



# ATLAS applications and plans

**LCG Database Deployment and Persistency  
Workshop**

17-Oct-2005,CERN

Stefan Stonjek (Oxford), Torre Wenaus (BNL)



# Outline

- Databases at ATLAS
  - Online
  - Geometry
  - Conditions
  - Other
- Distributed databases
- Outlook
- Summary and Conclusions

# ATLAS Database / Data Management Project

- Responsible for DB/DM activities across ATLAS:
  - DBs for detector production & installation, survey, detector geometry
  - Online configuration, bookkeeping, run conditions
  - Online and offline calibrations & alignments
  - Event data and metadata
  - Offline processing configuration and bookkeeping
  - Distributed data management (file-based data)
  - Distributed database infrastructure and services
- All of these rely on DB services at CERN and at external institutes, and a distributed DB infrastructure knitting these together
- **Consequently we strongly support and rely on the 3D project!**

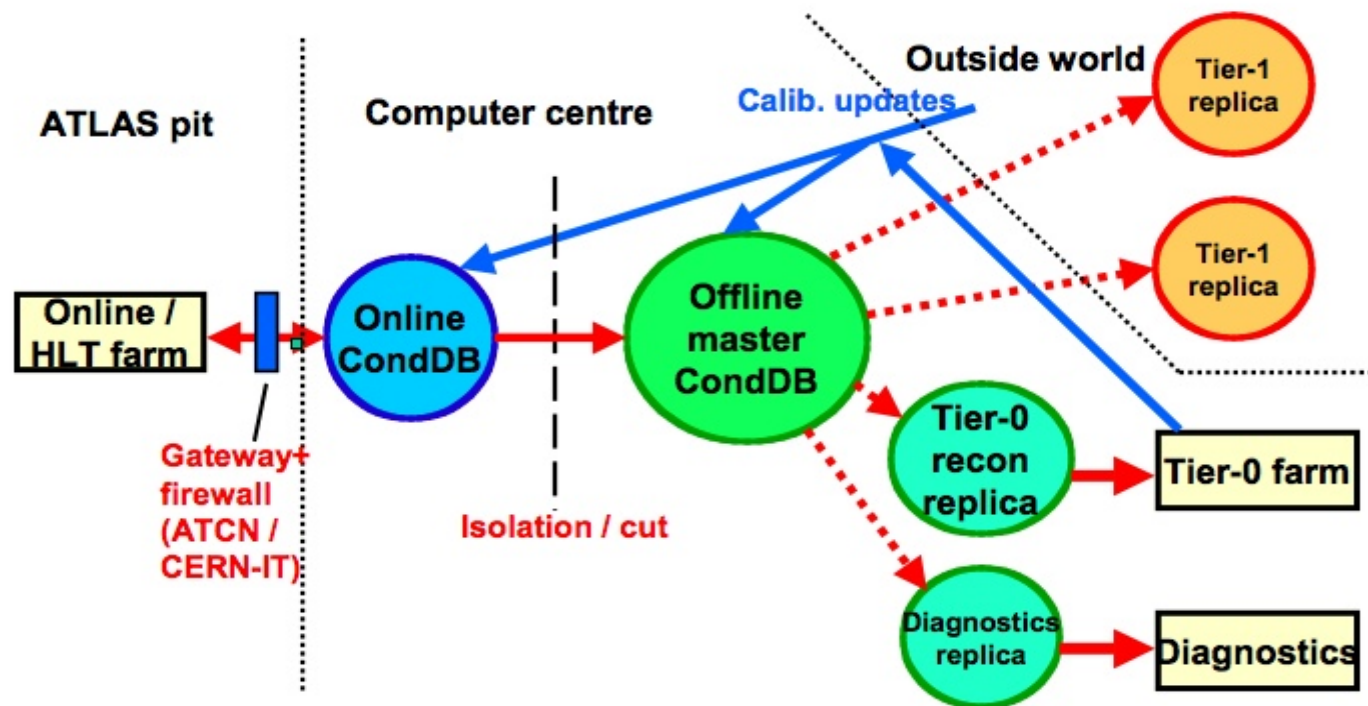
## Subprojects

Production DBs  
Installation DBs  
Geometry DB  
Online DBs  
Calib/Align  
Conditions DB  
Event Data  
Distributed DM  
Offl Processing  
DB Services  
SW Support

