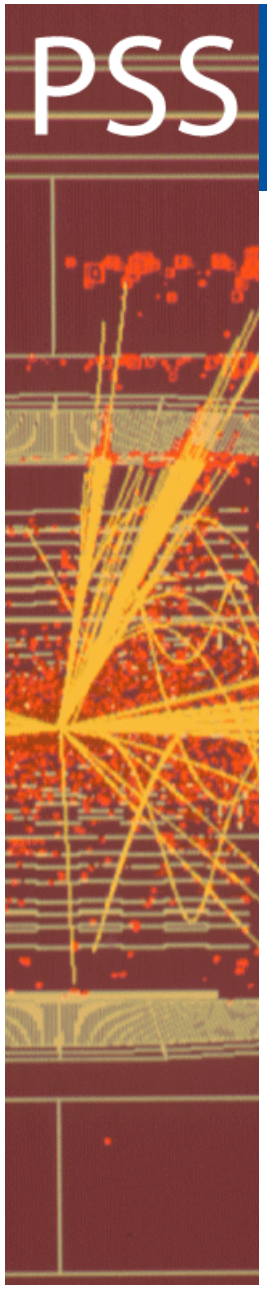


PSS Physics Services Support

CERN IT Department

CERN ORACLE database for physics status

Maria Girone, CERN IT-PSS

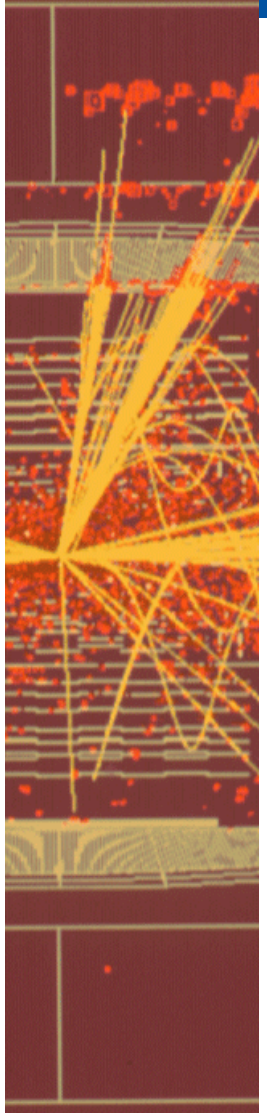
Outline

- Service operational aspects
- Hardware set-up in 2007
- Replication set-up
- Test plan
- Backup and security policies

Introduction

- Mandate: offer a highly available and scalable database service to the LHC experiments and grid deployment teams
  - The Tier0 3D service responsibilities moved to the service team
- Evolution of the database set-up
  - Database software -> fully based on Real Application Cluster 10g (validated Oracle server 10.2.0.3)
    - Now moving it to production. Should be considered by the sites as well.
  - Operating system -> Linux RHES3, moving to RHES4
- 15 RACs clusters in 2007 - same team
  - doubling resources and services to prepare for the LHC start-up

Main Operational Aspects



Service Size

- 110 mid-range servers and 110 disk arrays (~1100 disks)
- In other words: 220 CPUs, 440GB of RAM, 300 TB of raw disk space

