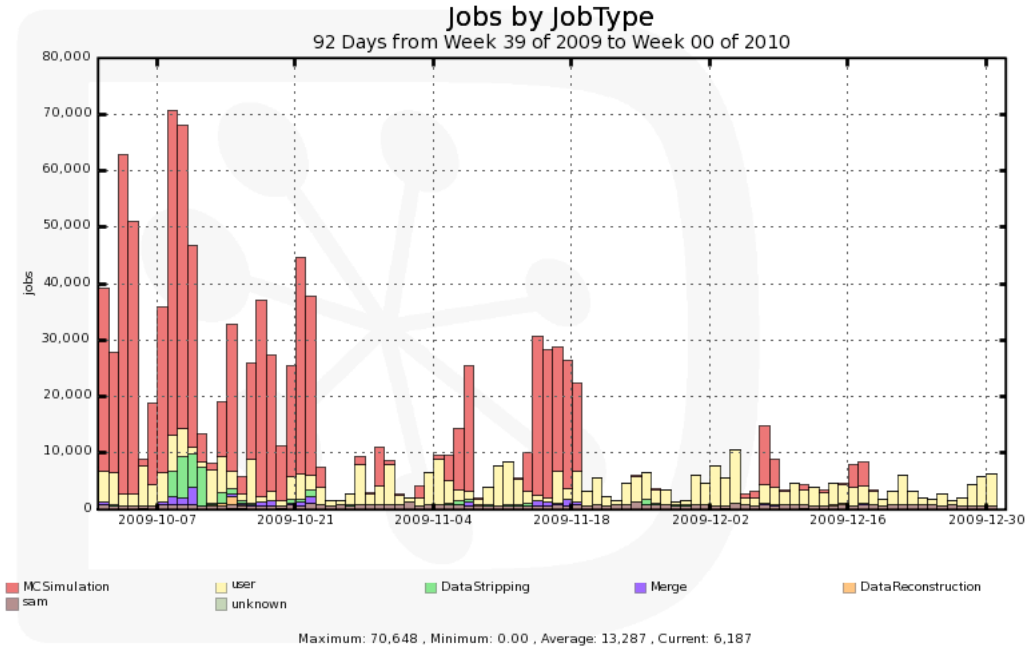


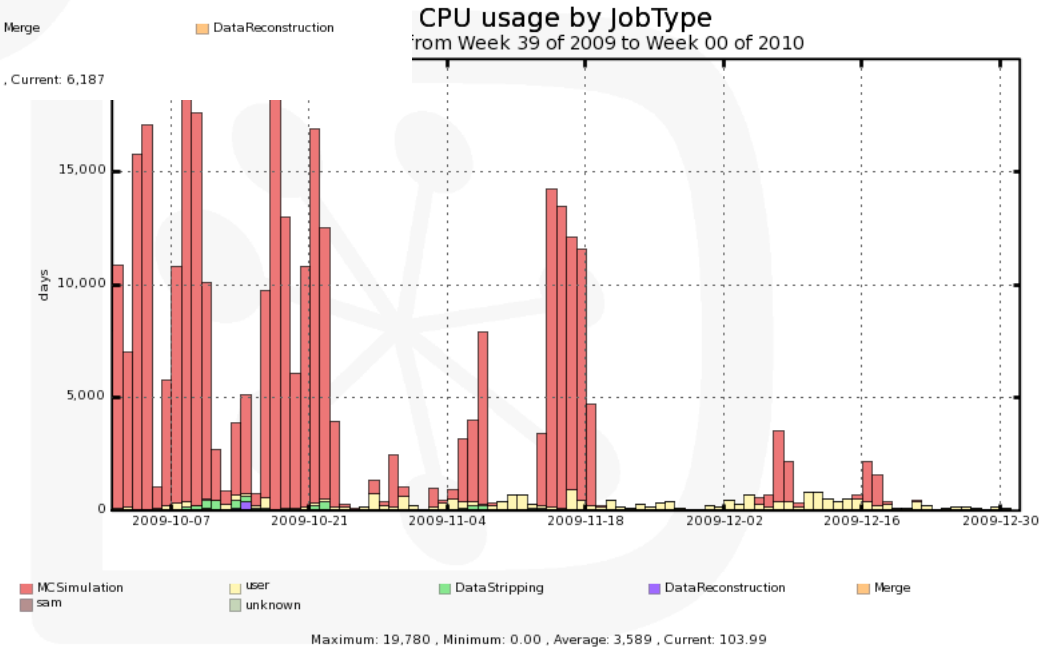
Activities in 2009-Q4

2009 Q4 REPORT

- Core Software
  - Stable versions of Gaudi and LCG-AA
- Applications
  - Stable as of September for real data
  - Fast minor releases to cope with reality of life…
- Monte-Carlo
  - Some MC09 channel simulation
  - Few events in foreseen 2009 configuration
  - Minimum bias MC09 stripping
- Real data reconstruction
  - As of November 20th

