



ATLAS

3D Requests

LCG3D Kickoff Workshop

CERN, Geneva, Switzerland

December 13, 2004

Alexandre Vaniachine, Jeff Tseng, Andrea Formica



Outline

- ATLAS Database Project and LCG3D
- Migration to Oracle and 2005 data volume
- 3D data inventory efforts
- Inventory: data vs. database applications
- Parade of ATLAS database applications
 - Offline
 - Technical Coordination (detector production DB)
 - Online
- Conclusions



ATLAS Database Project

- All ATLAS database domains are now under the umbrella of the new Database Project
 - Launched in May 2004
 - Led by Richard Hawkings and Torre Wenaus
 - <http://atlas.web.cern.ch/Atlas/GROUPS/DATABASE/project>
 - Emphasis on integration of different domains
 - Provides representatives to LCG3D



ATLAS and LCG3D

- ATLAS makes use of a distributed infrastructure of relational databases for access to various types of non-event data and event metadata
- Distribution, replication, and synchronization of these databases, which are likely to employ more than one database technology:
 - Oracle in larger centers
 - MySQL in smaller settings and for smaller applications must be supported according to the needs of the various database service client applications
- ATLAS favors common LHC-wide solutions and will work with the LCG3D and the other experiments to define and deploy common solutions



Migration to Oracle

- ATLAS acquire significant experience in MySQL database services, been critical for:
 - 2004 Combined Test Beam effort
 - Current Data Challenge 2 efforts
- Following CERN policy and encouragement to concentrate our database usage around Oracle, we are making increasing use of these services and support



Volume Request for 2005

- 'Psuedo-Cocotime' request made in October 2004
- Combined with CPU, AFS, tapes requests made by G Poulard

Usage	2005 space – GB
Geometry primary numbers	few GB
Conditions data	100 GB
TechCoord production data	100 GB
Event meta-data	250 GB
3D project data	?

- Volumes given as estimated raw data size, not including backup, indexing, mirroring



3D Data Inventory

- ATLAS started collecting data inventory and submitted required templates to 3D
 - A very sound initiative from 3D
 - Focus is now shifting from the simple data volume estimated to database applications inventory efforts
 - ATLAS parades its first database application ready for distribution at this workshop Vakhtang Tsulaia talk:
 - ATLAS Detector Description DB
 - Note: not only distributed application must be counted
- A sample of ATLAS database application follows next



ATLAS Database Applications

- 1) Inventory for (broadly defined) offline sw domain
 - Tag Collector – software infrastructure support
 - AMI – Bookkeeping DB (file-level metadata)
 - Production system DB
 - Detector Description DB
 - ConditionsDB IOV DB
 - Conditions data payload DB
 - Reliable File Transfer DB – ATLAS DMS: DQ
 - File Catalogs: POOL, RLS,...
 - Collections DB (event-level metadata)



Tag Collector DB

- Supported by LPSC Grenoble (*Solveig Albrand, Jérôme Fulachier*)
- Currently in MySQL
- Critical application for software development infrastructure
- Migration to CERN Oracle is in plans
- Not distributed, a geographically separate replica required as a failover service
- Application is merging within AMI framework



ATLAS Metadata Interface

- AMI is supported by LPSC Grenoble (*Solveig Albrand, Jérôme Fulachier*)
- Currently in MySQL
- Bookkeeping metadata database for:
 - DCs - Data Challenges
 - CTB - Combined Test Beam
 - ADA – ATLAS Distributed Analysis
- In 2005 migration to Oracle (may be not at CERN)
- Distributed Oracle services is in the plans
 - CERN replica server could be a failover node



Production DB

- ATLAS Production System database
- Data processing tasks configuration, jobs submission, bookkeeping
- Supported by CERN (*Luc Goossens*)
- MySQL prototypes used in DC1
- Designed in Oracle for DC2
- Hosted on pdb01
- Provided first Oracle operational experience in ATLAS



Production DB book-keeping remarks

- Significant performance problems seen
 - Some addressed with improved and tuned applications
 - Some show need for more servers, dedicated to specific tasks
- What lessons can we learn from this?
 - DB volume is not the whole story, usage patterns place heavy demands on the service
 - IT application consultancy needs to be involved to avoid poor design / inefficient DB usage
 - In the imperfect distributed world, load estimates are going to be very approximate – IT needs to provide appropriate server capacity headroom



Lifecycle Patterns

Disclaimer: observations by an outsider, not by the prodsys team

- Focus is on general patterns - what we should be ready for in next ATLAS database applications deployments
- ATLAS ProductionDB operational experience:
 - Initially designed for a few writers
 - writers count exceeded expectations by ten
 - Initially not designed for monitoring
 - used heavily to monitor production progress
 - Best Practices are hard to follow under pressure of DC2
- Should the database application be ready for:
 - ten times more users than was foreseen
 - will be used in a data access/queries pattern never foreseen
- Emerging/familiar pattern? (Already in Murphy's laws?)



