FAX Testing

Frederick Luehring

luehring@indiana.edu

Indiana University

US ATLAS Software & Computing/Physics

Wednesday August 20, 2014

Methodology

- The methodology is to make a simple cut flow skim of the 2012 Standard Model W/Z D3PD dataset.
    - Less than 1% of the events are written out.
    - The SMWZ D3PD example code developed by Doug Benjamin is used.
        - A standalone executable calls ROOT I/O and TTree libraries.
    - The dataset is quite large (220 TB uncompressed/90 TB as stored in D3PD files) but through use of the TTree mechanism, only about 1.4 TB (0.55%) are read in the job.
        - Even with this selective reading the jobs are IO and not CPU bound

- There are dataset contains 18856 files and all files were available in the US in April 2014.
    - The test reads some (but NOT all) data from all files.
    - All files were present at BNL and many of the files were present at the other US facility sites too.

Methodology

- When the test jobset fully processes all events, it reads 732,026,805 events in 18856 files using 1886 jobs (10 input files per job).
  - The jobs do a total of just over 1M ROOT read operations.

- I have been running the test jobset most days since mid-May and have done ~75 tests by flocking the jobs from IU Tier 3 site to MWT2 (ATLAS connect).

- The test jobset has also been tested successfully by:
  - Running the jobs locally on the IU Tier 3 reading the data from the WAN over a 10 Gbps Ethernet connection
  - Submitting the jobset using prun
  - Logging into the ATLAS connect login server and submitting the jobs there. (Still debugging this approach.)

Test Results

- To analyze the jobs, I scrape the log files and count the number of full size output ROOT files (when fully successful, the job produces 1886 ROOT files that are ~7 MB each containing the skimmed data).
  - There are certain errors that result in a file being skipped.
    - The most likely XRD error code causing this is error 3012: "Failed to open file (java.lang.IllegalStateException: Replica already exists [10011])".
    - Several other XRD codes cause files to be skipped.
    - The most frequent XRD code is error 3011: "Server [atlas-xrd-central.usatlas.org:1094] declared: session not found" seems harmless and the data is normally read on a retry.
  - Some jobs fail with a segmentation fault caused by an XRootD issue and Ilija Vukotic filed a bug report with the ROOT team.

Current Status

- Recently the number of segmentation fault failures has been small but most test jobsets do still skip a few files (less than 10 typically).
    - As of the most recent XRootD version (4.0.4) this segmentation fault still happens occasionally.
    - The fix for this issue is expected soon.
- So far there have been about 5 fully successful jobs that read all input events.
    - Most jobs read nearly all the events (more than 99%).
    - I need to follow up on why the FAX retry mechanism does not successfully work to find these datasets at other sites.

Providing a FAX Health Test

- I have cleaned up the my scripts, condor command files, and text file containing the dataset URIs.

- I provided them to Alden Stradling to be used to create a test of the FAX system health.

  – The test job now has job options that allow one to set:

    - The number of events and type of job (flocking, local, connect) and working directory path

    - The file containing the input URIs

    - The ROOT and XRootD versions

    - The full path of the X.509 proxy and the RUCIO user name

- As my time permits, I will put these scripts into the ATLAS Connect GitHub area and provide documentation on the ATLAS Connect Webpages to run the example.