CNAF Site Review

Barbara Martelli

Outline

- CNAF database service status (non-CASTOR DBs)
  - CASTOR databases service status details
- ATLAS LFC failover solution just implemented:
  - DataGuard between CNAF and INFN – Roma1

HW Resources for DB Services @ CNAF

- 30 servers, 12 of them dual-core
- 10 clusters
- 42 TB of raw disk space
  - 10 TB of Fibre Channel disks 2 Gbit on EMC CX3-80, RAID1 LUNs, ASM Disk Groups with external redundancy
  - 10 TB of Fibre Channel disks 4 Gbit on EMC CX3-80, RAID1 LUNs , ASM Disk Groups with external redundancy
  - 20 TB of SATA disks on EMC CX3-80, RAID5 LUNs, ASM Disk Groups with external redundancy
  - 2 TB of OCFS2 file system on FastT900 (only for OEM DB back-end and test DBs ➡ **to be dismissed**
- New server procurement ongoing:
  - 10 quad-core servers. Tender still open, don't know the brand yet
- No storage acquisition for DBs foreseen this year
  - Very important to know as precisely as possible experiments needs for 2009 and 2010
- 2 DBAs

Distributed DBs S/W

- All clusters except LHCb
  - RedHat Enterprise 5.1.2-21
  - EMC PowerPath 5.1 (path failover/load balancing, device persistency)
  - Oracle 10.2.0.4 plus January '09 CPU patch
- LHCb RAC
  - RedHat Enterprise 4 U5
  - EMC PowerPath 5
  - Oracle 10.2.0.3 plus January '09 CPU patch

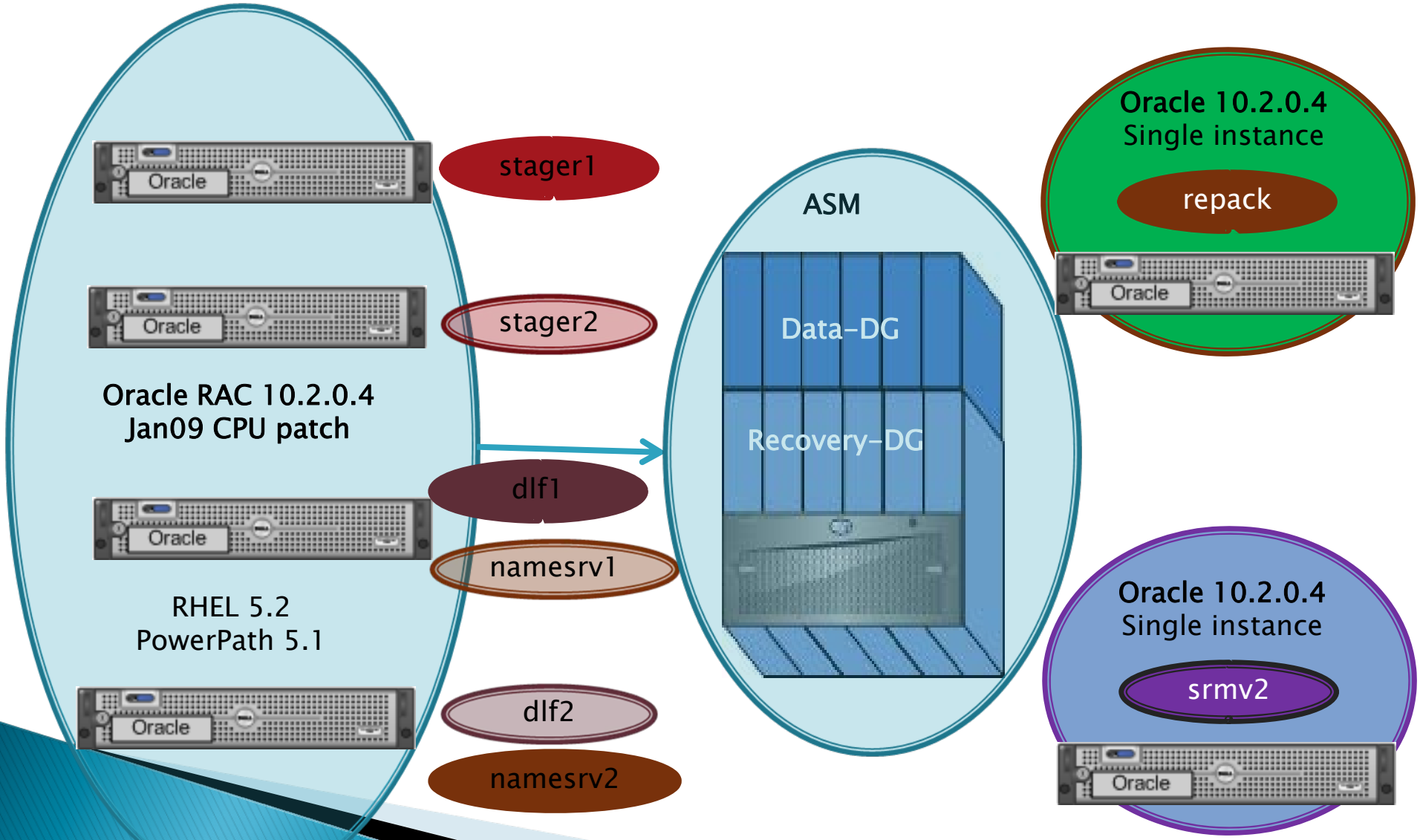
- Storage: 2 TB of raw disk space on EMC Clariion CX3-80 for each RAC
  - ASM with 2 disk groups (DATA, RECOVERY)

