LCG Database Deployment &
Persistency Workshop

Workshop Goals

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

LHCC questions to LCG TDR

- Databases were a problem in the beginning of RUN II and the consequence is that they are often not used. Please clarify the plans…
- The 3D project incorporated the experience from FNAL and other sites
  - non-linear response of database systems to load variation
  - late integration of database applications leads to uncertainties about their resource requirements
- Database strategy
  - Keep service implementation as flexible as possible
    - Eg extensible linux based database clusters at T0
  - Introduce a joint validation step between development and production deployment
    - Optimise against reference workload
    - Size the resource requirements
    - Part of service level agreement

Proposed Test and Implementation Plans

- Perform series of distribution tests (after T0 validation)
  - **COOL, FronNtier**, meta data, geometry, tags, file and meta data catalogs
    - Including online->offline tests
  - Database testbed with 6 LCG Tier 1 sites and monitoring available
- Upcoming database workshop will be used to turn performance results into h/w estimate
  - Result will go as proposal to PEB and GDB
  - Propose for implementation by LCG sites after approval

3D Participants and Responsibilities

- Will the deployment count on manpower provided by the experiments?
- LCG 3D is a joint project between
  - Service users: experiments and s/w projects
  - Service providers: LCG tier sites including CERN
- Project itself only coordination resources (2 FTE from CERN)
  - Rely fully on experiment manpower to define and validate application requirements and to participate in deployment
  - Rely fully on site resources for implementation and deployment

Session Goals: Experiment/project plans

- Review the experiment plans, priorities and requirements for production in the LCG database services
  - End-to-end: Online->offline(T0)->T1->T2
  - Be as concrete as possible:
    - Complete list of apps and priorities and production dates for each tier
    - Requirements: volume, number of clients, number of accesses (eg IOV lookups) required
    - Status of the software
      - Dates for apps and workload tests (writer/reader) available
- Would appreciate a summary slide with a table from each experiment and project so that we can collect all info for the resource discussion towards the end of the workshop

Project/Experiment feedback

- Feedback on the current services at CERN Tier 0
  - What needs to be improved? What works well?
  - Is the balance between experiment development support and service provision/development ok?
  - Is the balance of resources between T0 service and external site coordination ok?
- Not the place to ask for resources, but to help us to move the existing resources to the right task

Session Goals: LCG T1 sites

- Overview of state of the current services
- Statement of interest in providing LCG database services and list of supported experiments
  - Assumption all interested sites implement database requirement for LCG grid services (eg Jamie's list)
    - Will you participate only in LCG production or only in LCG apps validation?
  - Which resources are currently available/planned for production phase?
    - Be concrete: number of servers, CPU/Memory/storage
    - Testbed resources should not be counted as we'll need an independent validation testbed also during the production phase
  - Which local site changes are required to be ready?
  - Which which information is required for service implementation?
- Would appreciate one slide with a list of supported experiments and current resources for LCG production

LCG Site feedback

- Similar to the experiment feedback
  - Is the coordination effort in 3D sufficient?
  - Which areas would need to be improved?
  - Are installation scripts and documentation provided useful or rather complicating local service procedures?
  - Wiki, savannah, OEM setup - worth the effort?
- Goal: re-balance available 3D resources

Session Goals: Database Services and Impact on Applications

- Will present the current service and upcoming improvements at CERN
  - Many of these topics will directly be relevant for LCG tier sites
- Point out conventions and areas which require joint planning between service and application developers
  - Insure that applications which need high availability are prepared to minimise application downtime during expected service interventions
  - Insure the application which require large data amounts plan for efficient use of the backup/recovery infrastructure
  - Insure that application developers are aware of monitoring facilities and make best use of those for application diagnostics and optimisation

Session Goals: Distributed database Deployment

- Plan the continuation of application validation tests for apps requiring data distribution
  - Online->Offline
  - T0->T1->T2
- Walk through the proposed tests and schedule them together the experiments and sites
- Outcome: schedule for the 3D validation testbed

Session Goals: LCG Software and Experiment Experience

- Present the work plans of the relevant LCG projects in the persistency and database area
  - Focus software improvements to integrate with the database service and minimise the impact of service interventions on applications
  - Prototype implementation for applications which are not based on LCG common software packages
- In addition we would like to expose the first performance and scalability tests (Tier 0 only) with the LCG common software

Last Session: Production Setup

- Agree based on experiment requests and site availability and validation test results
  - List of sites which will provide support for the validated applications
  - Estimated hardware requirements as input to site h/w planning
  - Schedule for the implementation of the 3D production service
- And a short summary..