

A decorative white crosshair consisting of a vertical line and a horizontal line intersecting in the upper left quadrant of the slide.

# CERN Physics Database Services and Plans

Maria Girone, CERN-IT

[physics-database.support@cern.ch](mailto:physics-database.support@cern.ch)

# Outline

---

- Service Structure and Architecture
  - Validation and Production services
- Service consolidation plans
- Deployment model for 2006
- Resource Constraints
- Conclusions

# Database Services for Physics

---

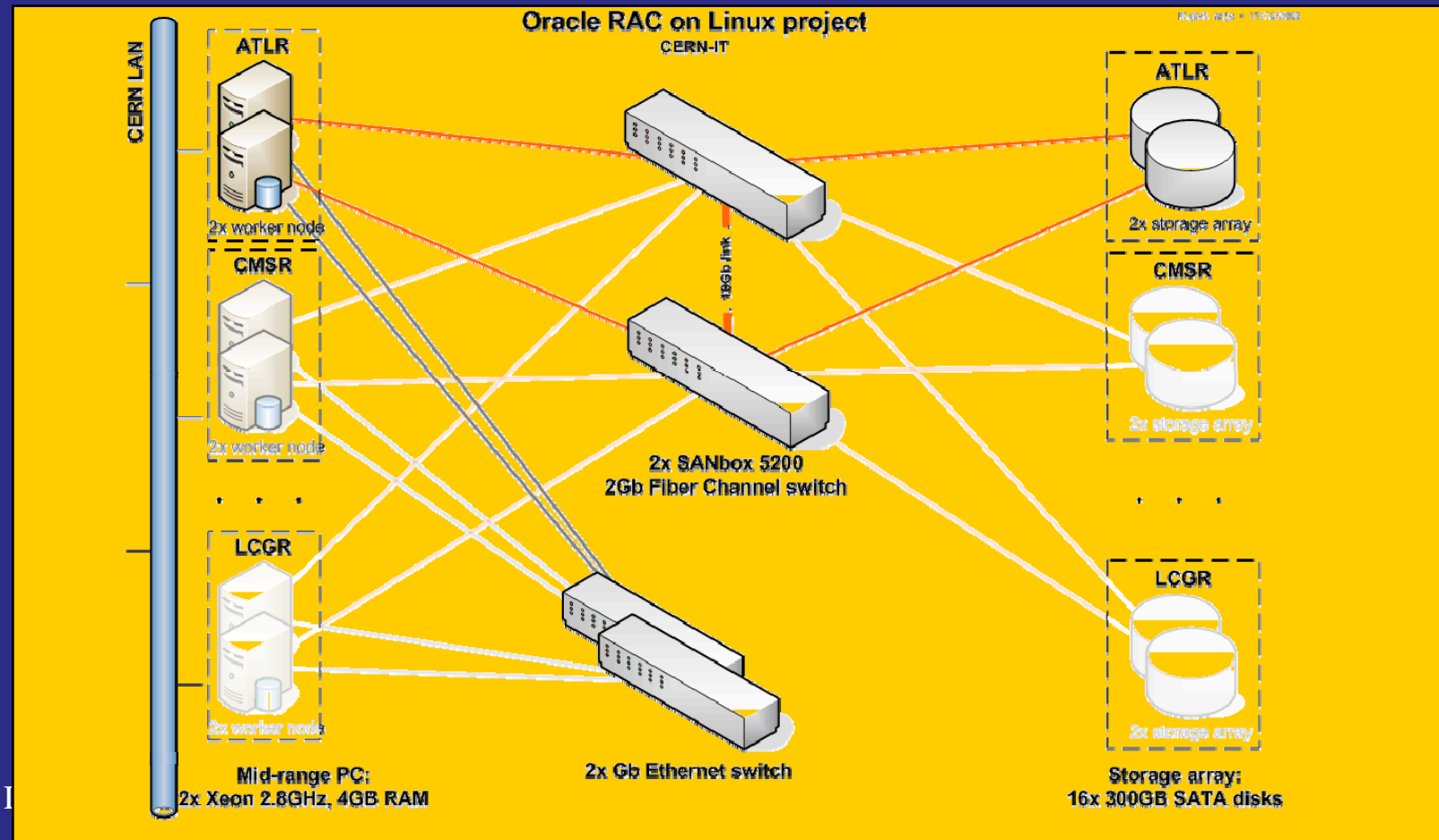
- **Mandate**
  - Coordination of the deployment of physics database applications
  - Administration of the physics databases in co-operation with the experiments or grid deployment teams
  - Consultancy for application design, development and tuning
  - Involvement in 3D project and LCG Service Challenges
- **Provide database services for LHC and non-LHC experiments**
  - applications related to book keeping, file transfer, physics production processing, on-line integration, detector construction and calibration

# Service Levels

- **Development Service (run by IT-DES)**
  - Code development, no large data volumes, limited number of concurrent connections
  - Once **stable**, the application code and schema move to validation
- **Validation Service (for key apps)**
  - Sufficient resources for larger tests and optimisation
  - Allocated together with DBA resources consultancy
    - Needs to be planned in advance
  - Limited time slots of about 3 weeks
- **Production Service**
  - Full production quality service, including backup, monitoring services, on call intervention procedures
  - Monitoring to detect new resource consuming applications or changes in access patterns
- **OS level support provided by IT-FIO**

# Service Architecture

- The Physics Database **Production** and **Validation** services are mainly deployed on **2-node RAC/Linux**, in failover mode