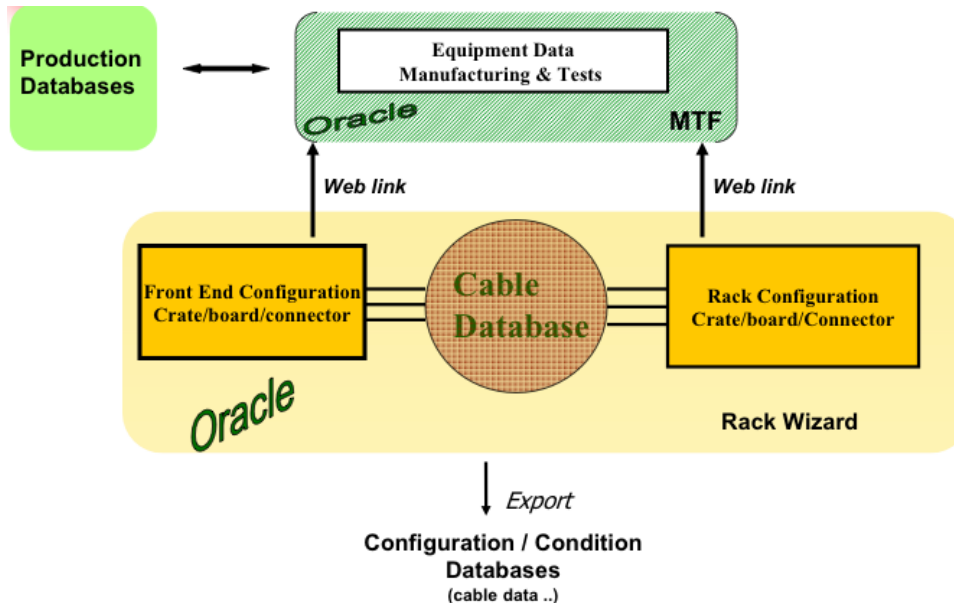
# Online Databases

- ATLAS Online uses standardized DB tools from the DB/DM project (and mostly from LCG AA)
- RAL used as standard DB interface, either directly or indirectly (COOL)
  - Direct usages: L1 trigger configuration, TDAQ OKS configuration system
- COOL conditions DB is successor to Lisbon DB for time dependent conditions, configuration data
  - Links between COOL and online (PVSS, information system) in place
  - PVSS data sent to PVSS-Oracle and then to COOL (CERN Oracle)
- Strategy of DB access via 'offline' tools (COOL/POOL) *and* via direct access to the back end DB popular in online
- **Joint IT/ATLAS project to test online Oracle DB strategy being established**
  - Online Oracle DB physically resident at IT, on ATLAS-online secure subnet
  - Data exported from there to offline central Oracle DB, also IT resident

