



# POOL Status and Plans

---

Dirk Düllmann,  
IT-DB & LCG-POOL  
Application Area Meeting  
10<sup>th</sup> March 2004



# POOL Release V1.6



- Has been produced yesterday..
  - <http://pool.cern.ch/relnotes.html>
  - Using SEAL\_1\_3\_3, Root 3.10.02, edg-rls-client 2.2.1
  - Now build binaries for Linux and Windows
    - rh73\_gcc32, rh73\_gcc32\_dbg, rh73\_gcc32\_gcov, win32\_vc71, win32\_vc71\_dbg
- Main new features in V1.6
  - Composite Catalogs
  - New documentation structure
- SPI distribution kits and updated installation instructions will appear on our web shortly



# Next Steps



- Will move to new IT-CVS service
  - Scheduled for tomorrow 10:00
  - Will take the opportunity to clean and rearrange the POOL repository



# 2004 Work Plan



- Draft document has been discussed in POOL and the Architect Forum
  - [http://pool.cern.ch/POOL\\_Program\\_of\\_Work\\_20040307.doc](http://pool.cern.ch/POOL_Program_of_Work_20040307.doc)
  - Based on WP workplans and experiment priorities
  - No big objections received so far
- Will finalize the plan in the next week and present it to PEB and CS2
  - Thanks to all experiments for their concrete and detailed input
  - Significant overlap between the different experiment requests



# Main Focus in 2004



- Stabilise POOL s/w products
  - Focus on performance improvements rather than large functionality changes
  - In line with the experiments plans for the data challenges
- Help to simplify the integration into experiment frameworks
  - Tight coupling between POOL and experiment development and production teams
  - Automated schema loading, usability tools, documentation improvements
- Establish the new ConditionsDB
  - After a initial consolidation round
- Achieve POOL independence of the RDBMS backend
  - And extend the set of supported RDBMS systems



# POOL usability tools



- Received requests for tools which would simplify the daily development work of pool users
  - Create or recreate a POOL catalog from a set of interrelated POOL files (Q1)
  - Provide command line tools for consistent file manipulation of POOL files (Q2)
    - Eg copy, move and rename in the local file system together with the associated POOL catalog changes
- Provide eg in the POOL contrib area a central repository of scripts developed by an experiment
  - To share the experience in the deployment of POOL (Q1)
  - May later support and distribute some of these tools if there is sufficient agreement



# Storage Manager



- Optimisation required in several areas
  - Client side resource usage - memory, CPU, file handles (Q2)
  - Mass storage handling (minimise costly requests) (Q3)
  - "Transparent" double to float mapping (Q3)
- Automated schema loading (Q2)
  - Based on SEAL service
  - In cooperation with ROOT team to allow late integration of data types for already open files
- Bug fixes - more complex cases
  - Eg std containers with user defined allocators which define local data - aka CLHEP Matrix (Q1)
- RDBMS backend based on the RDBMS Abstraction Layer
  - Storage of simple data structures into a RDBMS via the same interface as for objects stored on the streaming layer
  - Two step plan:
    - First allow for objects which can trivially be mapped to SQL tables (Q2/Q3)
    - Possibly later an extension to more complex C++ objects (Q1 '05?)



# File Catalog



- Significant development completed already (Q1)!
  - Support for LCG-2 (V1.5)
  - Support for Composite Catalogs (V1.6)
- File Catalog as model for handling and exchanging data could be a prototype for other (very similar) meta data catalogs (Q2/Q3)
  - Collection catalog and Collection entries
  - Condition Folder catalog and Condition Data
- Cataloguing, extraction based on meta data, publishing are all very similar
  - Even the component implementation could be factorised out and shared
  - Performance of XML as exchange format for larger data amounts needs to be evaluated
- Would like to start an activity to propose a common approach at least for the persistency framework projects
  - Closely coupled to a possible emerging LCG activity of deployment of heterogeneous databases





# Collections & ARDA



- **Joined Work Package with ARDA**
  - Still some uncertainties concerning the ARDA side of the work plan
  - Will continue work to address the outstanding issues on the POOL side
  - POOL has asked experiments for principal contacts in this area
- **Collection cataloguing, extraction and publishing tools**
  - Can we achieve a single baseline model for distributed meta data catalogs?
    - File Catalog, Collection Catalog, Conditions Folder Catalog
    - One basic mechanism of data exchange across RDBMS vendor boundaries based on the POOL relation abstraction layer?
- **Separation of logical and physical collection identification**
  - Introduce a Collection (Fileset?) catalog
  - First implementation could simply be based on the existing File Catalog components, but as a separate service
- **Integrate POOL collections with ARDA provided services**
  - Migrate to ARDA provided catalog and meta data(?) services



# POOL Infrastructure



- SEAL Component Model (Q2/Q3?)
  - Once picked up by the experiments
- Parallel build and test machinery (Q2)
  - Including effective build on windows
- Automated data format regression testing (Q1)
  - Incorporation of experiment defined test suites into the POOL release test procedure (Q2)
- Evaluate Appworks and possibly migrate to it (Q1)
  - Depending on the future of SCRAM support
- Complete move to QmTest (Q1)
- Next Ports
  - ICC 8, ECC, MacOS(?)
  - Schedule to be review in one of the next AF



# More Internal Review Responses



- Restructure the POOL documentation (Q1)
  - All documentation formats derived from a shared set of DocBook text modules
  - Minimises the overlap and possible inconsistencies between design and user documents.
- Provide ROOT plugins for Ref<T> and POOL Collections
  - Allows to use POOL functionality inside ROOT as an interactive analysis environment (Q2)
- Schema Evolution (Q2/Q3)
  - Will start from the ROOT support and try to confirm that POOL does not restrict the ROOT functionality significantly
  - Will need to address also schema evolution on the RDBMS layer



# Conditions Database



Still preliminary and to be discussed with all Project Participants

- First IOV interface and implementation (Q2)
- Connection to POOL Data (Q2)
- Review of extensions to the basic interface (Q3)
- Condition Folder Catalog review (Q3)



# Summary



- POOL V1.6 has been released
  - Thanks to experiments for the quick turnaround in validating our internal releases
- POOL work plan is being finalised
  - Effort estimation of tasks still underway
  - Minor shifts in the proposed delivery dates may still occur until PEB/SC2 presentation
- Proposed Focus of the Year
  - Consolidation and Optimisation
  - RDBMS vendor independence
  - Common POOL model for distributed, heterogeneous meta data catalogs (incl. data exchange across vendor boundaries)
  - ConditionsDB production release and integration with POOL