Jobs in 2009-Q4

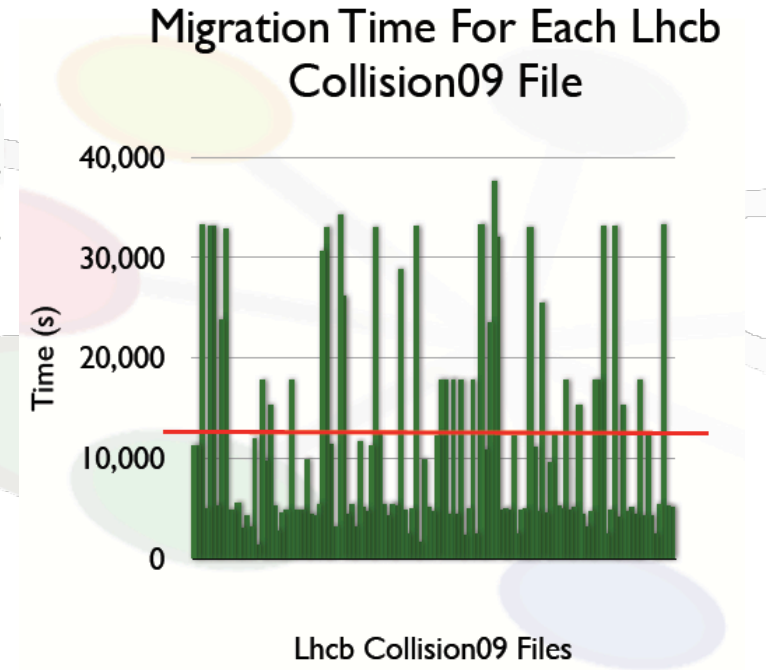




First experience with real data

- Very low crossing rate
  - Maximum 8 bunches colliding (88 kHz crossing)
  - Very low luminosity
  - Minimum bias trigger rate: from 0.1 to 10 Hz
  - Data taken with single beam and with collisions

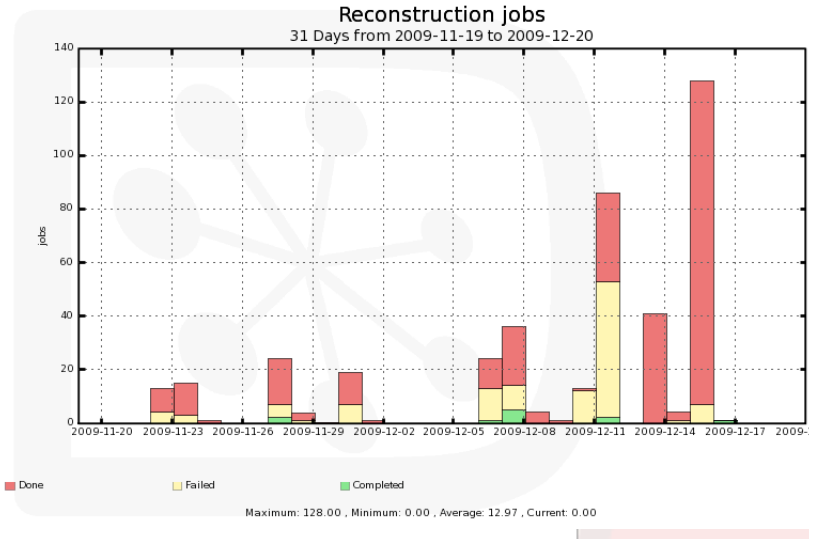
	Number of RAW Files	Total File Size
Lhcb Beam1	36	0.4 GB
Lhcb Collision09	120	217.4 GB

