

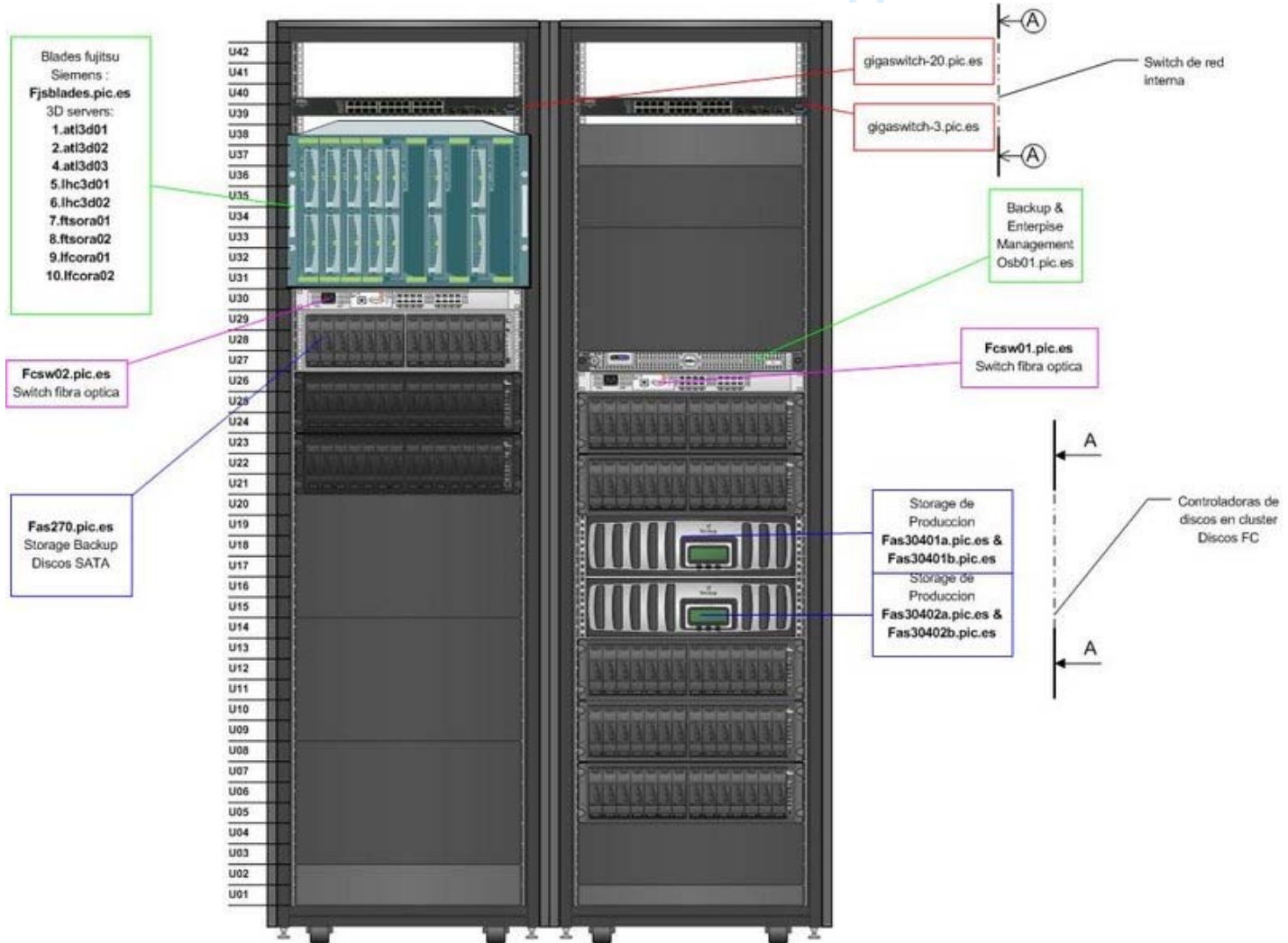
PIC
port d'informació
científica

Databases services at PIC: review and plans

Luis Diaz (PIC)#

- Global view
- Databases hosted
- Structure, general information
 - Servers
 - Redundancy
 - Storage
- Detail of each cluster
- Plans

