Phase 3 Spring and Beyond

HLT AND EXPRESSRECO

2019; August 27th

Nils Braun | IETP - KIT

HLT DURING OPERATION

Since beginning of Phase 3, many crucial bug fixes have been implemented thanks to the help of many people

- stable signal handling
- faster stopping time
- many DQM plot iterations, fixes, improvements
- configuration file management
- totally new steering file setup
- inclusion in the zabbix setup

Additional to that, we have validated and implemented multiple new basf2 releases and many database updates

HLT operation is much more stable in the recent times

STABLE SIGNAL HANDLING

- Stopping of run and aborting of framework is done by sending Linux signals to processes
  - Very important: processes need to go down cleanly
  - Check that processes are down
  - Close ring buffers correctly
  - Typical flow: main loop
    - instead of just aborting, break out of main loop
    - make sure to call destructors
    - e.g. `daq_slc`: Guards
- Important: blocking pipes are always evil
  - e.g. seen processes stuck in termination, because logging can not be flushed
  - also writing out to `nfs` can be a problem

FAST STOPPING, DQM PLOTS

Change in the DQM histogram file storage   R.Itoh

Underlying problem

- HLT response (at run stop) is very slow. Takes ~3min to stop.
  * The time is mostly to store DQM histograms in a file.
  * Independently of the live DQM histogram transport to
    dqmserver, when stopping basf2 at run end, the histograms
    are saved once in files corresponding to each basf2 process
    and then added up.
  * Currently 40 processes / one worker node is running.
  * Number of DQM histograms is more than 7500.
      -> Need to collect/add 3 million histograms! which takes time.
  * The histogram addition is done in multi-step
    1) store histograms in each process in a file (40 processes)
    2) add 40 files to a one file (each worker node)
    3) add files from 16 workers to one (on main server)
  * The files in 1) are now written to /tmp instead of to a common
    NFS directory to avoid the network conjestion.
  * 2) and 3) are now pipelined.
  * But still taking time.
  * Further tuning is being tried.

FAST STOPPING, DQM PLOTS

- Number of histograms has been reduced down to 2800 (thanks to the effort by TRG group), and the time for HLT stop was reduced to 2min. or so. But still long.

- Final solution:
  * To give up the histogram collection at each run end.
  * Instead, store the live histograms already transferred to the main node at run end.

  * For the purpose, a new NSM node called "DQMMASTER" is installed.
  * It receives the STOP signal from run control master and dump histograms on TMemFile into files.

  * It was implemented yesterday and now being operated.

     -> The stopping time of HLT was reduced to 30sec but additional 30 sec is required to store histogram files in DQMMASTER
        -> Total of <1 min. to stop.

FAST STOPPING, DQM PLOTS

- The histogram files are copied to
  qasrv01:/data1/dqm/dqmsrv1/e0008/dqmhisto/
           hltdqm_e0008rxxxxxx.root, and
           erecodqm_e0008rxxxxxx.root.

- The directory structure of the file is different from previous one.
  * No subdirectories corresponding to subsystem.
  * Instead, the subsystem labels are attached to histogram titles.

- The histogram files are copied to kekcc by Boqun
           kekcc:/group/belle2/phase3/dqm/dqmsrv1/e0008/dqmhisto/.

- One Note:

  * In this implementation, we cannot guarantee that "all the
    events" are in the stored DQM files. Only up to the events
    where the histograms are transferred to the main node.
    (Events in a last few minutes before STOP are not
     accumulated in histograms)
  * Should be OK, since it is a kind of "snapshot" of online DQM.
  * For the detailed study, you can check full events in offline.

CONFIGURATION/STEERING FILES

- The HLT configuration files are store on git now: `https://stash.desy.de/projects/B2DAQ/repos/hlt_config/browse`
- The `basf2` steering files are reworked to refer to the same common base class
- All steering files are tested via integration tests on every PR, has already prevented stupid bugs

DOCUMENTATION

Help wanted

Documenting such a complex system is not simple.
Your help for filling out troubleshooting pages, providing input and **giving feedback** is needed (both from experts as well as from newcomers).

- I still did not receive neither feedback on the graphs I created nor feedback from experts on the correctness of the documentation
  - When you read and us it, please always remember I have written the documentation by myself – it might be wrong!
- **Don't just restart**
- Use gdb
- `https://confluence.desy.de/display/BI/HLT+operation+in+Phase+3`
- `https://confluence.desy.de/display/BI/ERECO+operation+in+Phase+3`

PRESENT

- Filtering turned on
    - Did first studies with HLT filtering turned on
    - Large trigger performance studies on data before that
    - No show-stoppers seen
    - Expected data reduction approx. factor 8-9
    - Open discussion item: What to do with dismissed events?
- CVMFS
    - Installed already on all HLT machines, quick analysis server, ExpressReco and monitored by zabbix
    - Used for `daq_slc`, `basf2` and conditions db
        - Also want EB software :-)
        - Still not using same `daq_slc` as the rest. Some discussion items .
        - see Oskar's talk

FUTURE

- Improved HLT trigger menu and reconstruction speed
  - Work in progress, JIRA issues and discussions
  - Will be a continuous effort
- ØMQ (see other talk)
- Updated `cvmfs` (see Oskar's talk)
- More HLT units

Personal Note: I will not be the online coordinator for the upcoming data taking. Please ask questions to Chunhua or (starting from October) to Markus Prim.