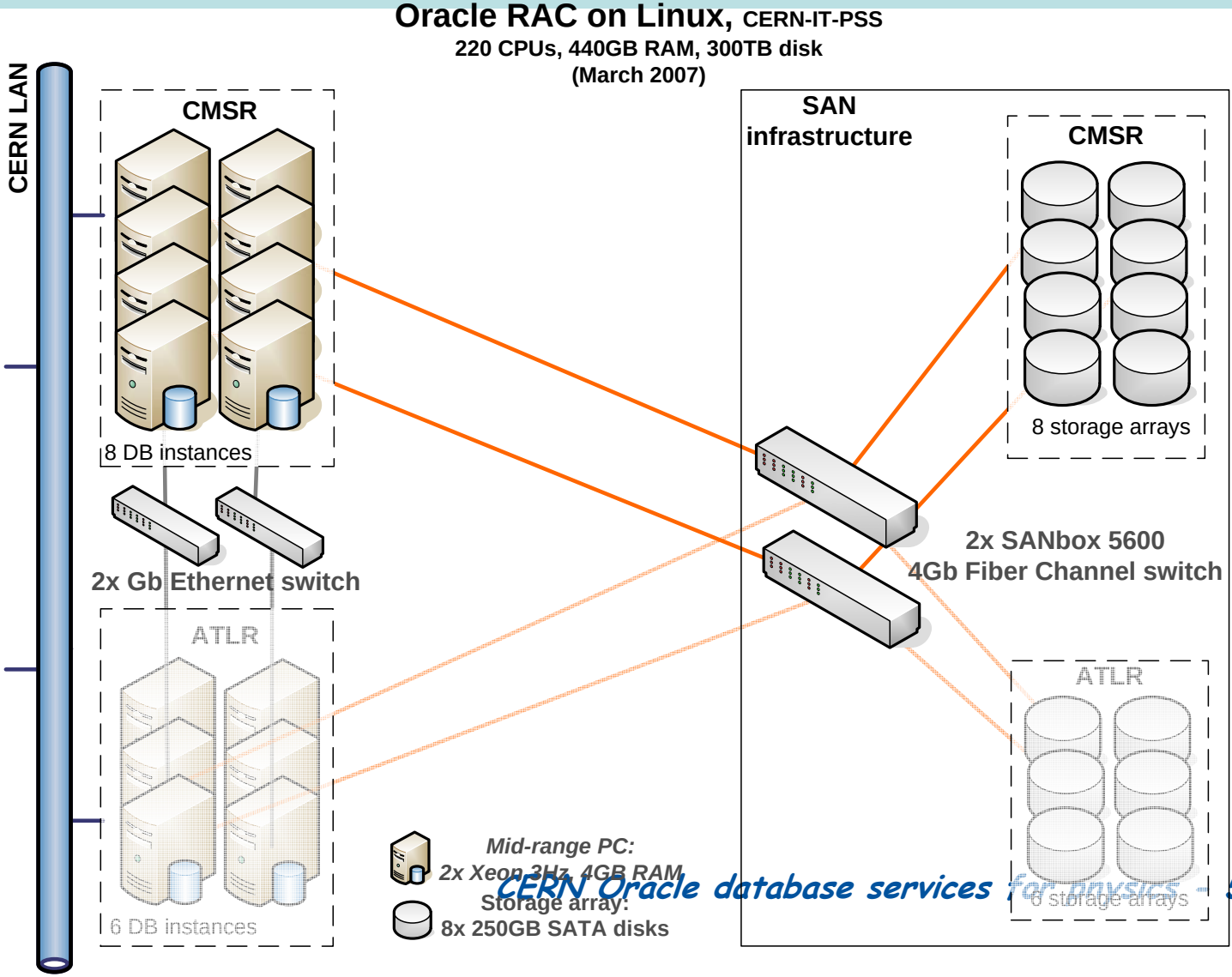
Service Procedures

- Proposing a piquet team for 24x7 coverage of the production service
  - 6 DBAs (and 5 developers for the on-call procedures)
- Backups on tape and on disk
- New Monitoring for ASM and Oracle Service availability introduced (see Dawid's talk)
- Replication procedures and monitorning for 3D being included

Maria Girone
CERN, IT-PSS

Service Architecture

HA DB requires redundant h/w : db server, storage array, ethernet networks, fiber Channel networks (SAN), redundant power supplies and UPS



