



Recent Developments in Panda

Torre Wenaus
US ATLAS Tier2/Tier3 Workshop
Fermilab
March 8, 2010

Status of the November Priority List



- Migration of Panda server to dedicated machines - done
- PandaMover monitoring - done
- Event/alarm logging in monitor - done
- Monitor-independent queue data info to pilot - done
- New schedconfig management system - pending (soon)
- autopyfactory: pyfactory + autopilot's DB-based config/monitoring - pending (not soon)
- multi-job pilots - next release (1-2 weeks)
- glxec in production - pending (by end March)
- pcache based WN-level file brokerage - under test
- DQ2 based site mover deployment (Europe) - pending
- Error stats and raw data dump mode added to stats monitor



Other Activities

- Panda pilots in CERNVM
 - CERNVM team (Predrag Buncic et al) implemented current Panda pilot in CERNVM-based virtual WN
 - With a new pilot management system for CERNVM they've developed
 - Targeted at BOINC but generally applicable to VMs
 - Free effort (with some consultation/support from us)
 - Follow-through to a full VM/cloud-based Panda site up to us
- Panda monitoring consolidation/coherence with ATLAS
 - No progress (worse; negative progress)
 - CERN ATLAS monitoring team leader left (R. Rocha)
 - Monitoring effort being reconstituted under Alex Read
 - New technical expert coming up to speed; long ramp
 - We will participate/track the activity and look (again) for ways to offload parts of monitoring to ATLAS



New Priorities Since Nov

- Squid proxy in front of Panda monitor - in progress
- Enhanced WN space management
 - Pilot limits total input file size based on site config param; default 14GB
 - Pilot capable of monitoring output file sizes (not yet used)
- Panda support for T3g - under test
- The usual steady stream of ops-driven optimizations and changes, lately driven by real data (re)processing
 - Brokerage optimizations for rapid pickup of high priority tasks
 - Brokerage optimizations for improved task distribution among clouds



Worker Node Level Brokerage

- New scheme to extend Panda data-driven brokerage down to the WN level
 - Uses Charles' pcache system to manage WN resident files
 - Plan to then extend it to xrootd
- Panda adds all job input files to a WN-level file catalog
 - memcached catalog collocated with the Panda server
- pcache/cleanup informs Panda when files are deleted, so Panda can maintain its list
- Also an API to populate Panda's catalog directly
 - eg. for preplacement or external management of WN data
- Currently under test

Panda @ T3g



- Panda now supported on Tier3s with ~no grid support
 - No DQ2, no gatekeeper/CondorG
- Pilots submitted locally to the site, direct to condor
 - Only condor local batch supported; others could be added
- Pilot authentication to server based on regexp match to WN name, not https/proxy (proxy not available)
- User input file specification via file list, not dataset
- Output 'dataset' maps to directory
- No file registrations or catalog dependence in pilot
- Full Panda monitoring available
 - Including, soon, log tarball access from monitor
- Can use Panda's access controls to limit usage to defined list of locals+friends (site managed)

Panda @ T3g Usage



- conventional pathena:
 - `pathena jobOpt.py --inDS dsname --outDS outname`
- pathena to T3g:
 - `pathena jobOpt.py --site ANLASC --pfnList list.dat --outDS outname`
- where list.dat is a list of input files
 - either conventional filenames
 - or supported remote I/O specs, eg. `xrootd://...`, `dcache:/pnfs/...`
- pathena must be issued from environment with grid proxy, as usual
- output files found at
 - `/SiteSpecificRoot/2010/UserName/DatasetName/FileN`
- Output files owned by the pilot submission account

Panda @ T3g Status



- Queues set up at Duke, ANL; pilots flowing
 - Testing with 15.6.3 jobs at ANALY_ANLASC
 - Waiting for SLC5 at ANALY_DUKE
- Data source and sink functionality in DQ2-free mode validated
 - Outputs to a directory structure: Year/Username/Dataset/File
- pathena build jobs work; athena run jobs fail in athena, still diagnosing
- PandaTier3 wiki taking shape:
 - <https://twiki.cern.ch/twiki/bin/view/Atlas/PandaTier3>
- Ready to add more T3s; get in touch!

Panda Based Analysis



- Much work to analyze and optimize usage patterns
 - In particular to drive jobs to sites other than BNL
- Removed cloud constraint in analysis brokerage to use grid more efficiently
- Explicit warning to pathena/prun user when jobs are hardwired to a busy site with --site option
- Implementation in progress for automatic output delivery to user-designated site
 - Extension to DaTRI: pathena uses DaTRI API
 - User monitoring of subscriptions
- Support added for analysis project caches in pathena



Near/Mid Term Priorities

- Finish/deploy Panda@T3g, respond to feedback
- Finish/deploy WN level brokerage at dcache site(s)
- Extend WN level brokerage
 - xrootd
 - T3 sites with preplacement of WN data
- Finish multi-job pilot and deploy on analysis queues
- Finish/deploy glexec in production
 - Gradual rollout as sites ask for it. A few Euro sites ready now; BNL may be soon
- Finish/deploy new job recovery system
- Monitoring adaptation/integration with AGIS, new ATLAS monitoring effort



Summary

- Panda continues to be stably functional, no scaling or performance issues
 - Long planned improvement in service configuration -- separation of Panda server & monitors -- now done
- Development team kept busy nonetheless!
 - Steady stream of improvements and optimizations driven by real data (re)processing and analysis
 - Major refactoring of the pilot ongoing (and near completion) for code cleanup and support for glexec, multi-job pilots
 - Major extension of scope to cover T3g -- looking forward to working with T3s to make Panda a valuable T3 service
- More of the same: optimization/extension for prod/analysis layered over support for production running