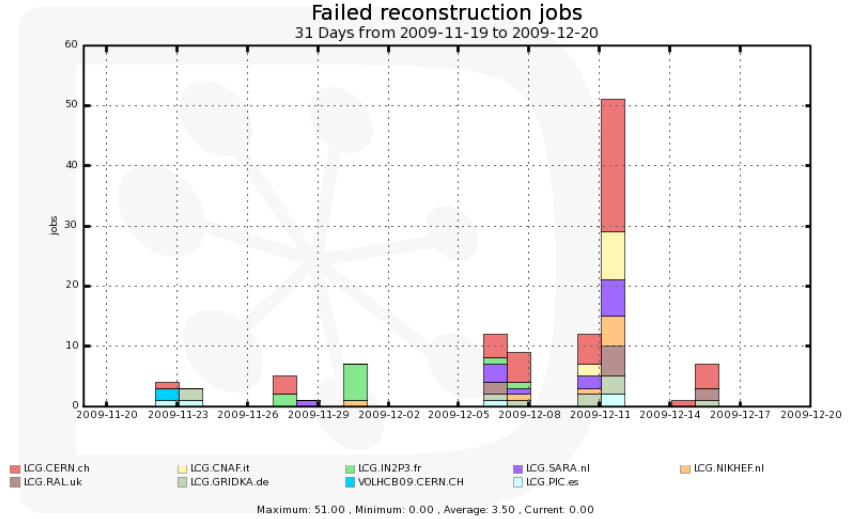


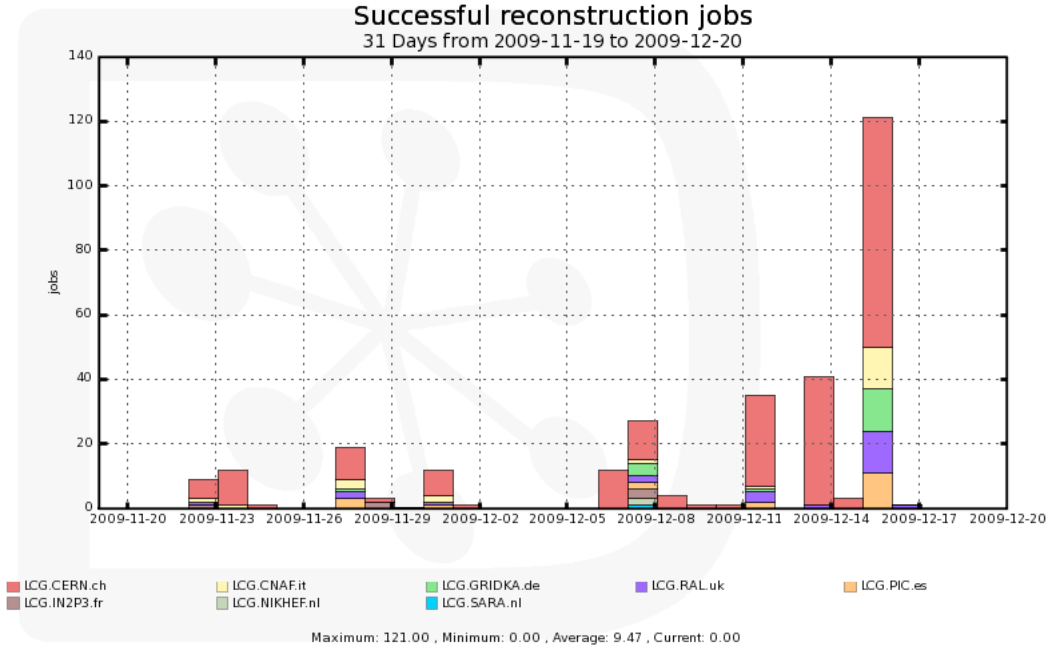
Real data processing

- Iterative process
  - Small changes in reconstruction application
  - Improved alignment
  - In total 5 sets of processing conditions
    - Only last files were all processed twice
- Processing submission
  - Automatic job creation and submission after:
    - File is successfully migrated in Castor
    - File is successfully replicated at Tier1
  - If job fails for a reason other than application crash
    - The file is reset as "to be processed"
    - New job is created / submitted
  - Processing more efficient at CERN (see later)
    - Eventually after few trials at Tier1, the file is processed at CERN
  - No stripping ;-)
    - DST files distributed to all Tier1s for analysis