Maria Girone
CERN, IT-PSS

Hardware Evolution in 2007

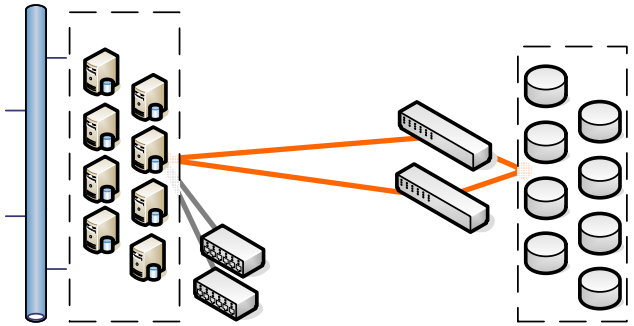
	Q1 2006			Q1 2007		
Experiment/ activity	**Production**	**Integration/ Pilot/Test**	**On-line test**	**Production**	**Integration/ Pilot/Test**	**On-line test**
ALICE	On PDB	-	-	On PDB	-	1 x 2 nodes (PVSS tests)
ATLAS	1 x 4 nodes	1 x 2 nodes	1 x 2 nodes	1 x 6 nodes	1 x 2 nodes	1 x 6 nodes
CMS	1 x 4 nodes	1 x 2 nodes	-	1 x 8 nodes	1 x 2 nodes	1 x 6 nodes (PVSS tests)
LHCB	1 x 4 nodes	1 x 2 nodes	-	1 x 6 nodes	1 x 2 nodes	1 x 6 nodes (PVSS tests)
WLCG	1 x 4 nodes	1 x 2 nodes	-	1 x 8 nodes	2 x 2 nodes	-
PDB (alice, harp, totem)	1 x 2 nodes	-	-	1 x 4 nodes	-	-
COMPASS	1 x 4 nodes	-	-	1 x 4 nodes	-	-
3D	-	1 x 2 nodes	-	3 nodes (downstream capture)	1 x 2 nodes	-

Maria Girone
CERN, IT-PSS

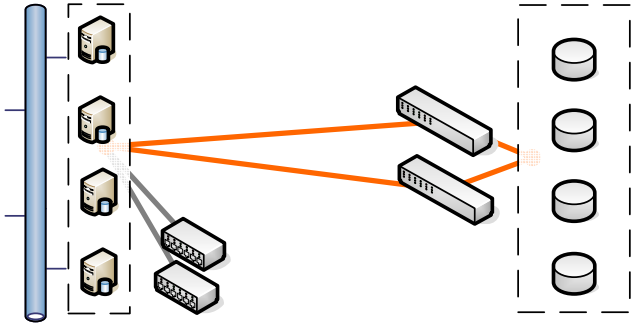
Production and Validation set-up

Oracle RAC on Linux, CERN-IT-PSS
220 CPUs, 440GB RAM, 300TB disk
(March 2007)

PRODUCTION

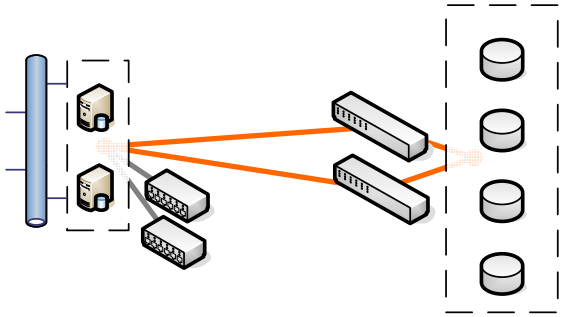


ATLR	ATLAS (6 instances)
CMSR	CMS (8 instances)
LCGR	Grid/WLCG (8 instances)
LHCBR	LHCb (6 instances)



COMPR	Compass (4 instances)
PDBR	Other physics apps (4 instances)

