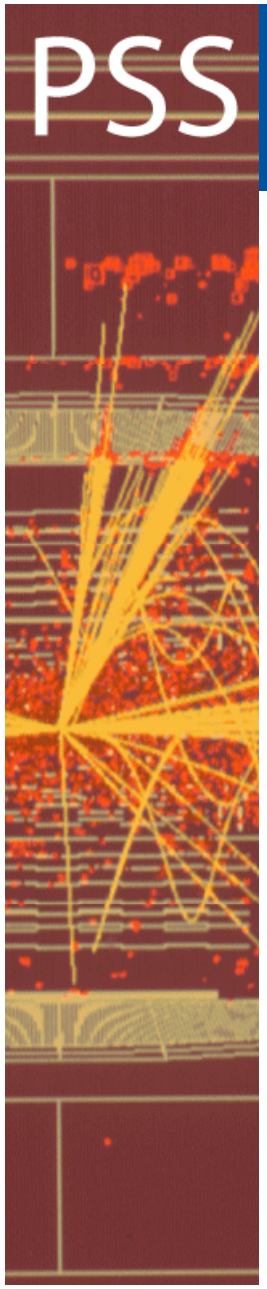


PSS Physics Services Support

CERN IT Department

Tier0 database extensions and multi-core/64 bit studies

Maria Girone, CERN IT-PSS

LCG 3D Database Workshop, 12-13 June 2007

CERN - IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

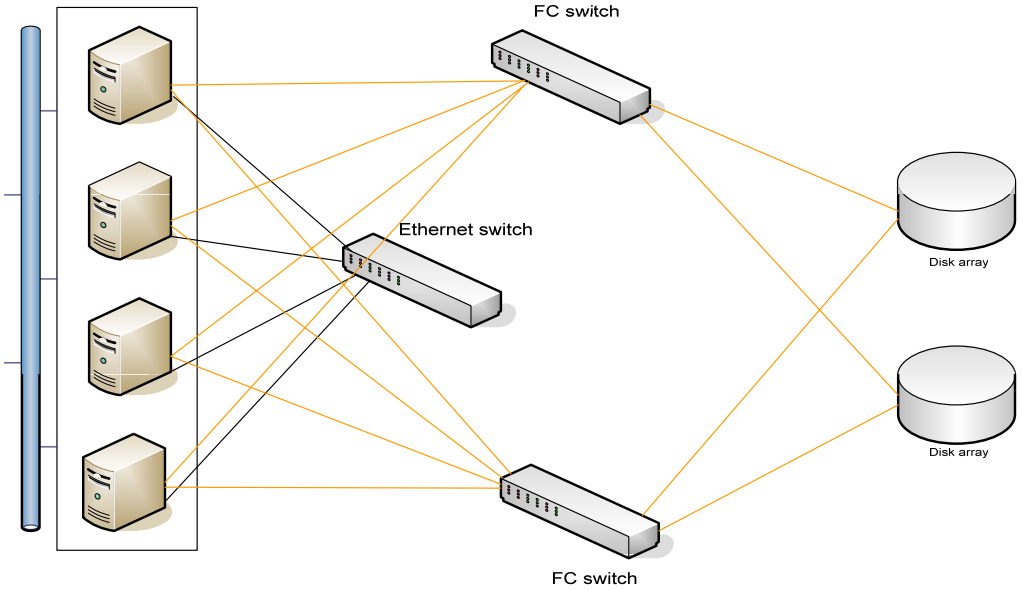
Outline

- Current set-up
- Quad-core/64 bit tests
- Resource allocation for 2008

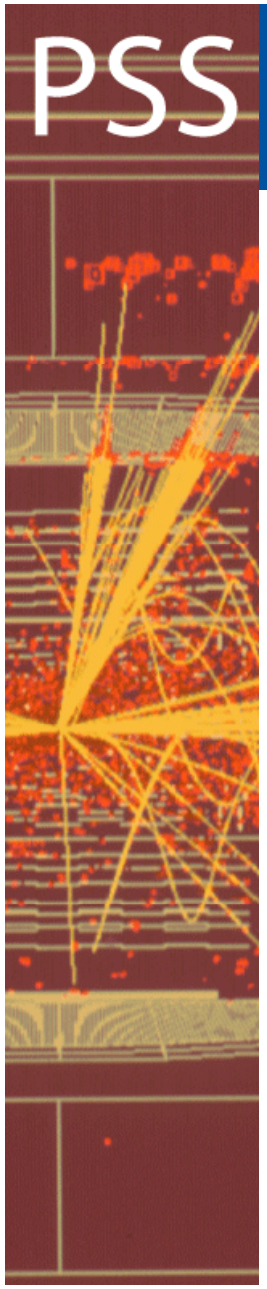
CERN - IT Department
CH-1211 Genève 23
Switzerland
maria.girone@cern.ch

Current set-up at Tier0

- RAC on commodity hardware
  - Full redundancy
  - Linux RHES4 32bit as OS platform
  - Oracle ASM as volume Manager
  - 10g version of Oracle (10.2.0.3)
  - Dual-CPU P4 Xeon servers with 4GB of DDR2 400 memory each



CERN - IT Department
CH-1211 Genève 23
Switzerland
maria.girone@cern.ch



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

- Several production clusters
  - One production cluster per LHC experiment for offline applications, varying from 4 to 8-node clusters
  - Online test Atlas cluster
  - COMPASS cluster
