

Tier0 DB Service Upgrade

Maria Girone, IT-DM

WLCG Workshop – Database Track

April 25th 2008



- Migration of services to 2008 h/w
- Readiness for CCRC'08 May run and LHC start-up
- Review of 2008 requests
- CCRC'08 Post-Mortem
- Summary

- New hardware arrived in Jan (RAC5 & 6)
 - 34 Dual-CPU quad-core 2950 DELL servers, 16GB memory
 - Intel **5400-series “Harpertown”**; **2.33GHz clock**
 - 60 Infortrend storage arrays: 16 SATA disks, 400GB each (more in Luca’s talk)
- Some installation delays due to availability of “rails” – solved in Feb
- Hardware acceptance tests Feb / March
- O/S installed and Oracle RAC s/w (10.2.0.3) in Apr (week 1 and 2)

DM

Current set-up

CERN IT
Department

RAC1



RAC2



RAC3



RAC4



DM

```
// stop all pools,  
for(tp = m.thr  
if(tp->second-  
busyTPools.p  
  
// Reap child pr  
pid_t pid;  
while ((pid = w  
if(!beGraceful)  
// on a SIGINT  
return; }  
  
// now loop wait  
while(busyTPool  
sleep(1); //  
for(unsigned i  
if(busyTPools  
// it's little no  
busyTPools  
  
else
```

New hardware: RAC5 and RAC6

RAC5



RAC6



- 1/2 of Servers + all storage Installed in Critical Power area: battery UPS and diesel backup

Experiment	# servers	# arrays	Data volume
ATLAS offline	4	7	10TB
CMS	4	7	10TB
LHCb	3	5	7.2TB
WLCG	4	5	7.2TB

(1 disk array ~ 6.4 TB raw space
~25% effective space (mirroring and on-disk backups))

- Resource allocation based on experiments' requests (**please check!**)
 - ATLAS requests at <https://twiki.cern.ch/twiki/bin/view/Atlas/DatabaseVolumes>

- Stress tests revealed a high rate of controller failures (~10% of disk arrays affected)
- Final migration performed in Apr (week 2) using Data Guard (minimal down-time of two hours)
 - ✓ WLCG, LHCb, CMS clusters on RAC6
 - ✓ ATLAS cluster on RAC5
 - ✓ Two downstream capture servers (ATLAS and LHCb) on RAC5
- Retained “old” hardware (on RAC2 and RAC3) as Data Guard “stand-by” to the new one
- Running Oracle s/w version 10.2.0.3 (64 bit)
 - Deployed on the validation RACs since Dec

- Oracle have recently delivered **10.2.0.4**
- This includes numerous fixes:
 - Numerous library cache bugs
 - Several streams related bugs
 - All 10.2.0.3 patches (we had to apply a few)
- Can still expect some surprises – extensive testing is important, but unlikely to catch all issues without realistic production load
- **10.2.0.4** is now deployed on all experiment **validation** RACs
- Oracle **April CPU** also published (included already in 10.2.0.4) and already tested by our team on 10.2.0.3 and available to the experiments for further testing
- Need extensive tests of this major Oracle release
 - Possibly **prior** to 2008 data taking

- **online-offline** streams replication for conditions for ATLAS, CMS and LHCb
- Experiments online clusters are ready
 - LHCb still waiting for the final hardware configuration
- **T0-T1** streams replication for LHCb (conditions and LFC) and ATLAS (conditions)
- More in Eva's talk

- Post-mortem workshop to review CCRC'08 will take place June 12-13 at CERN
- A session on **DB services** was requested this week – clearly needed with increasing emphasis on DB services in May phase!
- Also fully consistent with important (crucial) role of DB services vis-à-vis experiment “**Critical Services**” lists
- Still some concern that some aspects (e.g. ATLAS/LHCb conditions) still not fully tested in May
- A July run of CCRC'08 is not excluded...
- (Not to mention CCRC'09...)

- No change in SLA expected with the current resources
 - 24x7 still “best effort”, streams support during working hours
- Hardware upgrade – very tight schedule but we managed to put it in place prior to CCRC’08
- Oracle 10.2.0.4 – already on validation RACs
- Need joint decision on its possible deployment before summer
- Now working on a “**DB availability and performance page**” (more in Miguel’s demo)
- **DB services are a full and essential component of WLCG!**