



3D Workshop Outline & Goals

Dirk Düllmann, CERN IT

More details at <http://lcg3d.cern.ch>



Why a LCG Database Deployment Project?

- LCG today provides an infrastructure for distributed access to file based data and file replication
- Physics applications (and grid services) require a similar services for data stored in relational databases
 - Several applications and services already use RDBMS
 - Several sites have already experience in providing RDBMS services
- **Goals for common project as part of LCG**
 - increase the availability and scalability of LCG and experiment components
 - allow applications to access data in a consistent, location independent way
 - allow to connect existing db services via data replication mechanisms
 - simplify a shared deployment and administration of this infrastructure during 24 x 7 operation
- **Need to bring service providers (site technology experts) closer to database users/developers to define a LCG database service**
 - Time frame: First deployment in 2005 data challenges (autumn '05)

Workshop Logistics



- **Workshop will take place at two locations**
 - Mornings: Bat 160-1-009
 - Mo & Tu Afternoon: 513-1-024 (VRVS available)
- **Wireless Network**
 - MAC registration is required
 - Follow the procedure on the given on the web page displayed in your web browser
- **Workshop Dinner on Tuesday?**
 - If sufficient people are interested on Tue evening 19:30
 - Please send an email by today lunch time if you plan to join
- **LHC experiment visit?**
 - Trying to arrange a visit to an LHC experiment site for Wed afternoon
 - Please let me know if you'd be interested to join

Workshop Outline



- Monday Morning - Experiment Requirements
- Goals
 - Clarify which database applications exist and need to be deployed in the LCG by autumn 2005
 - Main volume and functionality requirements
 - check service architecture
 - input to resource planning at T0/1/2
 - Which distribution mechanism is used (if any)?
 - Status of the application development
 - Input to a realistic test and deployment plan
- Result
 - Adapt service definition, collect a list key technologies to be tested against applications
 - Schedule for joint test of applications in the 3D testbed

Workshop Outline



- Monday Afternoon - Site Services
- Goals
 - Which services and policies are in place today and can be relied on for 2005?
 - Site production experience with physics applications
 - Available deployment and software development services
 - Overview on available resources for LCG work
 - Real resource discussion will be done via LCG channels
 - Short term plans until autumn 2005
- Result
 - Check with deployment plan to identify missing resources

Workshop Outline



- Tuesday Morning - Application-Service coupling
- Goal
 - Identify main areas of which required s/w development to couple applications to LCG database services
 - Overview of ongoing development activities in experiments and projects
- Result
 - Check project development plans against the provided services
 - Identify activities / applications which do not fit into a standard architecture and therefore require separate deployment plan and/or dedicated deployment resources

Workshop Outline



- Tuesday Afternoon - 3D distribution technologies
- Goal
 - Describe the two main technologies proposed for a standardized 3D service
 - Oracle Streams Replication
 - FroNtier based database access
 - Overview of experience with both
 - Identify main uncertainties which will need to be resolved in tests in the 3D test bed
- Result
 - Proposed items for a 3D test plan