# Validation Service

---

- Based on **two 2-node RACs**
- Reviewed about 10 **key** applications in about 4 months
- It requires a significant effort from both sides, but
  - Sizeable **performance improvements**
  - Better understanding of **resource requirements** achieved
    - In some cases, a reference workload is still missing
  - **Positive feedback** from the experiments
- Adding this service level has been a good idea!
  - Reduces the **risks** in production deployment
  - DBAs have a better **knowledge** of the key applications

# Validation experience

- Validation/Production levels is an iterative process for new application software versions
- Next step: use the results as a part of Service Level Agreements
- We capture snapshots of the query mix and resource consumption of a given application
- Can be compared later to similar snapshots at production level
  - Useful for identifying changes in access pattern or problems

# Production Service

---

- **End of 2005:** phasing out of the old **2-node 9i Sun cluster (PDB)** and most of individual disk servers
- Many **new** applications/instances have been requested in the last 6 months, many in the pipeline
  - Deploying them on **RAC**, after **Development/Validation** cycle
    - **Flexible** architecture to cope with increasing demand
    - **Redundant** architecture for high availability
- We are currently migrating all the LHC experiments and grid applications to RAC
  - Cooperation needed for validation on the new system



# RAC is Production now!

---

- Two RACs for the Validation service
- Three LHC experiments (ATLAS, CMS and LHCb) dedicated RACs
- One RAC for LCG applications (FTS, LFC, etc)
- One RAC for the ATLAS online tests (so far, with time limited allocation)
- One RAC for the service development
- We are happy to collaborate with LCG sites who are interested in deploying RACs in their services, in the context of the 3D project

# Applications moved to RAC already

---

- **ATLAS**
  - ATLAS\_COOL, ATLAS\_Event Tags, ATLAS\_da, ATLAS\_TO
  - ATLAS\_ProdSys, ATLAS\_Muon Cert, ATLAS\_Muon (migrated from PDB)
- **CMS**
  - CMS\_transfermgmt\_SC, CMS\_transfermgmt\_TEST, CMS\_PXL and CMS\_HCL
  - cms\_muon\_endcap (migrated from PDB)
- **LHCb**
  - LHCb\_COOL
  - LHCb\_bookkeeping, LHCb\_ecal, LHCb\_richhpd (migrated from PDB)
- The migration from the PDB cluster is half way through

# Current Requests...

---

- Consolidation of the **Production** service in the RAC architecture
- Experiment dedicated **Validation/Test** services
- **Development** service on ORACLE 10g **Release 2**
- **3D** service
- **Online** Database test at the computer center for ATLAS on RAC
- Possible consolidation of the service for COMPASS into RAC with about **10 TB** of data, including 2006 running and full data re-processing
  - valuable experience in handling **large** volumes data

## ... And Issues

- **Database Structure Conventions:** naming conventions, roles, profiles for achieving better organization and **smoother transitions** across service levels (see M. Anjo talk)
- **Storage:** studying different scenarios for the storage layout in order to increase the system **I/O performance** and make best use of the available capacity (see L. Canali talk)
- **Backups:** need to have **scalable** and regularly **validated** recovery procedures from backups with **minimal** recovery latency (see J. Wojcieszuk talk)
- **Security:** how to get securely connected to a database in a grid environment (see K. Zajackowski talk)
- **Monitoring:** provide **database** and **application** level monitoring information to be used by both developers and DBAs (see R. Chytracsek talk)
- **High Availability:** Planned interventions for applying OS and ORACLE release upgrades and security patches are our main reason of service **downtime**. This results in a significant **impact** at application level (see D. Duellmann talk)

# Hardware evolution for 2006

- Ramping up of the hardware resources in 2006-2008

Current State							
ALICE	ATLAS	CMS	LHCb	Grid	3D	Non-LHC	Validation
-	2-node offline	2-node	2-node	2-node	-	-	2x2-node
	2-node online test						
Proposed structure in 2006							
2-node	n-node or nx2-node	n-node or nx2-node	n-node or nx2-node	n-node or nx2-node	2-node	2-node (PDB replacement)	
	2-node valid/test	2-node valid/test	2-node valid/test	2-node pilot		Compass??	
	Online?						

# Online Databases

- Several experiments foresee their **online** databases to be located at the experiment sites
- Currently we offer database services **only** in the computer center
  - Resource **constraints** don't allow us to take up this significant additional task
- We would like to avoid possible software and hardware **divergence** and help the experiments to setup their own **online** services on a best effort basis
  - Keep the experiment informed about hardware choices
  - Database s/w kits and **consultancy** are provided also for **online** applications
  - We can help in organizing **OCP** training for experiments **DBAs**

# Resource Constraints

---

- Still very busy period
  - Many experiment applications still **ramping up**
  - **Consolidation** of DB service for physics still ongoing
- Can not pick up any larger **additional** tasks at the moment
- Recent hire:
  - Expert Oracle DBA with significant application optimisation experience
- Another way of **freeing** resources is retiring services which are not used anymore

# Conclusions

---

- In the last 6 months we made a **big step** towards a flexible infrastructure for LHC
- New **validation service** introduced for the **key** applications
  - Better understanding of resource requirements and applications deployed. Needs considerable effort on both sides
- **RAC/Linux** is now in production
  - The **consolidation** phase is still going on. Additional hardware resources expected to be available from November 2005
- Can now extend the database service but need some planning
- We think we are progressing well to phase the LHC start-up ☺





Questions?