Databases and POOL

Storage and Data Management Technologies Session

29 June 2004

http://cern.ch/jamie/dti-june2004.ppt

Jamie Shiers, IT-DB, CERN

http://cern.ch/db/

Introduction

- Overview of Database Services at CERN
- The POOL Persistency Framework
- Outlook for the future
- Summary

Database Services at CERN

- Currently based on Oracle technology
  - Oracle DB Enterprise Edition, Partitioning + RAC Options
  - Oracle Internet Application Server, Enterprise Edition (iAS)
  - Oracle Enterprise Manager (OEM)
- Choice of Oracle dates back to early 1980s for LEP construction
  - Primary requirement then: ***relational*** as opposed to ***hierarchical*** or ***network*** DB
  - Most recent contract amendment December 2002 following Market Survey and FC paper
  - Server deployment platforms:
    - Primarily Solaris for non-Physics services
    - Primarily Linux for Physics (Red Hat Enterprise Linux ES)
- Data volumes in the hundred GB – few TB range
  - Expected to grow significantly as LHC enters production
    - At least one order of magnitude – could be 2 – 3 (event-related data, as for COMPASS + HARP)

DB Services - Overview

- Infrastructure Databases (+iAS)
  - Extension of original services to cover all aspects of lab's technical infrastructure
    - Accelerator assembly, disassembly, run parameters, …
    - Engineering documents (EDMS), …
    - Network Database, Security Event logging, Machine monitoring, …
- Administrative Infrastructure Services (AIS)
  - Underlying DB + iAS services for lab's "e-business"
    - Expenditure tracking (CET), Orders, shipping, leave, training (EDH), Human Resources Toolkit (HRT) + many more…
    - Combination of commercial s/w + in-house developments (IT-AIS)
- Physics Database Services
  - Detector assembly, monitoring, conditions, etc, Book-keeping, …
  - File Catalogs + File-level Meta-data
  - Event-level Meta-data

} ***Link to the Grid***

POOL – Persistency Framework

- Project in the LCG Applications Area:
  - Provides storage and retrieval experiment data and associated meta data in a distributed and Grid enabled fashion (100PB +)
  - Data comes in different volumes
    - Event data, physics and detector simulation,
    - Detector data and bookkeeping data
  - Data comes in various forms
    - Bulk data
    - Time dependent data
    - Metadata
- This challenge is faced by a hybrid technology approach
    - C++ object streaming technology for bulk data
      - Using ROOT framework
    - Transactionally-safe services for catalogs, collections and meta data
      - Using RDBMS systems such as Oracle, MySQL,
- Follows on from work done in 1990s using Object Databases
  - Market did not take off as predicted –too many risks!
  - Single technology not optimal for all different data types and access patterns
  - Not Grid-enabled!

POOL Related Services

- POOL is distributed through LCG Applications Area channels
  - Software Process and Infrastructure (SPI)
    - Shared libraries and command-line tools
- POOL users and applications depend on services such as:
  - LCG File Catalog + File-level Meta-data
    - Currently service based on Oracle iAS + DB
  - Storage services, including Castor
    - Data streamed to / from files using ROOT I/O
  - (Eventually) Database Services
    - RDBMS backend for conditions and other data

DB Services - Deployment

- Deployment platforms: Solaris and Linux
  - Expect future Physics services to be Linux based
  - Has Linux (and PC h/w) attained sufficient stability for other services?
- Linux services:
  - Farm nodes for stateless services (iAS)
  - Disk services for DB servers
- Neither optimal: re-evaluating SAN-based configurations with PC servers
  - Currently determining optimal configuration for both iAS servers and DB servers in a RAC environment

Future Outlook

- Taking advantage of synergy created by CERN re-organisation to re-engineer services to be as standard as possible
  - Use common building blocks – avoid bespoke solutions and unnecessary complexity
  - Use most appropriate technology for the job in hand
    - The most expensive (and complex) solution is not necessarily the best!
- Understand how to offer services in Distributed Environment for LHC experiments
  - Data exchange between CERN and Tier1 / Tier2 sites
- Understand how to handle potential explosion in data volume with strictly limited resources

Summary

- CERN benefits from close relationship with Oracle which has enabled us to influence product
  - Support for ultra-large databases (ULDB), support for IEEE floating point numbers (single and double) etc.
- Primary challenges are scale and distributed nature
  - Databases up to 100TB, deployed at some 10 sites (Oracle) plus exchange with ~100 smaller sites (MySQL)
- POOL fills a gap that is not satisfactorily addressed by commercial market-place
  - ODBMS potential not realised – C++ niche market?