ORB Report

Wednesday 2002 November 20

Central Analysis Backend (CAB)

- Small & Large tests ➔ Volume production of d0mino can be moved to CAB
  - Batch jobs should run ~120 times faster
- For people who know what they are doing
  - Support & troubleshooting is limited
  - No parallelization for ROOT yet
    - Not good for ROOT analyses
  - Optimized for people using d0tools
    - Good for stripping thumbnails, making root-tuples, tmb-trees, etc.

CAB & Farms

- Sign up for the d0cab-users@fnal.gov mailing list
- Prioritize work for Physics groups
  - Send requests and feedback
- How-to page available

http://www.nuhep.nwu.edu/~schellma/cab/cab_doc.html

- Data Farms - running smoothly all week
  - 10-12% crash rate with p13.02
  - About 10% problems - processing taking too long
  - p13.03 should take care of a lot of this
    - New version should go to farms end of this week

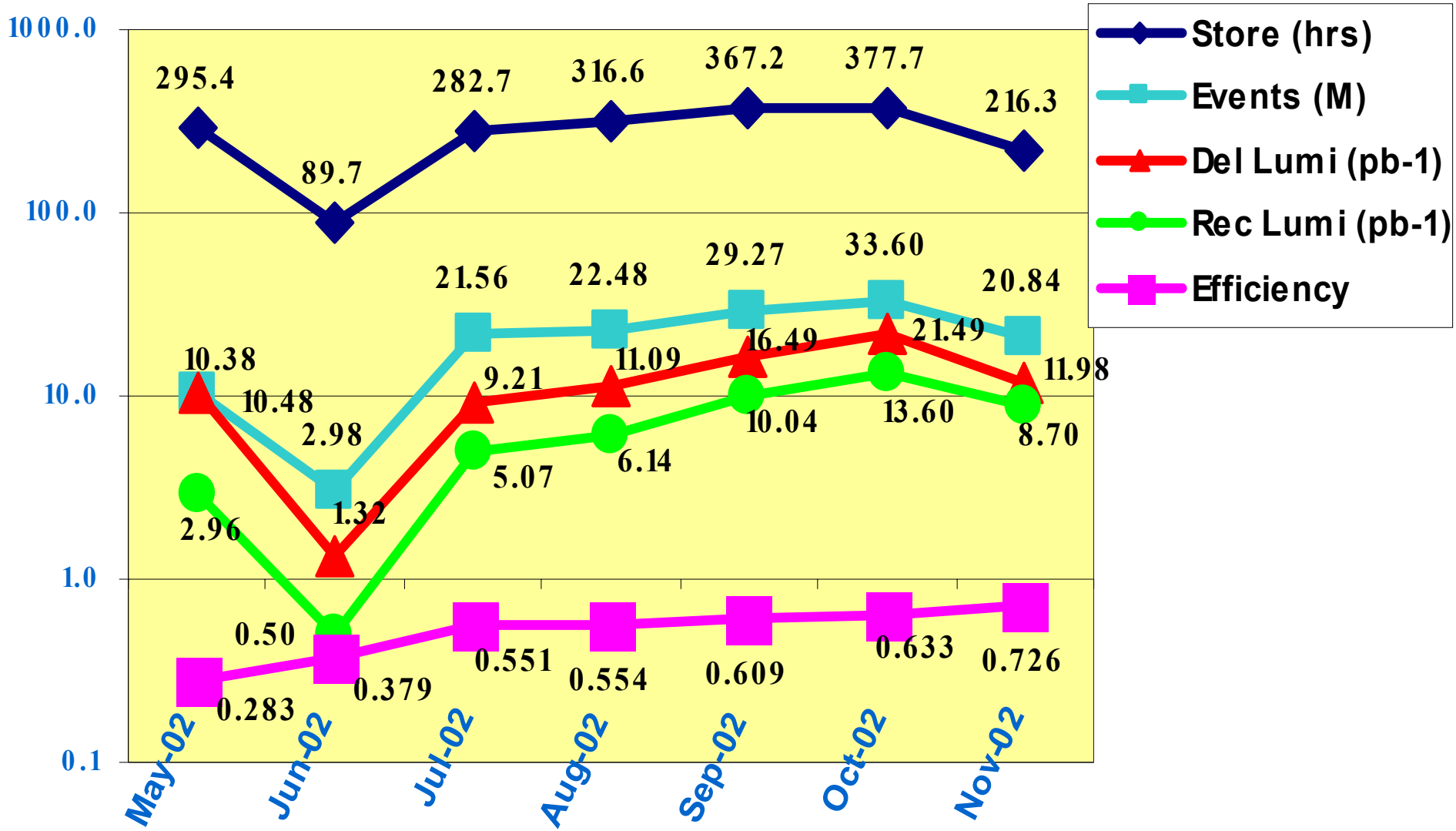
Farm nodes & ENSTORE

- Last batch of farm nodes being installed
  - 140 new nodes + 122 old ones
    - About six nodes need disk work
  - 12 racks of nodes now
    - Three racks of nodes today (Wed)
  - Turn off really old nodes (500 MHz) - about 20 of them
    - This is a power utilization issue
    - Should only have a 2-3% effect on processing
- Problem with ENSTORE ~17:00 on Tues Nov 19th
  - Probably a result of maintenance work
  - Writing files to multiple tapes
    - "No Access" tapes

