



# Database Summary

Dario Barberis

Genoa University/INFN



# Database Software (1)

- Update on LCG Persistency packages (A. Valassi)
  - Current release LCG60b has many bug fixes and improvements, including:
    - Patch for CORAL reconnection problem in case of network glitches
    - ROOT 5.28.00b
    - Robust Frontier client 2.8.0 with better performance
  - API extensions to CORAL and COOL in the pipeline
    - e.g. many payload per IOV, replace ATLAS CoraCool (task #10335)
      - This breaks binary compatibility to previous releases
      - When will be a good time for ATLAS (offline/online) to release this?
    - Reviewing other pending requests for enhancements to plan ahead
      - What are the highest priorities from ATLAS (e.g. in COOL)?
    - CORAL performance studies and optimizations
      - e.g. comparison of Frontier and CORAL server
      - e.g. studies of solid state disks with IT-DB
  - Is there any other R&D that could be useful for ATLAS?



# Database Software (2)

- News on the Conditions DB and related tools (M. Borodin + D. Banfi)
  - Many improvements to python tools `AtlCool*.py`
  - Default connection to databases at CERN through Frontier server
    - except for writing of course!
  - New schemas for FWD detectors in COOL
  - Started work for conditions files access through CVMFS
- Conditions tags (M. Plamondon + P. Laycock + L. Sharmazanashvili)
  - Improving the model after experience with 2010/2011 repro campaigns
    - Decrease the number of global tags
    - Needs discussion with detector conditions people
  - New COOL Tag Browser (9.0) with several new functionalities
    - Improved search function, menus and GUI
- Quality Defects Database (P. Onyisi + P. Waller)
  - Contains list of defects by IoV instead of just quality flag
    - Auxiliary tools developed/converted to handle new system
    - Using well-defined, fast Python API for access
    - Underlying DB manipulation tools available for general use



# DCS, PVSS, Online tests

- DCS Data Viewer (S. Winkelmann)
  - DDV - a flexible, user friendly tool fetching ATLAS DCS data
  - 'Search Engine' and 'Relational Queries' useful functionalities
  - Can host on top of it other applications
  - Long list of additional features to be implemented
- Sliding window on PVSS data in ATONR and ADCR application issues (G. Dimitrov)
  - 12 months sliding window set to the PVSS data on the ATONR database
    - The ATONR database has a controlled max volume after introducing the 12 months sliding window for the PVSS data
  - New segment organization of the PanDA data
    - The ATLAS\_PANDA schema has increased scalability after setting up a set of Oracle jobs for adding, copying and dropping on partition level.
  - New procedures for better column statistics to reflect the reality
    - Important work on adjusting the Oracle column statistics for the "data or timestamp" type columns to reflect the reality.
- Online stand-by database and DB disconnection tests (F. Viegas + M. Blaszczyk)
  - Tests done mid-February
    - Several minor issues found in the transition to the stand-by database and solved
  - TGC database did not reconnect after switch — work needed
    - New tests to be scheduled after fixes implemented
  - Switch to stand-by database needs 30 minutes + time to take this decision



# Metadata

- AMI (S. Albrand)
  - Report on operation issues
    - Server problems, ORACLE problems, Security & Information Protection.
  - Developments
    - General, Real Data, MC, Data Periods through COMA
  - Plans
    - Improved synchronization with DDM
- Data Periods in COMA (E. Gallas)
  - Available interactively through AMI and programmatically through pyAMI
    - For RunQuery, BeamSpot, Data Quality etc.



# Geometry DB and DB releases

- Geometry DB (V. Tsulaia)
  - Optimizing number of queries (schema changes)
    - The effect of this optimization is clearly visible for the jobs, which read geometry data from Oracle (RTT, interactive jobs with nightlies).
  - Geometry DB in Frontier
    - First had to test that it does not add big overhead to the existing Frontier load caused by the conditions data access
      - Such test was carried out last week (March31) and it proved that the additional impact on Frontier caused by the geometry data is small
    - For the moment the Geometry DB is available via Frontier only at CERN
      - The next step is to put Geometry DB into Oracle streams replication from T0 to T1s. As a result it will become available via Frontier worldwide
    - It's not unrealistic to have this done by release 17
- DB Releases (V. Tsulaia)
  - Database releases (both MC and Conditions) play a key role in the production system (reprocessing, MC production)
  - Having the geometry DB available in Frontier offers several possibilities to rethink the overall strategy of DB Release usage in ATLAS
    - We can drop the current hybrid database access model in analysis jobs
    - The analysis jobs should be able to read both Geometry and Conditions from Frontier
    - No more need to download & unpack 1.2GB DB Release only for getting access to 50MB Geometry DB replica



# Frontier and AGIS

- Frontier news (J. DeStefano)
  - Frontier servers deployed at CERN and 6 Tier-1 sites
    - >100k analysis jobs/day access conditions data through Frontier
  - Concerns
    - Site Squids CERN server maintenance
    - Monitoring Declaring downtimes
    - Squid sharing
  - Plans
    - Geometry data
    - Optimization of RPMs for Frontier servers and Squids
    - AGIS, CVMFS
- AGIS news (A. Anisënkov)
  - After many discussions with DDM experts AGIS schema have been updated to satisfy DDM needs more precisely
    - AGIS API and WEB I/F need to be fixed and updated according to schema changes, partially updated, work in progress.
  - Recently implemented changed in AGIS schema allowed to improve downtime calendar
  - Frontier/Squid activity:
    - Still need some feedback and clear understanding on which are the Frontier/Squid requirements for AGIS
    - Will implement client API covered basic needs in ~2-3 weeks, so packages could be deployed and tested on client side



# TAGs

- TAG DB Monitoring, Performance, Scalability (F. Viegas)
  - Improvements in database performance and in service performance
    - Catalog integration in the services has given them capabilities for parallelizing queries across remote sites.
    - We can take advantage of this to query faster, even at the expense of more data redundancy across sites.
    - ELSSI Suite of Services is today in a position to be « distributed aware » and take full advantage of data and service catalog.
    - Queries can be broken down across servers, and parallelized for optimization.
- TAG related services (J. Cranshaw)
  - 2010 established a set of services with core functionality that are deployed and being maintained.
    - Extensions of this functionality are being developed.
  - Monitoring and user feedback are being incorporated into the development.
    - Better logging and testing.
    - Better integration with user groups: DAST, tutorials, physics groups.
  - Focus is now on performance and response to user needs.
  - Manpower issues!
- See also the talks in the PAT session:
  - QZ Zhang on the ELSSI suite
  - W. Ehrenfeld on TAG validation
  - The TAG discussion led by A. Farbin





# Announcements

- New version of Oracle (11g) will be introduced at the end of 2011
  - A test RAC is now available for application tests
  - All database application developers must get in contact with Gancho to schedule these tests during next few weeks
  
- Next step of the consolidation of distributed databases
  - The 3D databases at NDGF, SARA and CNAF will be decommissioned next week
  - There should be no consequence on any application as Squids are active at all these sites, but just in case...



# Thanks Florbela!



😊 A special “THANK YOU” slide 😊



A special, warm  
**OBRIGADO**  
*and*  
**БЛАГОДАРЯ\***  
to  
**FLORBELA**

for being a great colleague in all these years we worked together!

\* Благодаря = ‘Thank you’ in my native alphabet and language ( Bulgarian )

