

PanDA Server

Tadashi Maeno (BNL)

Beyond-Pledge Resources Management (1/2)

➤ New additional policy

- Runs jobs for users who belong to a particular country group if those jobs are waiting in the queue. Otherwise, runs any user's jobs
 - Resources are always being used effectively
 - Ratio between Pledge/Beyond-Pledge can be defined in shedconfig

Beyond-Pledge Resources

Management (1/2)

➤ Implementation

1. availableCPU, pledgedCPU, availableStorage, pledgedStorage in schedconfig
2. The pilot sends a request with countryGroup=XYZ and allowOtherCountry=True
 - Single pilot stream
3. The panda server does
 - calculate the ratio of running jobs with countryGroup=XYZ to all running jobs at the site
 - Get the number of activates jobs with countryGroup=XYZ

This could be moved to memcached from Oracle if heavy

4. A job with countryGroup=XYZ is sent to the pilot when the ratio is less than $(1 - \text{pledgedCPU}/\text{availableCPU})$ and there are some activated jobs with countryGroup=XYZ. Otherwise, any job is sent

Database Archiving

- Jobs in ATLAS_PANDA are copied to ATLAS_PANDAARCH as soon as they are finished/failed/cancelled, and then are deleted from ATLAS_PANDA 3 days later
 - To keep the number of jobs in the live tables small
- Multiple processes were doing the above in parallel
 - caused contention in ATLAS_PANDAARCH due to bitmap indexes
- Changed to run a single process
 - Cannot catch up with the rate the number of jobs in ATLAS_PANDA is increasing
- Gancho has created PL/SQL procedures
 - Copy
 - bulk INSERT works fine
 - Delete
 - bulk DELETE is not very efficient due to indexes in filesTable4
 - Creates daily partitioning and drops old partitions

DDM-related Issues

- Now dis datasets are created even if input files are already available at T2
 - dis datasets were created only when input files needed to be copied to T2
 - Files were sometimes deleted during jobs were waiting in the queue since DQ2 didn't know that those files were going to be used
- Coming changes
 - dis and sub datasets will set lifetime and the 'hidden' attribute of metadata
 - sub datasets will be deleted from EGEE T1s as soon as subscriptions are completed
 - Already deleted from T2s
 - File-level callback via ActiveMQ
 - If DQ2 SSS send callbacks directly to panda server, the panda server would be killed
 - To freeze output datasets for analysis jobs when all jobs finished/failed
 - Required for DaTRI when --destSE is used