PRE-PRODUCTION

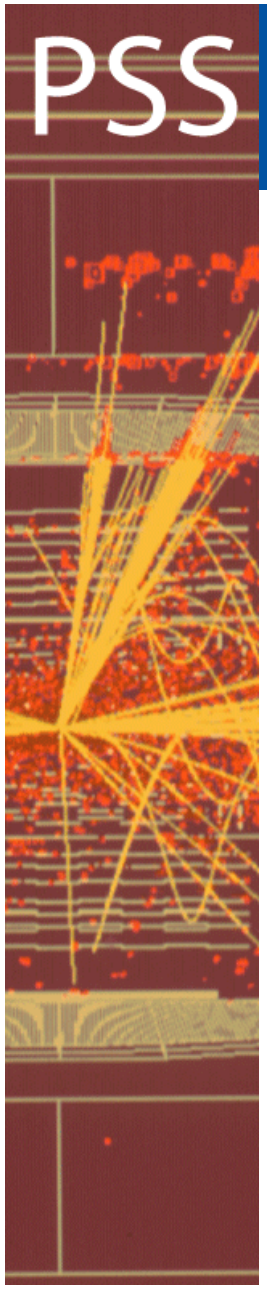


ATONR	ATLAS online (6 instances)
D3R	3D project (2 instances)
INTR	ATLAS validation
INT2R	CMS validation
INT3R	Grid/WLCG validation
INT4R	LHCb validation
INT5R	Other apps validation (4 instances)
INT6R	Grid/WLCG pilot
TEST1	Test RAC / 64bit tests (8 instances)

Temporary allocation:

INT7R	Alice PVSS tests
CMSONR	CMS PVSS tests (6 instances)
LHCBONR	LHCb PVSS tests (6 instances)

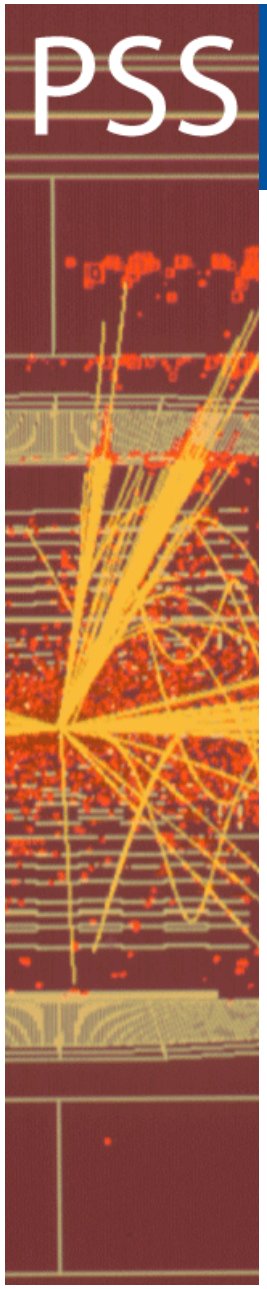
Maria Girone
CERN, IT-PSS



Test Plan for 2007

- Multi-core and large memory servers are expected to significantly increase the database performance
- Medium term
  - Planning a systematic test in Q1 with dual-core and quad-core based servers to prepare for upcoming service extension towards the end of 2007
  - Need to understand additional I/O and RAM requirements for balanced system (and the price)
  - Natural point to introduce 64-bit Oracle servers
- Short term
  - Next acquisition will be based on dual-core and expected by Q2/Q3
  - Gathering experiments/WLCG service requirements now

