

ProdSys Development

- Rod Walker, LMU Munich,
6th April 2011

- Introduction
- Development
 - Short term
 - Long term
- Manpower
- Conclusions

What is ProdSys?

- ProdDb, AKTR and job definition scripts
 - Task stats update, increase max attempt, abort jobs, finish tasks
 - Plus Panda monitor web pages for tag & task definition
- Everything between TRFs and Bamboo
- A lot of stuff

Development

- Short term
 - Day to day maintenance
 - New trf, new params, new workaround – often needs script changes
 - SL5 migration
- Long term
 - Consolidation to ease/share maintenance
 - r and f tag unification
 - New use cases needing more work

Short term

- AKTR maintenance by Alexei
- As for the rest - still 'getting by' in Pavel's absence.
- Necessary changes
 - SLC5 host migration
 - fairly painless, with a few surprises
 - tweaks from Borut, for new trf parameters and fixes
- Done some tidying and documenting but more ToDo
 - <https://twiki.cern.ch/twiki/bin/viewauth/Atlas/ProdSys>
- Pending development work
 - process 10% ESD output(assumes 100%, no DDM)
 - Trigger reco(formerly done on CAF)
 - Prompt T1 ESD processing (formerly T0)

Open Problems

- Not all output datasets deleted when task obsoleted
 - task request wants subset of allowed formats in ctag
 - _ only requested formats get dataset created by AKTR
 - _ jobdefs produce other formats – panda creates missing datasets
 - extra datasets unknown to AKTR/AMI – not deleted
- Slow to finish old tasks – follow strict priority
 - daily increase in job priority is broken since slc5 move
 - replace with high priority when 90% done
- Getting allowed parameters to trf sometimes fails
 - Higgs group repro stuck – fixed now
- Super-merging of AOD to respect Lumi Blocks
- Multiple jobs per input based on AMI #events
 - done for HI repro but not yet integrated
- Support Generate_trf for mc11
 - involves tag request page

Longer Term

- De-cronification
 - Ordered steps to process task done by independent crons
 - e.g output task created by one cron. Jobdef cron checks for existence before processing task.
 - could trigger dataset creation while processing request
 - leads to natural delays in processing requests
 - Are other things such as containers, freezing for finished tasks
 - would like to deal with all of these at once

Longer Term

- Db consolidation
 - ProdDb (ejobdef,ejobexe) and Panda job table have much duplicated information
 - can lead to inconsistencies
 - Still on wish list, but no progress
- r tag to f tag
 - f tags used by T0, rtag by prodsys – want only 1
 - f tags created in AMI. Trf args in pickle file.
 - makes sense to use these, but needs some dev.

Manpower

- Short term support
 - Rod and code snippets from Borut. Sasha dev to submit repro
- Short and longer term
 - Dmitry arrived at CERN for 3 months
 - need to discuss, but likely assist in maintenance while planning improvements
- Operation and development
 - almost complete overlap
 - good to break this

Conclusions

- Coping well, without major development
 - learned a lot from the migration
 - mostly patient and helpful customers
 - simple jobdef fixes made, but larger development work halted
- Can improve
 - response to new requests (de-cron)
 - maintainability, operability, monitoring
 - help is on the way