# Online/T0 DB System Architecture



# Technical Coordination Databases



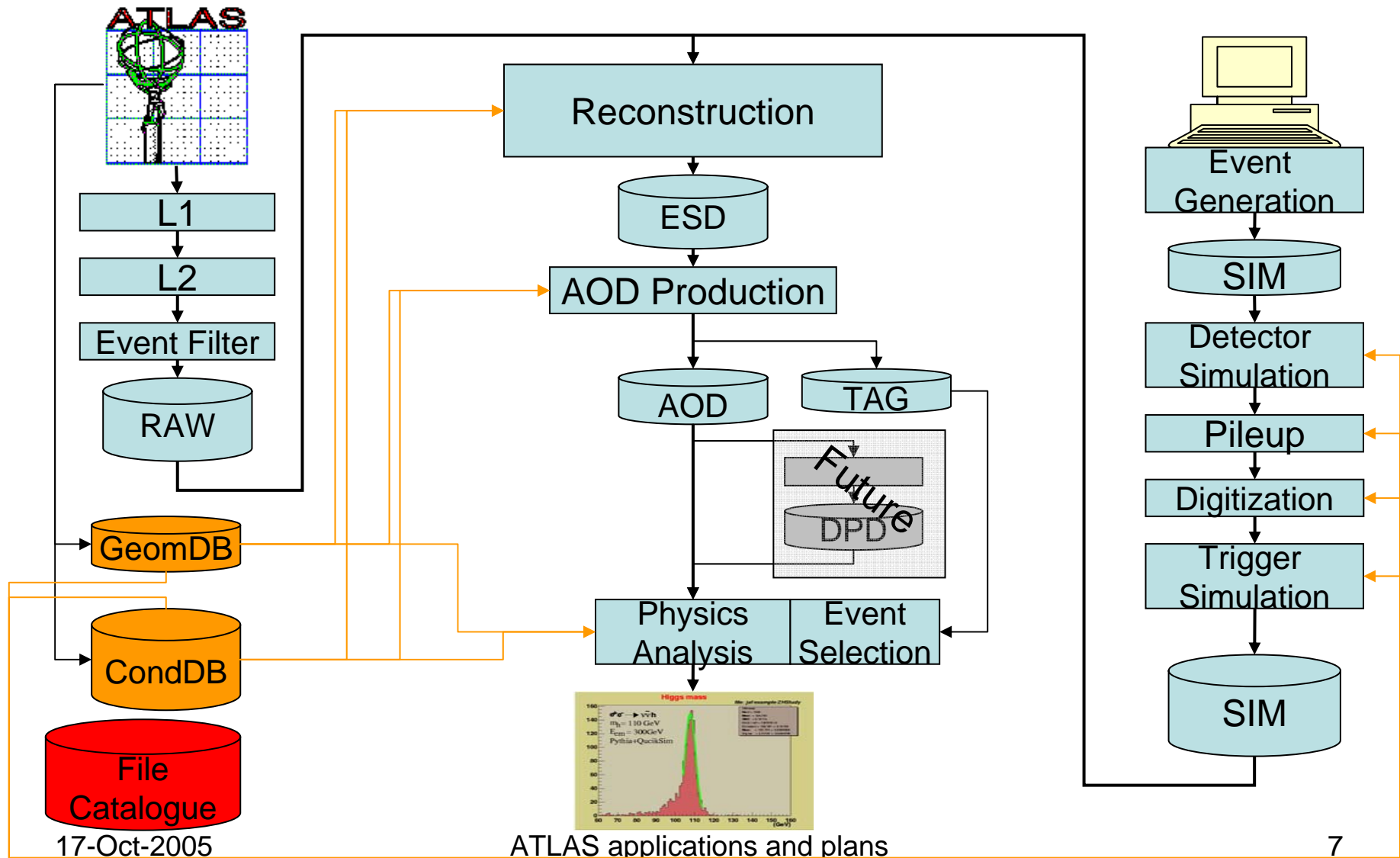
**CERN Oracle based**

**Long since in production;  
relatively stable and well  
advanced; relatively light usage  
load**

To come: selective export of data  
needed offline to conditions DB

- Detector production and installation data (CERN Oracle MTF)
  - Production data, registering/tracking installed equipment
  - Data being gathered in CERN Oracle from disparate subsystem DBs
  - Interfaces and import tools in development
- Rack and cabling databases (CERN Oracle)
  - Rack configuration and content, cables, cable trays, patch panels
  - Web-based interface (Rack Wizard, jointly developed with CMS)
  - Implementation largely complete and population well advanced
- Survey DB (CERN Oracle) in development

