Doing TRT Analysis at IU Tier 3 and ANALY_MWT2

Frederick Luehring
Indiana University
luehring@indiana.edu
3-Mar-2009
ATLAS Tier 3 Meeting at OSG All-Hands Meeting

Introduction

- Using both RAW (ByteStream) and ESD files for single beam and cosmic runs, analyses were done using the "classical" data analysis model:
  1. Get sample data files on the Tier 3 using dq2-get.
  2. Repeatedly edit-compile-test modifications of the TRT_Monitoring package.
  3. Subscribe the datasets to the MWT2_DATADISK.
  4. Use pathena to run against the full dataset.
     - Both ESD and RAW datasets were analyzed
- The Tier 3 in this talk is **not** a Tier 2-like site.

Tier 3 Setup

- The IU Tier 3 is a departmental cluster shared between dzero and ATLAS (the da cluster).
  - Mixture of ~10 servers with ~20 cores
    - No two servers have the same hardware
  - SLC4 / Rocks with PBS queue
  - 2 RAID arrays totaling about 12 TB
  - 1 Gbps connectivity
  - All recent versions of Athena
  - wLCG and DQ2 client software
  - **NO** grid server middleware
  - Managed by a part of the departmental sysadmin.
  - 50 miles from the closest part of MWT2

MWT2 Setup

- ANALY_MWT2 located on MWT2_UC
  - By default 200 cores are reserved to handle analysis jobs with high priority though production jobs run when no analysis jobs are available.
  - Currently ~20 TB available to analysis jobs in MWT2_DATADISK (which is physically on MWT2_UC).
  - Storage served by dCache.
    - The dCache causes difficulty when the cluster is drained and many analysis jobs start at once.
      - This is being investigated

Conditions Database Issues

- The Conditions Database caused two issues:
  1. Very slow transfers of data from BNL Oracle during job initialization (affects T3 and MWT2).
     - Should be solvable using FroNTier and Squid.
     - Carlos Gamboa helped us tune the IU Tier 3 network and the Oracle setup **BUT** a job using ESD or RAW data still takes something like 15 minutes to read Oracle (~30 seconds for the same job running at BNL).
  2. COOL requires certain .pool.root files be present locally on the machine where the jobs run and an XML file specifying their locations.
     - This is separate and is not addressed by FroNTier/Squid.

What is Possible

- It is possible to work efficiently:
  - Subscribe your dataset to the Tier 2
    - If the subscription is not acted on, email Kaushik.
  - Use dq2-get to get the data onto the Tier 3
    - dq2-get from any US site is fast.
    - dq2-get from CERN is slow.
  - Develop/test your software package on the Tier 3
  - Run your Monte Carlo locally on the Tier 3
  - Then send hundreds to thousands of jobs to the Tier 2 using pathena
    - Have to be careful not to waste cycles.