CASTOR S/W

- 6 Oracle servers:
  - RedHat Enterprise 5.1.2-21
  - EMC PowerPath 5.1 (path failover/load balancing, device persistency)
- Storage: 2 TB of raw disk space on EMC Clariion CX3-80
  - ASM with 2 disk groups (DATA, RECOVERY) shared among DLF, Stager, Nameserver
  - Local storage for SRM and Repack instances
- Oracle 10.2.0.4 plus October '08 CPU patch
- Castor 2.1.7-17
- SRM 1.3-28

Oracle Deployment for CASTOR

For each service, TAF policy is basic, one node active, one available, the remaining two not used



Distributed DBs Storage Usage

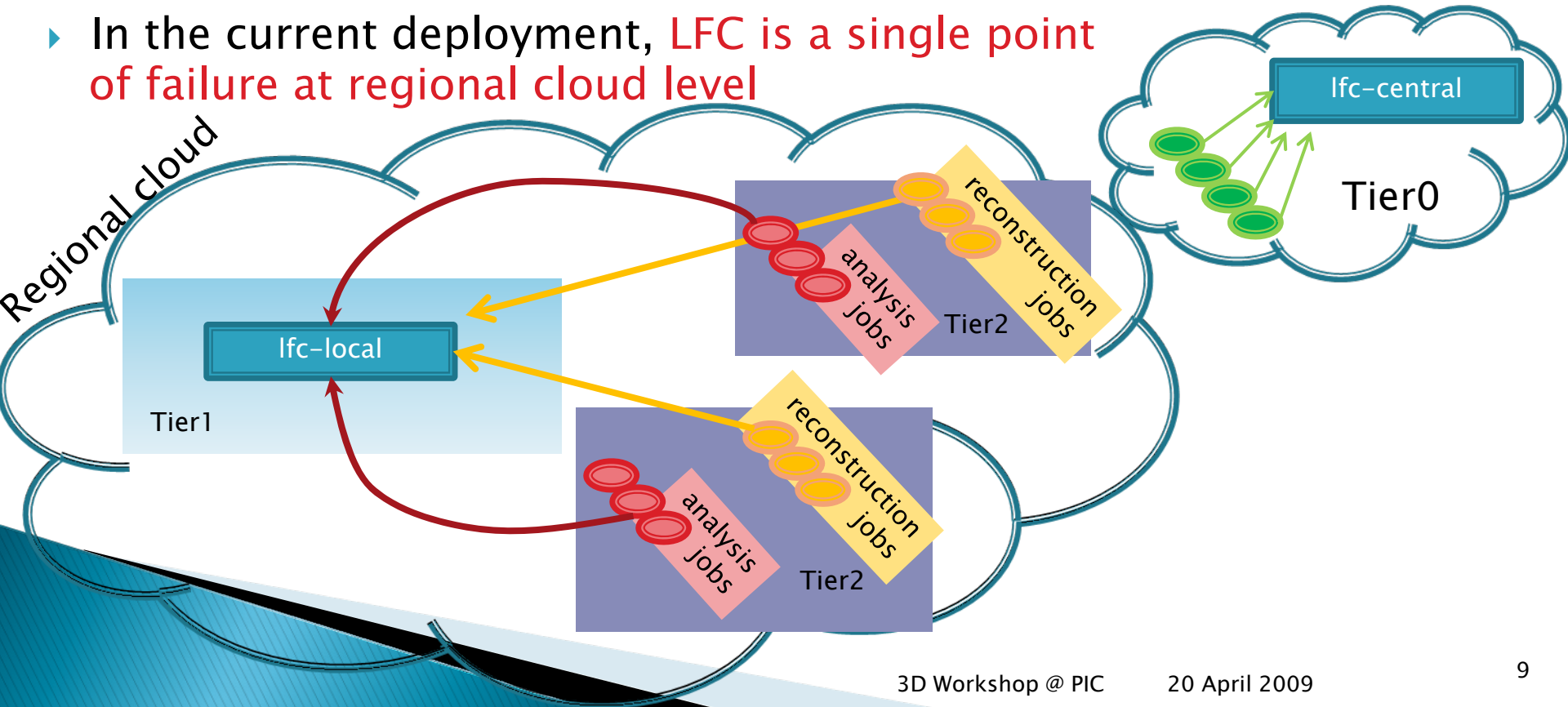
DB Service	Data /Flash Recovery Area	Redo Log activity per day	Active sessions
ATLAS Conditions	139 GB/ 286 GB	4 GB	0-1000
LHCb Conditions	7 GB/ 8 GB	1.2 GB	0-200
ATLAS LFC	9 GB/ 62 GB	300 MB	20
LHCb LFC	10 GB/ 33 GB	800 MB	20
FTS	72 GB/ 312 GB	900 MB	60

CASTOR Storage Usage

DB Service	Data /Flash Recovery Area	Redo Log activity per day	Active sessions
Stager	9 GB/301 GB	48 GB	170
Nameserver	3 GB/8 GB	512 MB	45
DLF	98/113 GB	6 GB	20
SRM	4 GB/9 GB	512 MB	130
Repack	2GB/15GB	/	/