# ATLAS data flow



# Geometry Database (1)

- Geometry DB based on RAL and Hierarchical Versioning System, jointly developed by ATLAS/LCG
  - in production for ~1 year
- presented at the first 3D workshop - first ATLAS database application ready for distribution
  - ready for the 3D testbed Slice Test
- Master instance (read/**write**) in CERN Oracle
  - Write usage light; read usage should become light!
- Read usage focused on distributed replicas, not the master



# Geometry Database (2)

- Infrastructure in place and in use to
  - Define and tag new geometry versions in the master DB
  - Replicate tagged geometries to MySQL and SQLite for use via RAL
  - Package and distribute replicas for use at external sites
- **MySQL replication/distribution mechanisms in use for 3D testing of Oracle T0/T1 -> MySQL T2**
  - **Octopus/JDBC based replication (in test)**

# Conditions DB – COOL (1)

- COOL used for conditions data across ATLAS subdetectors and online
  - ATLAS is the main experiment contributor to its development
  - Initiative started to use COOL to manage time-dependent aspects of subsystem configurations
- COOL now fully integrated with ATLAS offline framework Athena
- Deployment plan written and well received within ATLAS and by IT
- Usage model: write centrally (CERN Oracle), read mainly via caches/replicas
- **Scalability and performance testing/debugging of COOL and underlying DB services being tested with realistic ATLAS online workloads**

# Conditions DB – COOL (2)

- **CERN Oracle service in place; data content still small**
  - Offline scalability tests in distributed production planned for 2006Q1
- **Full ATLAS data volume and access patterns not yet well understood!**
  - Large volume data will be stored in (POOL) files, managed by DDM
  - Possible application of object relational POOL being explored
- **Supporting scalable distributed usage is main development focus now**
  - Copy/extraction tools under development
  - Interested in FroNTier or similar, no work done yet
- UI tools (interactive python, web browser) in development

# Event Data

- POOL/ROOT files store event data; well established in production
- Schema evolution system with distinct and independently optimized transient/persistent representations of the event model now in development
- Most recent large scale production produced all ATLAS-common event data model components: new were AOD and **event tags**
  - AOD too slow (no real surprise) and major focus now is its optimization
  - Performance study for tags based on POOL/ROOT files, Oracle and MySQL just completed. Improvements needed, but captured physicist interest
- Integration of Athena event I/O with DDM and **event processing metadata (AMI)** systems just beginning
  - Dataset-based discovery, replication, access and storage/publication of data
- Provenance data collection, storage and management in early development
- **DB service requirements from tags and event processing metadata/provenance**
  - Much remains to be done before we can be reliably quantitative
  - Tag requirement estimates for CERN Oracle provided for 2005 (with actual usage much less) and being updated for 2006
  - Master AMI Oracle DB instance currently at IN2P3

# Offline Processing

- offline production database application entering a second generation with the development of a new ATLAS production system
  - Development responsibility lies outside DB/DM project
- **Was the target of extensive performance and optimization studies in the earlier (DC2) generation; focus of new design (of DB and its clients) is performance and scalability**
  - **Good results so far in recent scalability tests on CERN Oracle**
- AMI metadata system being adapted for the new production system as repository for production (task) metadata and physics-level dataset metadata
- DDM system provides data handling services to the production system
- Completion, integration and deployment of the new production system underway now