Global view



- Oracle DB

- ATLAS
- LHCb/LFC
- LFC-ATLAS
- FTS

- All this DB are: 10G R2 (10.2.0.4), 64 bits -- RACs with ASM

- Mysql DB

- LHCbdb (dirac). Pic manage the following subcategories
 - LHCb DIRAC Web site
 - LHCb DIRAC Monitoring, Accounting and Logging system

Streams

Yes

Yes

No (see plans)#

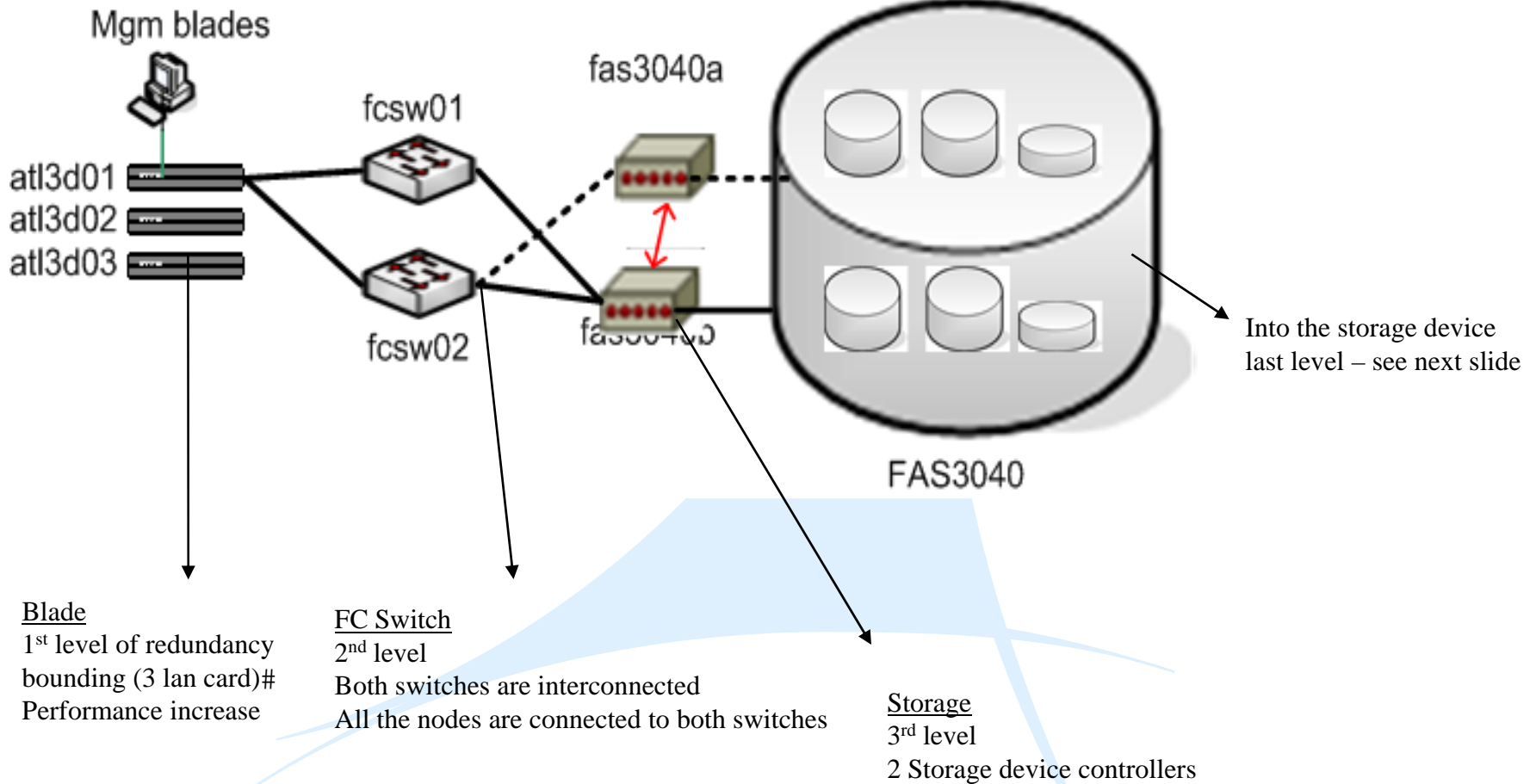
No

- Servers

- All the nodes hosting oracle databases are identical:
 - Linux RedHat 5.2 64 bits
 - Fujitsu siemens: Primergy Advanced Blade Ecosystem BX600
 - 8 G RAM
 - QuadeCore intel CPU 1,6GHz
- DIRAC mysql
 - Scientific Linux CERN SLC release 4.6 (Beryllium)#
 - Dell 1850
 - QuadeCore Intel(R) Xeon(TM) CPU 3.20GHz
 - 4 G RAM
- Repartition:
 - 3 blades for 3D Atlas
 - 2 blades for 3D de LHCb
 - 1 blade for DIRAC
 - 2 blades for FTS
 - 2 blades for LFC

