

LCG Database Deployment & Persistency Workshop

Workshop Summary (my impressions at least)

Dirk Duellmann, CERN IT

3rd Database Workshop

- After last two workshops which focused on Database functionality and technology...
- Focus on preparation for large scale deployment
 - 3 days (Tue in B40-S01-A01)
 - Capture current deployment plans
 - Discuss proposed services and impact on applications
 - Continue service validation with real apps/workload
 - Define and schedule production setup at LCG Tier sites

Requirements / Deployment Plans

- ALICE
 - Databases at T0, outside only for LCG middleware, conditions file based, impact on file catalog
- CMS
 - Databases at T0, conditions via FroNtier/Squid/POOL/ORACLE
 - Online->Offline oracle replication from P5

Requirements / Deployment Plans

- ATLAS
 - Databases at T0,T1,T2
 - T0→T1 (COOL, GEOM, TAGS) apps scheduled for test
 - Online→offline test (COOL via streams)
- LHCb
 - Databases at T0,T1
 - T0→T1 COOL and Bookeeping(?)

Requirements

- Volume, CPU and client number estimates still evolving
 - Will stay uncertain to guarantee that service size can be predicted well in advance
 - DB Services at CERN and T1 need to be able to react flexible
- Iterate via application validation tests
 - Determine DB resource requirements
 - Optimise key software
 - Adjust experiment access model and start again
- Needs significant effort and flexibility from all sides
 - Experiments, db service and development teams

Physics DB Services @ CERN

- Database clusters (RAC) are in production now !!
 - Big effort of Maria's team in parallel to ongoing services and significant support load
- Hope to re-focus resources to consultancy and service improvement
 - Standardized account setup for smother transition between service levels and more stable client side connection information
 - Proposed new storage organization provides improved performance and uses available disk volume to decrease recovery latency
 - Backup policy review and read-only data are important to insure that tape backup volume stays scalable as volume grows
 - DB Server monitoring for developers/deployment integrated into LEMON
 - Unavailability caused by security patches and overload are most significant - applications and services need to retry/failover

LCG Applications

- POOL/CORAL release aims to provide important deployment enhancements wrt authentication, service lookup, connection retries/waiting to higher level components.
 - Rapid production release required to support test activities
- ROOT work plan focuses on orthogonal areas (analysis performance) and integration of CORAL could add
- COOL T0 test have been performed with a parameterized benchmark - so far DCS data. Adaptation to offline use cases needed

Apps Tests at T0

- Some promising results have been obtained
 - But only after
- Need to get developers, testers and DBAs closer together and increase the speed of test->db analysis->code fix loop
- Conditions data is the biggest database application and will drive optimisation of code and sizing of the service
 - Smaller apps will profit with less effort

RAC

- Several sites are testing/deploying Oracle clusters
 - CERN, CNAF, FNAL, GridKA, IN2P3, RAL
- Several experiments foresee Oracle services at their pit
- Propose to organize meetings for more detailed DBA discussions
 - Tier 1 DB teams and Online DB teams
 - Coordinate test plans
 - RAC setup and existing test results
 - Storage configuration and performance tests

Streams

- Have been tested for simple workloads (file catalogs)
- Now moving to test for conditions data
 - Online→offline for ATLAS, CMS
 - Offline→T1 for ATLAS, LHCb
- Many tests upcoming requiring service and consultancy
 - 0.5 FTE DBA support the CERN side likely to become a bottleneck

FroNtier

- CMS Baseline for conditions data
 - No database outside T0 required
 - Squid setup on T1 and T2
- Now integrated with POOL via transparent plugin
 - Early testing phase in CMS
 - Interest from other experiments (ATLAS/LHCb)
 - FroNtier test setup available in 3D test bed

LCG Certificate Support

- Easier for 3-tier apps
- Still an issue for client-server db applications
 - X509 authentication in Oracle and MySQL
 - Proxy certs can be made work with MySQL but fail so far with Oracle
 - Authorization (mapping of VOMS to DB roles) still missing
- Stop-gap solution (read-only access outside T0) acceptable for security and experiments for ~6month
 - Monitoring will important and needs real user id (DN)
- Little manpower left in the project
 - ASCC contributing in the area of Oracle security infrastructure setup
 - Need a test system and review manpower @ CERN

Tier 2 Setup

- Only little effort from 3D so far
 - Will change with full slice test for COOL now
- BNL / ATLAS have most deployment experience
- Propose to define and document standard s/w installation
 - Need more participation of prototype Tier 2 sites

FNAL DBA Team

- Does not need to provide tier 1 database services in the CMS model
- Has contributed significant streams and physics db deployment experience
- Is heavily involved in CMS applications validation
- Would be very appreciated for consultancy

Proposed Tier 1 Architecture

- Experiment plans firming up on application list
 - DB volume and access predictions still rather uncertain
- Several sites looking at db clusters
- Propose to buy modular h/w which at least can be turned into clusters later - eg
 - Shareable (SAN) storage (FC based disk arrays)
 - Dual CPU boxes with 2GB with mirrored disk for system and database s/w

Proposed Tier 1 DB Server Setup

- Propose to setup (preferably as cluster)
 - 2 nodes per supported experiment (ATLAS, LHCb)
 - 2 nodes for the LCG middleware services (LFC, FTS, VOMS)
 - As Oracle server node if site provides Oracle service
 - As specified by GD for MySQL based implementation
 - Note: LFC may come with Oracle constraint for LHCb replication
 - Squid node (CMS - specs to be clarified)
- Target release Oracle 10gR2 (first patch set)
 - Sites should consider RedHat ES 3.0 to insure Oracle support

Proposed Setup Schedule

- November: h/w setup defined and plan to PEB/GDB
- January: h/w acceptance tests, RAC setup
- Begin February: Tier 1 DB workshop
- February: Apps and streams setup with T0
- March: production starts