# Distributed Data Management (DQ2)

- Redesigned in light of DC2 experience for scalability, robustness, flexibility
  - Fewer middleware dependencies
  - Dataset (hierarchical, versioned file collections) based logical organization of files
- Designed to meet data handling requirements of ATLAS Computing Model
  - From raw data archiving through global managed production & analysis to individual physics analysis at home institutes
  - Aggregate data volumes of 10s of petabytes/year from 2008
- Basic file handling middleware at the foundation (FTS etc, SRM, LFC)
- Above, loosely coupled distributed services providing logical/physical cataloging and file movement
  - RDBMS back ends; wide use of POOL FC and its RAL implementation
  - MySQL now; beginning to mix in Oracle
- Scalability tests over the summer; now being deployed on ATLAS SC3 (Tier 0 test)
- Integrated and operating in US part of new production system (Panda)
- Much to learn operationally before we can be quantitative on DB service requirements
  - **Many scaling 'knobs' in the system to be explored: mix of Oracle, MySQL, grid catalogs, catalog partitions and/or replicas, caching, system instances, ...**
  - **The learning begins now...**

# Distributed DB Services

- The DB/DM subproject charged with providing distributed DB services to ATLAS
  - Led by Sasha Vaniachine
- Close relationship with 3D
- Has done a heroic job with too little manpower (with some welcome increases recently)
- Most recently, Data Challenge 2 production, Combined Test Beam analysis, Rome Physics Workshop production, regular onslaughts from power users, ...
- As the picture from a recent production post-mortem talk indicates
  - *we progressed even further than expected!*
- Sasha will talk on Wednesday



# Summary of apps, requirements, priorities and production status/plans

## Applications

- Reconstruction
- AOD Production
- Physics Analysis
- MC Generation

- For test plans see talk on Wednesday
- Distributed Data Management is different issue

## Databases

- Geometry
  - Known to produce load
  - Plan to distribute as SQLite files
- Conditions
  - In the process of switching to COOL



# 3D-Relevant Timeline

- All CERN Oracle ATLAS applications have been migrated to RAC!
- Objective is to commission and validate a scalable production DDM system (DQ2) in late 2005, so that
  - CSC activities can operate on the foundation of a stable, preferment and low-maintenance DDM system
  - DDM experience in CSC drives (hopefully small scale) tuning and redefining of the system, not a major overhaul
- Principal context, and dependency, for DDM commissioning and validation is Service Challenge 3

# Some ATLAS DB Concerns (1)

- Scalable distributed access to conditions data
  - COOL copy/extraction tools in development, and in good hands; will come but aren't there yet
  - FroNTier approach of great interest but untouched in ATLAS for lack of manpower
    - FroNTier/RAL integration is welcome, we need to look at it!
  - DDM already deployed in a limited way for calibration data file management, but needs to be scaled up and the divide between file- and DB-based conditions data better understood
  - Role of object relational POOL and implications for distributed access still to be understood

# Some ATLAS DB Concerns (2)

- Manpower, of course
  - ATLAS in-house Oracle expert(s) clearly essential
    - a hire is in progress
  - DB services and DDM operations under-resourced
  - Does IT have enough resources? 3D? A great team, but big enough?
- Will Oracle be sufficiently scalable? Will its behavior be comprehensible?
- Our 3D resource requirement/plan information is not what it should be!

# Conclusion (1)

- ATLAS DB/DM is well aligned with, draws heavily from, and contributes to LCG 3D and LCG AA/persistency; we depend on them being strongly supported and will continue to support them as best we can
- The ‘easier’ applications from the DB services point of view are well established in production, reasonably well understood, and relatively light in their service/distribution requirements
  - TC databases (except survey), geometry database

# Conclusion (2)

- For the most critical of the rest, the ‘final’ applications now exist in various states of maturity, but more scale/usage information and operational experience is needed before we can be reliably concrete
  - Conditions DB, event tags?, production DB, DDM
- For the remainder, applications and even strategies are still immature to non-existent (largely because they relate to the still-evolving analysis model)
  - Event processing metadata, event tags?, physics dataset selection, provenance metadata

*The End*

# Further Information

- Wiki
  - COOL in ATLAS
    - <https://uimon.cern.ch/twiki/bin/view/Atlas/CoolATLAS>
  - COOL in Athena
    - <https://uimon.cern.ch/twiki/bin/view/Atlas/CoolAthena>