Maria Girone
CERN, IT-PSS

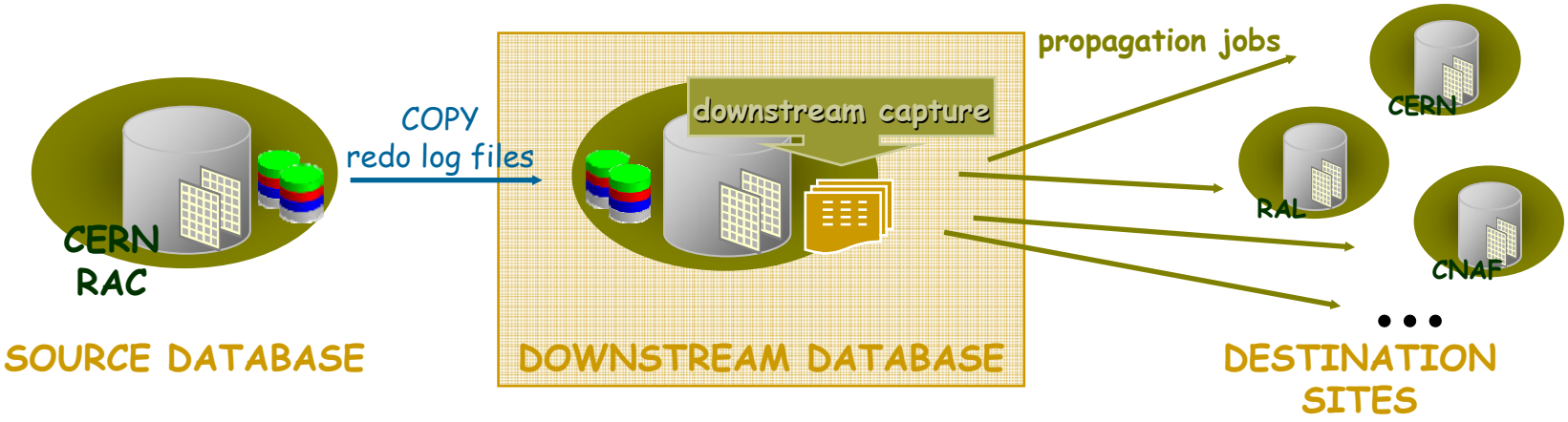


Current Replication Set-up

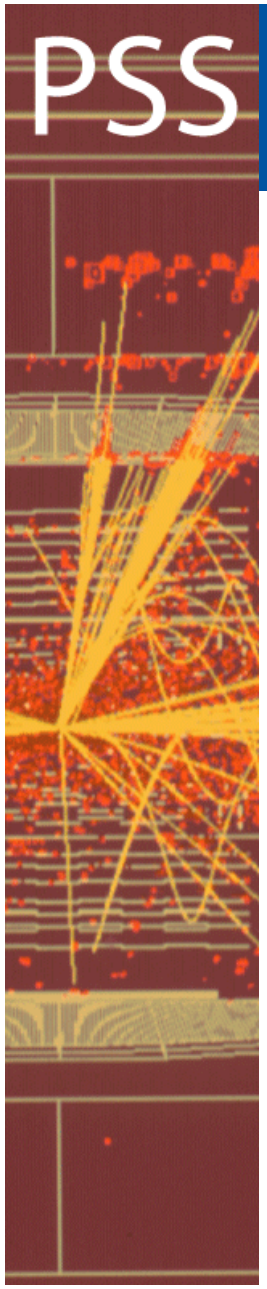
- ATLAS replication setup
  - online -> offline -> Tier1's
  - successful tests with 4 Tier1 active sites
    - CNAF, IN2P3, RAL, GridKA
  - Tier1 sites being added:
    - BNL, ASGC and first phase 2 sites Triumf, SARA
- LHCb replication setup
  - online (pit) -> offline -> Tier1's
  - successful tests with 3 Tier1 active sites
    - IN2P3, RAL, GridKA
  - Tier1 sites being added:
    - CNAF
  - LFC replication setup
    - source database: LHCb RAC
    - first Tier1 destination site: CNAF

DOWNSTREAM Capture

- The set-up has been implemented and will be used in production
- The source databases will be deployed on the experiments production RACs



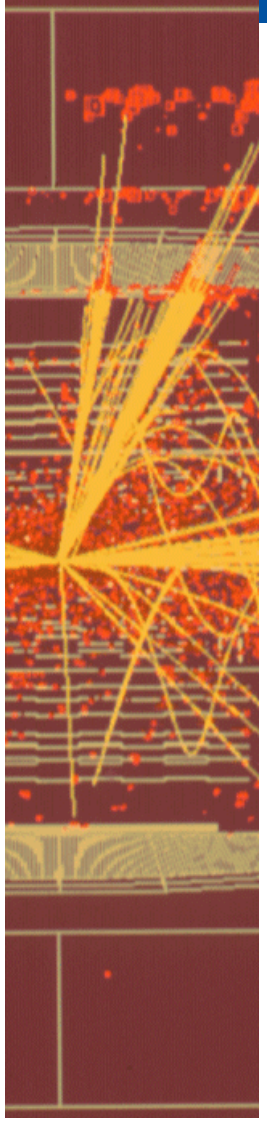
https://twiki.cern.ch/twiki/bin/view/PSSGroup/DownstreamDatabaseConfiguration