STRUCTURE/GENERAL INFO

- Redundancy (bounding, multipath, rac)#



- Storage

- Provider: NETAPP

- 1st : Production device

- 5 array of 14 discs, 300 Go each one
- 3 spare disks, 2 parity disks and 2 double parity disks
- Total space for data: 12,4Tb

- 2nd : Backup device

- 1 array of 14 disks, 750 G SATA II
- 1 array de 7 disks, 750 G DATA II
- 2 spare disks, 2 parity disks and 2 double parity disks
- Total space for data: 9Tb

- Storage

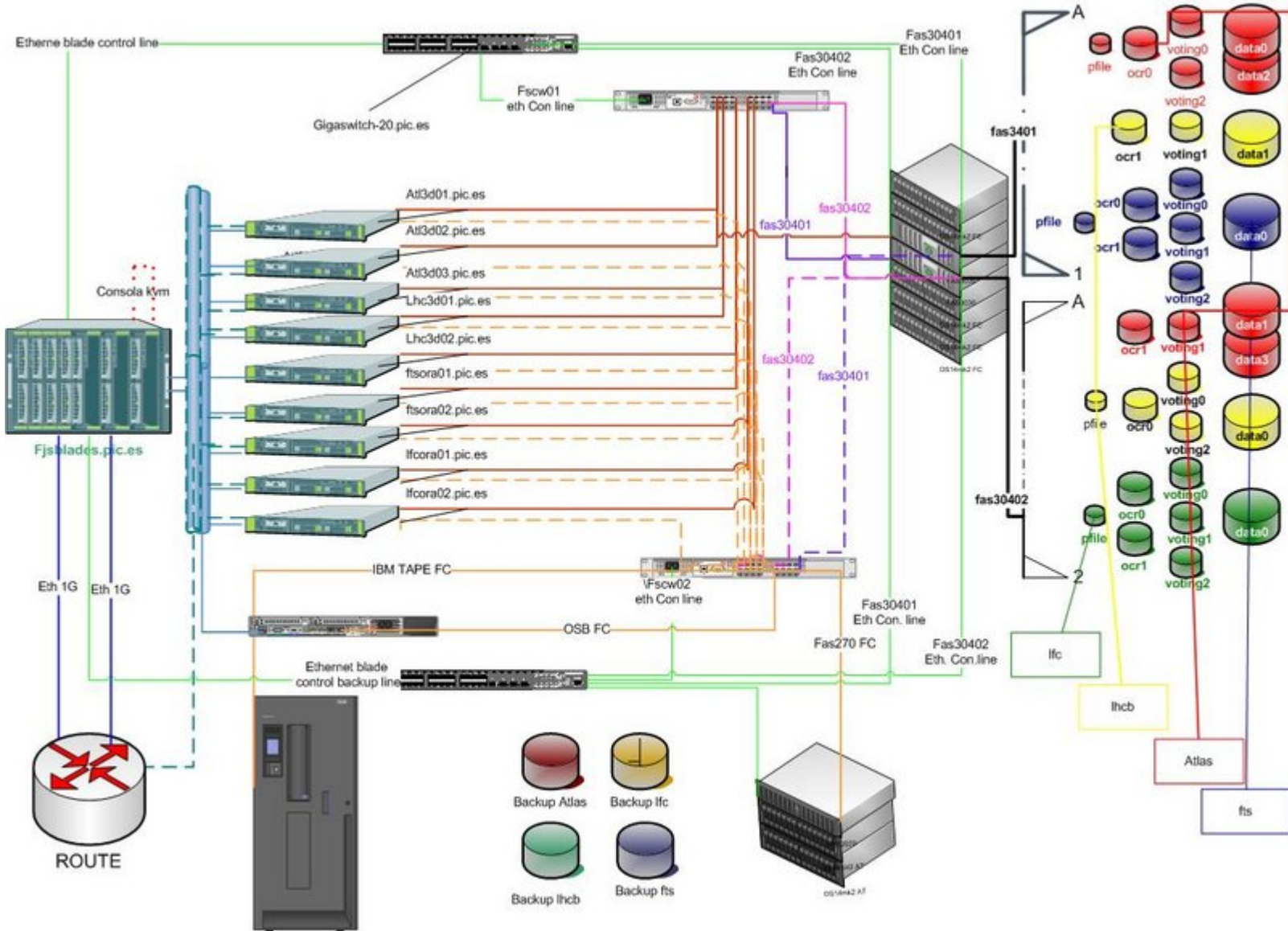
- Performances test / ORION

- `./orion_linux_em64t -run advanced -write 0 -matrix basic -duration 120 -testname mytest -num_disks 50 -cache_size=10000`

- Small IO size: 8 KB
- Large IO size: 1024 KB
- IO Types: Small Random IOs, Large Random IOs
- Simulated Array Type: CONCAT
- Write: 0%

- Cache Size: 10000 MB
- Duration for each Data Point: 120 seconds
- Small Columns:, 0
- Large Columns:, 0, 1, 2, 3, 4, 5, 6, 7, 8, 16, 24, 32, 40, 48, 56, 64, 72, 80, 88, 96
- Total Data Points: 51
- Name: /dev/dm-8 Size: 1099529453568
- Name: /dev/dm-3 Size: 1099529453568
- Name: /dev/dm-9 Size: 1099529453568
- Name: /dev/dm-4 Size: 1099529453568
- 4 FILEs found.
- Maximum Large **MBPS=384.62** @ Small=0 and Large=96
- Maximum Small **IOPS=17537** @ Small=228 and Large=0
- Minimum Small **Latency=1.20** @ Small=6 and Large=0

STRUCTURE/GENERAL INFO



- ATLAS
 - Service description
 - Metadata of ATLAS experiment cond DB.
 - Access requirements
 - All opened
- LFC-LHCb
 - Service description
 - Metadata of LHCb experiment cond DB.
 - Access requirements (not FULL opened like ATLAS)#
 - Opened from PIC to all other TIER1
 - Open for CERN: OEM monitoring

- LFC-ATLAS
 - Service description
 - Catalogue/index of files used by ATLAS experiment.
 - Access requirements
 - Full opened
- FTS
 - Service description
 - Transfert Log, history, service given to PIC's VO for LHC