Event Processing

- 135 million events (MEvts) on tape since the June shutdown (resumed data taking June 18th)
  - ~5% of events from daq_test, zero_bias & special runs
- 15 MEvts processed by p13.02
- Estimate 20 MEvts/wk to be processed with p13.03
- Since July 1st we have recorded an average of 6.5 MEvts/wk *(see next slide)*
  - Last week (Nov 11-17): New Best of 10.4 Mevts
  - Estimate 10-14 weeks to process all date since June with p13.03 for Moriond
    - If January shutdown begins on 13th (eight weeks away)… well you can do the rest of the math
- New data is processed the next day
  - Backlog is processed going backwards day by day

D0 Monthly Data Taking



Monte Carlo Status

- Problem with Meta data in p11
  - SAM complains if file name has non-integer number - cannot return match with 90.0
    - Needs changed in SAM database
- P13.03 MC certification files (reco output) went into SAM on Thursday
  - `sam translate constraints --type=mcrun --dim='global.requestid 2187 and data_tier XX'`
    - XX = reconstructed, simulated, digitized, thumbnail
  - Distribute to remote farms early next week
- Switching to 200 Gb tapes in ENSTORE
  - Need to merge files to ~1Gb size in order to access shorter list of files

More MC & SAM

- Discussion on weighting of requests
  - Which groups get priority, and in which order
  - This lead to heated discussion on how fast events are processed
    - Iain promised numbers (in evts/Hz or day) shortly
- SAM on Clued0 is near production
  - People have been using it for 9 days (Wed)
  - By end of week, it should be ready for everyone
  - Discussion that Lucas is leaving in a couple of weeks & Dugan in May
    - Some hand-wringing of course

Scriptrunner "bug"

- Logical flaw that occasionally caused farm nodes to incorrectly identify the proper triggers in the event headers (used by Luminosity, SAM, and Examines)
  - Hint of the problem for weeks but not conclusively found until a few days ago
    - Fix was made & has been online since Tuesday
  - Affects data taken from April 2002 ➔L2 was turned on
    - Small but non-trivial fraction, particularly in terms of normalization
    - Example: November data taking efficiency would drop from 72.6% to 61.7% (as of Wed Nov 20th) if "bad" data was discarded
      - That is a "loss" of 1.4 pb^{-1} of data in just three weeks!

How does this affect you?

- There is sufficient information in the raw data to fix this problem
  - Using L1 trigger mask from the TFW, L2 trigger mask from L2 Global, we can determine if the L3 trigger bit should have been set
    - The "bad" triggers could then be removed
- All DSTs & p13.04 Thumbnails can be repaired, but none of the root-tuples
- To recover "lost" luminosity, it will be necessary to generate new "event catalogs"
  - Important to perform unitarity cross-checks
    - Expect to take a couple of months to spin through all the data
  - This option will allow us to repair SAM as well

Unitarity Checks

- If even one event failed then the entire block is declared bad. The basic unitarity criteria that are associated with data provided by L3/ScriptRunner (recorded information comes from the event catalog, accepts come from the TFW; all information is integrated over an LBN):
  - if l3Recorded > l2Recorded: BAD
  - if l3Recorded > l1Recorded: BAD
  - if l3Recorded > l2Accept: BAD
  - if l3Recorded > l1Accept: BAD
  - if l2Recorded > l1Recorded: BAD
  - if l2Recorded > l1Accept: BAD
  - if l2Recorded > l2Accept: BAD
  - if l1Recorded > l1Accept: BAD

Summary

- P13.04 will have the fixes for the Scriptrunner bug
  - Requires processing all the data…again
- Try to identify runs which were unaffected
  - Bug bit us when other runs were started or stopped during a global physics run (even configured)
    - Usually commissioning or calibration runs
      - SDAQ, CFT, L1 CTT, L2 STT, etc.
- All trigger efficiency studies that used data and a L2 trigger will need to be re-examined
- All data prior to Feb 9th 2002 is already "junk" because of two earlier Scriptrunner bugs
  - "junk" in the sense you will not be able to normalize your data with