Workshop Outline



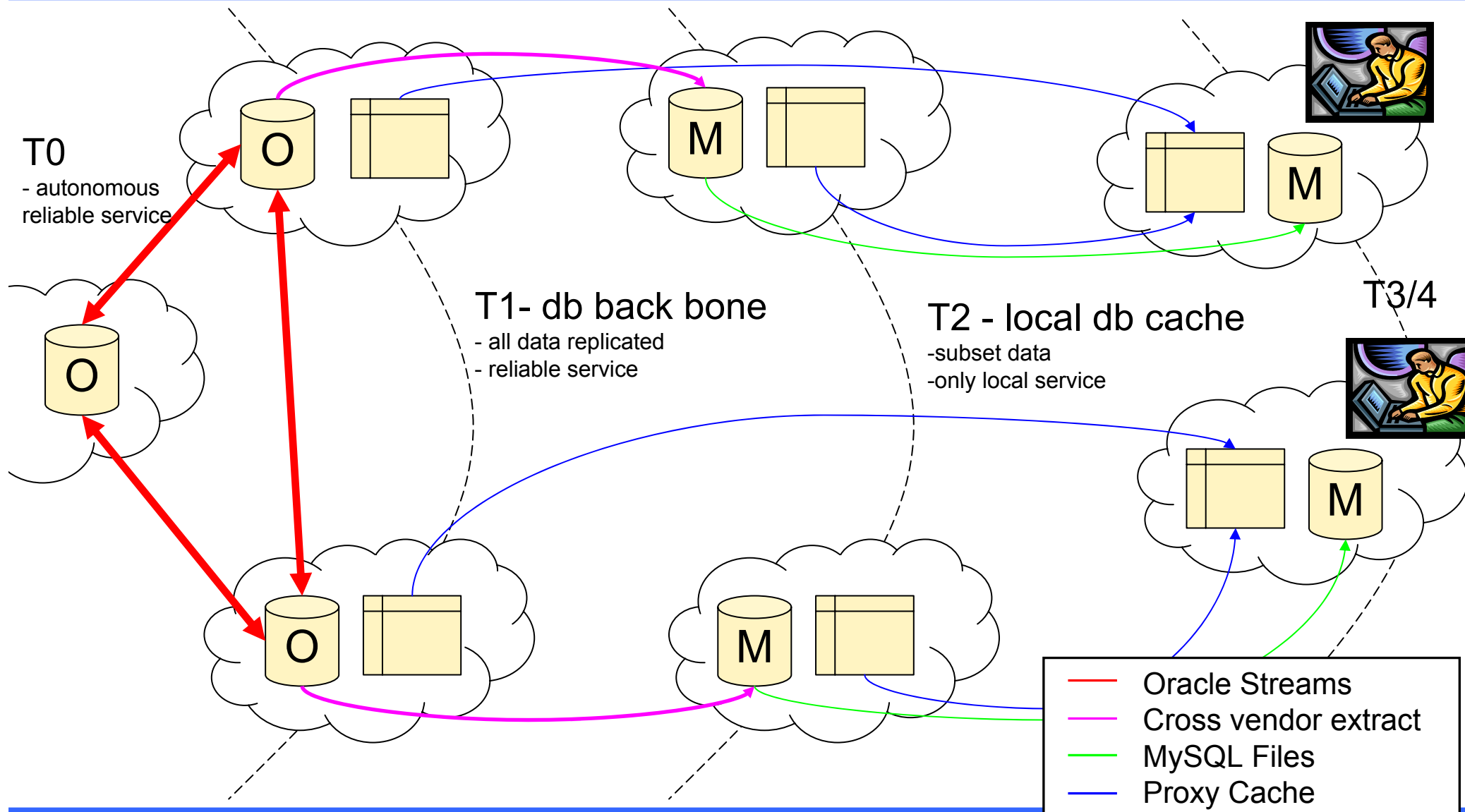
- Wednesday Morning - Persistency Framework
- Goal
 - Overview of developments in POOL/CondDB which impact on distributed database services
 - Describe the development plan for other remaining features of LCG-PF
- Result
 - Items for 2005 Persistency Framework work plan

Workshop Outline



- Wednesday - Workshop Summary
- Goal
 - Figure out if it all (or any) of that really happened :-)
 - Summarize work items (together with available / required resources) for the 3D work plan

Starting Point for a Service Architecture?



Site Contacts



- **Established contact to several Tier 1 and Tier 2 sites**
 - Tier 1: ASCC, BNL, CERN, CNAF, FNAL, GridKa, IN2P3, RAL
 - Tier 2: ANL, U Chicago
- **Visited FNAL and BNL around HEPiX**
 - Very useful discussions with experiment developers and database service teams there
- **Regular meetings have started**
 - (Almost) weekly phone meeting back-to-back for
 - Requirement WG
 - Service definition WG
 - Current time (Thursday 16-18) is inconvenient for Taiwan
- **All above sites have expressed interest in project participation**
 - BNL has started to setup Oracle setup
 - Most sites have allocated and installed h/w for participation in the 3D test bed
 - U Chicago agreed to act as Tier 2 in the testbed

Requirement Gathering



- All participating experiments have signed up their representatives
 - Rather 3 than 2 names
 - Tough job, as survey inside experiment crosses many boundaries!
- Started with a simple spreadsheet capturing the various applications
 - Grouped by application
 - One line per replication and Tier (distribution step)
 - Two main patterns
 - Experiment data: data fan-out - $T_0 \rightarrow T_n$
 - Grid services: data consolidation - $T_n \rightarrow T_0$
 - Exceptions need to be documented
- Aim to collect complete set of requirements for database services
 - Also data which is stored locally never leaves a site
 - Needed to properly size the h/w at each site
- Most applications don't require multi-master replication
 - Good news - document any exceptions