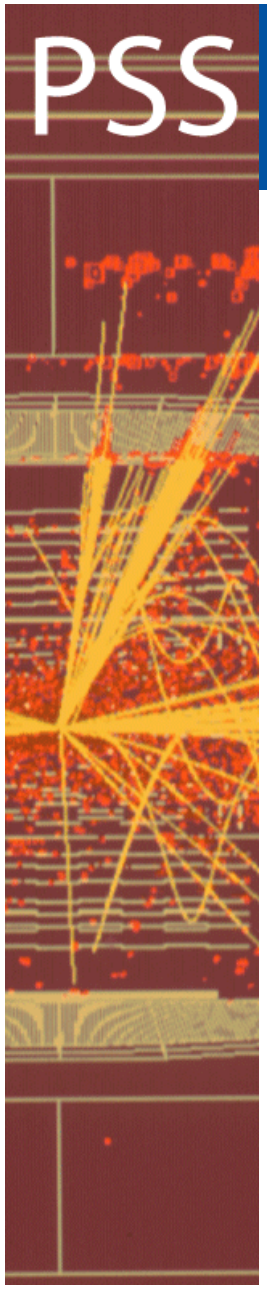
Backup and Security Proposals

Maria Girone
CERN - IT

Backup and Recovery Proposal



- Backup based on ORACLE RMAN 10g
  - Risky changes should be tried on the validation set-up first
  - Backup on tape: uses TSM
    - retention set to 31 days
      - A full backup is systematically performed and kept before any ORACLE software upgrade
      - **Full** - every 2 weeks
      - **Incremental** (differential or cumulative) – daily
      - **Archive logs** - every 30 minutes
  - Backup on disk: all data files into the flash recovery area
    - retention set to 2 days
      - **Full** - at database creation
      - **Incremental** – daily

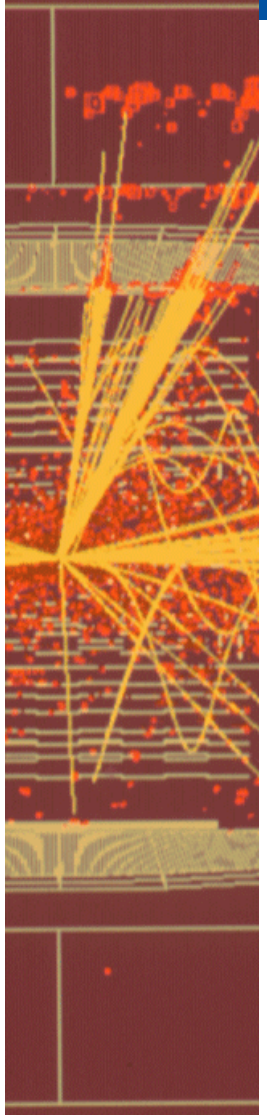


Security and s/w Upgrades

CERN IT Department

- Critical Patch Upgrade (CPU) are applied asap, typically within two weeks from the publishing date
  - validation period of typically one week
- Oracle software upgrades are typically performed once or twice per year
  - version installed on the validation RAC and tested by the application owners and Tier 1 sites for one month
- Minor upgrades discussed in LCG-SCM and 3D
- Major version updates deferred to the MB
- Oracle patches are only made for recent versions and therefore it is essential to update accordingly

Summary



- Security and backup proposals have been presented to the experiments/WLCG community
  - No objection received so far
  - Can the Tier1 sites agree on the proposal?
  - Will present it to the GDB and LCG MB
- We have included the Tier0 3D responsibilities within the service
- On schedule with the hardware expansion in Q1
- Preparing transition to multi-core servers and 64 bit RH & Oracle
- We need your input for any larger access patters and volume changes now