Reconstruction jobs





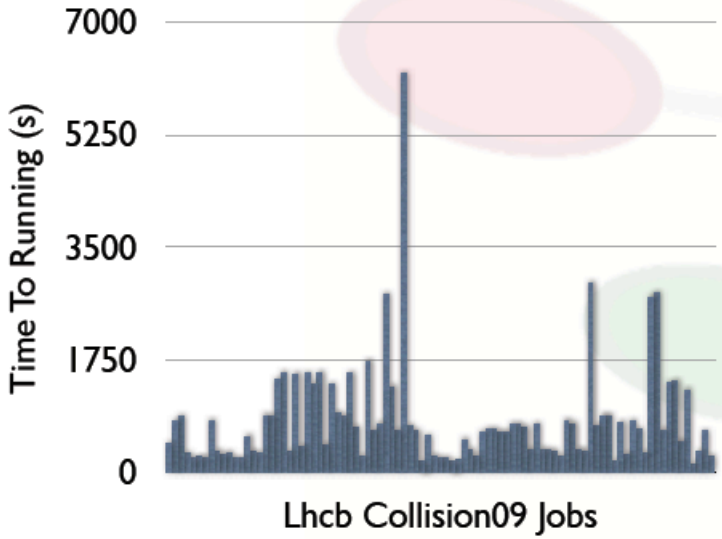


Issues with real data

- Castor migration
  - Very low rate: had to change the migration algorithm for more frequent migration
- Issue with large files (above 2 GB)
  - Real data files are not ROOT files but open by ROOT
  - There was an issue with a compatibility library for slc4-32 bit on slc5 nodes
    - Fixed within a day
- Wrong magnetic field sign
  - Due to different coordinate systems for LHCb and LHC ;-)
  - Fixed within hours
- Data access problem (by protocol, directly from server)
  - Still dCache issue at IN2P3 and NIKHEF
    - dCache experts working on it
  - Moved to copy mode paradigm for reconstruction
  - Still a problem for user jobs
    - Sites have been banned for analysis

Transfers and job latency

- No problem observed during file transfers
  - Files randomly distributed to Tier1
  - Will move to distribution by runs (few 100's files)
  - For 2009, runs were not longer than 4-5 files!
- Very good Grid latency
  - Time between submission and jobs starting running



Mean	13 mins
Minimum	133 secs
Maximum	1hr 43 mins

Brief digression on LHCb position w.r.t. MUPJs

- So-called *Multi-User Pilot Jobs (MUPJ)* are used by DIRAC on all sites that accept *role=Pilot*
  - They are just regular jobs!
- MUPJs match any Dirac job in the central queue
  - Production or User analysis (single queue)
  - Each PJ can execute sequentially up to 5 jobs
    - If remaining capabilities allow (e.g. CPU time left)
    - MUPJ has 5 tokens for matching jobs
  - *role=Pilot* proxy can only retrieve jobs, limited to 5
- A limited user proxy is used for DM operations of the payload
  - Cannot be used for job submission
- Proxies can be hidden when not needed
- DIRAC is instrumented for using gLexec (in any mode)
- First experience
  - Problems are not with gLexec but with SCAS configuration

MUPJs (cont'd)

- LHCb is not willing to loose efficiency due to the introduction of badly configured gLexec
  - Yet another point of failure!
  - Cannot afford testing individually all sites at once
- This topic has been lasting over 3 years now
  - Where is the emergency? Why did it take so long if so important?
- Propose to reconsider pragmatically the policy
  - They were defined when the frameworks had not been evaluated, and VOs had to swallow the bullet
  - Re-assessing the risks was not really done in the TF (yet)
  - Questionnaire leaves decisions to sites
    - We got no message that sites are unhappy with current situation
  - MUPJs are just jobs for which their owner is responsible
    - Move responsibility to MUPJ owner ("the VO")
    - Two tier trust relation (Sites / VO / User)
    - Apply a posteriori control and not a priori mistrust
  - VOs should assess the risk on their side

Conclusions

- Concentrating on real data
  - Very few data (200 GB)!
  - Very important learning exercise
  - A few improvements identified for the 2010 running
    - Run distribution (rather than files)
    - Conditions DB synchronization check
      - Make sure Online Conditions are up-to-date
- Still some MC productions
  - With feedback from first real data
    - E.g. final position of the VeLo (15 mm from beam)
- Analysis
  - First analysis of 2009 data made on the Grid
  - Foresee a stripping phase for V^0 physics publications
- LHCb definitely wants to continue using MUPJs!

2009 Q4 REPORT