 - Fts monitoring (web interface)#
 - Access requirements
 - Only open to 6 internal frontend servers

- ALL DB

- Disponibility requirement
 - 24/24 7/7
- Scheduled jobs
 - Backup rman jobs
- Monitoring
 - Nagios checks:
 - RAC check (include instance, listener and vip checks) , Block corrupt check and streams check
 - Oem
 - All the standard checks
 - Ganglia
 - DB Growing sum(bytes) from dba_segment where owner like '%ATLAS/LHC%'
 - sessions (and all the standard ganglia checks:
 - » Memory..cpu load, red load, etc...)#
 - Cern Monitor (<https://oms3d.cern.ch:1159/dstrmon/>)#

- **HardWare plans**

- During the period 2009/2010, we don't plan to upgrade any cluster.
- We will wait for the 1st year taking data from accelerator. If, as expected, the cluster give good results, no change will be done ; but at the same time, we will check for CPU load with ATLAS cluster, and add CPU if we see that it is needed.

- **Configuration plans**

- For security reason, we'll change ATLAS listener port from 1521 to another.
- Implement the backup/restore procedure with NETAPP storage device, as soon as it works.
- LFC-ATLAS streams: We plan to test streams internaly. Then we'll be able to decide if we'll implement it in production.

- Miscellaneous

- Periodical tasks:

- Restore tests: From backup file, we move all files in other cluster and from scratch we recover...

- The procedure to move files from ASM volume is:
DBMS_FILE_TRANSFER()#

- Performance report: Using data from OEM, Cern Monitor, AWR and Ganglia.

- Stress tests

- With ATLAS mainly (see Xavi Espinal talk)#

- Oracle VM

- We have one Oracle VM installed with FTS Preproduction on it.
 - We are testing this solution: the main objective is to reduce the electrical cost due to many servers connected.

- THANKS FOR YOUR ATTENTION