Detector Description DB

- DDDB – primary numbers for Detector Description
- Supported by U Pittsburgh (*Joe Boudreau, Vakhtang Tsulaia*)
- Initially in MySQL (NOVA)
- Redesigned for Oracle
- Focus on Best Practices from the start
 - fruitful interactions with IT/DB
- Being deployed for DC2 production use
 - Served data to thousands of reconstruction jobs on worldwide grids
- Replication tools
 - JDBC-based (Julius Hrivnac, see his talk at workshop)
 - RAL-based (*Vakhtang Tsulaia*) work in progress
- First ATLAS distributed database application ready for 3D testbed
- Expected to produce high (read-only) server load in production



ConditionsDB

- Note: separation of IOV data and payload data (NovaBlob DB)
- Supported by Lisbon group (led by *Antonio Amorim*) and BNL/ANL team (*Hong Ma, Alexandre Vaniachine*)
- Both currently in MySQL (with payload also in POOL files)
 - Integrated:
 - Web browser: CondDBrowser
 - Replication tools (*Sven Schmidt*)
 - Served well during Combined Test Beam
 - **This week: first large scale production exercise**
 - Deployed for Commissioning activities at point 1

Database applications requirements:

- Must be replicated within T0: online-offline
- Have to be distributed beyond T0 (scale to be defined)

Expecting Oracle implementation in March



Data Management System

- ATLAS Data Management System – Don Quixote (*Miguel Branco*) elegantly integrates a zoo of database applications:
- Mostly File Catalogs: POOL, RLS, ...
 - mixture of MySQL and Oracle
- Reliable File Transfer DB
 - Currently in MySQL (*Miguel Branco*)
 - Integrated in Globus GT4 RFT
- Not replicated/distributed (to our knowledge)
 - except failover needs



Collections DB

This database application is from POOL project

- significant ATLAS contributions (*David Malon, Kristo Karr*)
- Very data volume intensive
 - Event-level metadata (tag database)
- Currently in MySQL
- Heterogeneous replication tools (*Julius Hrivnac*)
- DC2 Collections data is ready for 3D testbed



Technical Coordination

- ATLAS equipment management database, used for managing racks, cables
 - Common project with CMS, runs on pdb servers – usage increasing as detector installation / commissioning is proceeding
- Use of MTF and EDMS for handling production data and documentation
 - User perception that database infrastructure is slow (but also slow interfaces)
- Subdetector production data in various technologies
 - Encouraging data migration to CERN Oracle for long term security
- No replication/distribution requests



Online application types

- DAQ Configuration
 - Tdaq current state (ReIDB should replace XML files)
- Sub-detector Configuration (Electronics config, Cable mapping)
 - XML files or ReIDB tables , versioning needed
- DCS (Temperature, HV, pressure,...)
 - Large use in H8 during 2004 testbeam (Condition DB)
 - Activity foreseen for 2005 during commissioning
- Calib/Align
 - Large amount of data in ReIDB (used in Muon system align. also as data source, i.e. for determining the corrections)
 - Sensors data available during commissioning in 2005 (Bfield, Align,...)
 - Usage of POOL(ROOT) files referenced in relational tables
- Monitoring
 - Histograms important during commissioning
 - ROOT files referenced in a relational DB



Online inventory remarks

- Inventory for 2005 (preliminar)
 - Conditions Data : requirements from several subsystems already available (data volume/year)
 - Central DCS : ~4GB
 - Muons (DCS+Align+BField): <15GB
 - Lar : ~100 MB
 - Configuration Data : not yet well defined
 - TDAQ : ~1GB
 - Monitoring data : mainly histogram files (large data volume)
 - Do they need replication ?
 - Number of clients and type of access still difficult to be estimated. Extrapolation of H8 situation not always realistic
- To Do : try to collect missing information and needs for replication/distribution in the online community



Conclusions

- ATLAS Oracle-CERN usage is now ramping up fast
- 2005 will see (at least) three very high profile activities:
 - Production DB for offline production
 - dozens of concurrent writers
 - Geometry DB for offline production
 - thousands of concurrent readers
 - Online (conditions) DB for commissioning focussed at point 1
- In all cases, will need significant dedicated resources
 - Database problems if any will be very 'visible'
 - Other smaller scale activities will continue
 - TAG database may require significant volumes