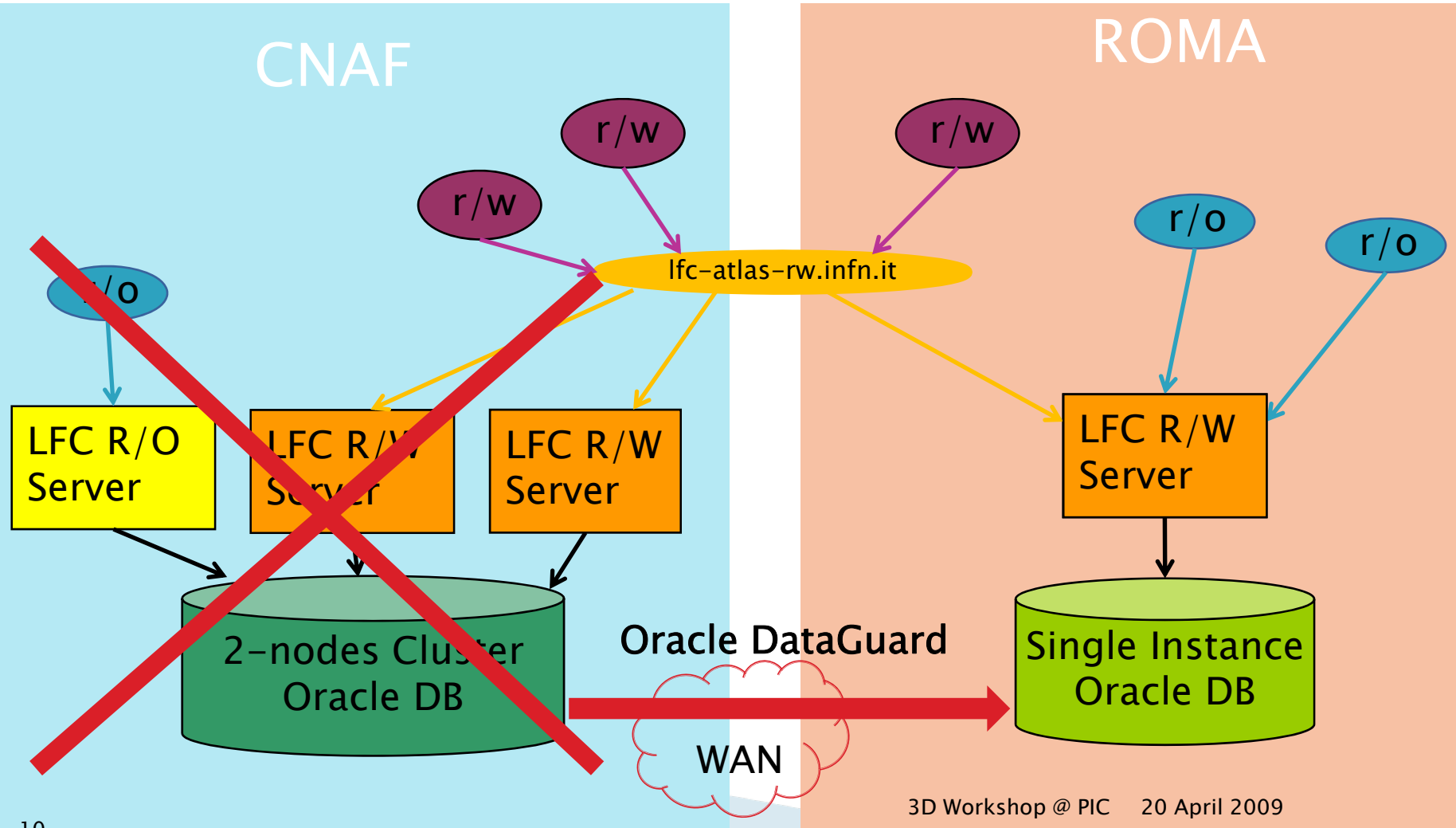
ATLAS LFC Failover to Roma

- ATLAS uses the LFC as both central and local catalogue:
  - **Central LFC at T0**: stores entries referring to production data
  - **Local LFC at each T1**: store entries referring to raw data, ESD, AOD, TAGs delivered to that *regional cloud*
- Therefore a local LFC down implies a total interruption of all reprocessing and analysis operations in that whole regional cloud
- In the current deployment, LFC is a single point of failure at regional cloud level



Deployment Outlook

Oracle DataGuard logical stand-by, maximum performancem, fast start failover



Future Work

- Move SRM and Repack instances to the CASTOR 4-nodes RAC
- Move the OEM DB backend to a new 2-nodes cluster on the EMC storage
- Upgrade the LHCb RAC (Conditions and LFC)
  - RHEL 5.3
  - Oracle 10.2.0.4
- Install the Nagios script in order to implement the ATLAS LFC automatic failover to Roma also at the application layer

Questions?