

ADC SW Dev Planning, Priorities, Timeline

S. Campana, T. Wenaus
S&C Week
April 6, 2011

ADC development meetings

Proposal: regular weekly meeting Wed ~17:00 CERN time

- Many projects and activities underway; use a regular weekly meeting for good communication, inter-project awareness and planning, meeting consolidation, engagement of non-ADC participants
- Duration: objective 1hr, limit to 1.5hr

Meeting structure:

- Focus topic
 - More in-depth discussion of a particular activity, with non-ADC invitees as appropriate
 - At the top of the meeting to be friendly to non-ADCers
- Hot topics, deliverable tracking
- Open floor across all activities for brief recent highlights

Prospective Near-Term Focus Topics

- Pre-Dubna objectives for ADC dev related Dubna topics
- Cloud computing evaluation - near term plans
 - Scope for near term: target platforms, functionalities, priority development (DDM, PanDA, ...)
 - Workshop plans & objectives
- File & event level caching - near term activities
 - Relationship and commonalities between the two
 - Event level caching approach & protocol
 - Testing plan, performance for representative use cases
 - Brokering for reuse - strategy and tools
- Rucio - architecture, scope, components, technologies
 - Rucio for DDM - overview and plans
 - Rucio for others - how can other areas leverage it
- Relational and NoSQL DBs

Prospective Near-Term Focus Topics

- PanDA evolution
 - Pilot - can/should we leverage GlideinWMS? If so how?
 - Pilot - efficient resource utilization in athenaMP/all core scheduling
 - Messaging & notification, ActiveMQ, leveraging/coherence with DDM
 - Panda-internal file/dataset tables needed in Rucio era??
- PanDA/site metadata rationalization
 - Future of schedconfig, AGIS integration
- Prodsys evolution
 - Prodsys/PandaDB consolidation/convergence?
- Coherence and collaboration in monitoring
 - Broader use of Dashboard infrastructure - assessment of experience

Priority dev deliverables

Incomplete strawman for discussion & adjustment

April

- T1-T1 PD2P implemented & under test
- new looping job killer in pilot
- post-Graeme support for autopyfactory defined

May

- JEM/PanDA deployment to users/shifters
- LFC registration in Panda server
- analysis output file merging in PanDA
- pandamon classic => Dashboard infrastructure prototype & evaluation
- Dashboard based task monitor

June

- autopyfactory support for local batch pilot submission at T3s
- pilot production testing moved to Hammercloud

Priority dev deliverables (2)

Summer

- ToA migration to AGIS?
- Cassandra job archive back end on PanDA job monitor
- S3 (API) as DDM site?
- PanDA cloud site accepting jobs and supporting data in/out
- PanDA Tier3 WN data affinity deployment at 2+? interested sites
- Active analysis job control via web monitoring interface (kill, retry, peek)?

Fall

- task support reworked to support analysis as well as production?
- prodsys dynamic job definition?
- AKTR de-cronified?

Later

- Oracle 11g migration: Dec '11 timescale, PanDA/DDM ready before then
- Workload sharing in all-core queues? (for efficient usage when athenaMP doesn't fill them)
- Role for glideinWMS in our pilot system?

ADC dev related topics in Dubna

PanDA/DDM/PD2P

Event level caching, xrootd federation/file level caching

Data storage evolution

Cloud computing

Relational and non-relational DBs