- Several validation and test clusters
  - 1 or 2 per LHC experiment of 2-nodes
  - Some hardware allocated for internal use/tests
- Service responsibilities
  - 6 DBAs in the team
  - 24x7 service on best effort for the production service
  - 8x5 coverage for streams

CERN - IT Department
CH-1211 Genève 23
Switzerland
maria.girone@cern.ch

PSS

Current set-up

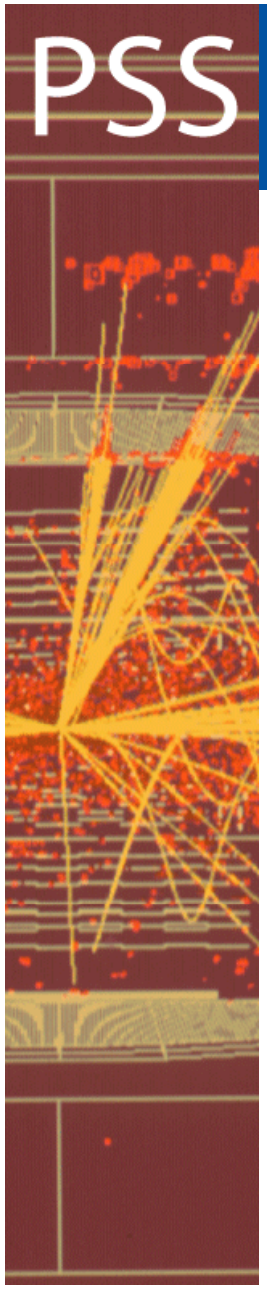
CERN IT Department

CERN - IT Department
CH-1211 Genève 23
Switzerland
maria.girone@cern.ch

2007 hardware allocation

	2007		
Experiment/ activity	**Production**	**Integration/ Pilot/Test**	**On-line test**
ALICE	On PDB	-	1 x 2 nodes (PVSS tests)
ATLAS	1 x 6 nodes	2 x 2 nodes	1 x 6 nodes
CMS	1 x 8 nodes	2 x 2 nodes	
LHCB	1 x 6 nodes	1 x 2 nodes	
WLCG	1 x 8 nodes	2 x 2 nodes	-
PDB (alice, harp, totem)	1 x 4 nodes	-	-
COMPASS	1 x 4 nodes	-	-
3D	3 nodes (downstream capture for atlas and lhcb) 1 node (OEM)	1 x 2 nodes	-
Test		1 x 4 nodes (coral) 1 x 6 nodes (cool) 1 x 2 nodes (recovery)	1 x 6 nodes (PVSS IT-CO)

CERN - IT Department
CH-1211 Genève 23
Switzerland
maria.girone@cern.ch



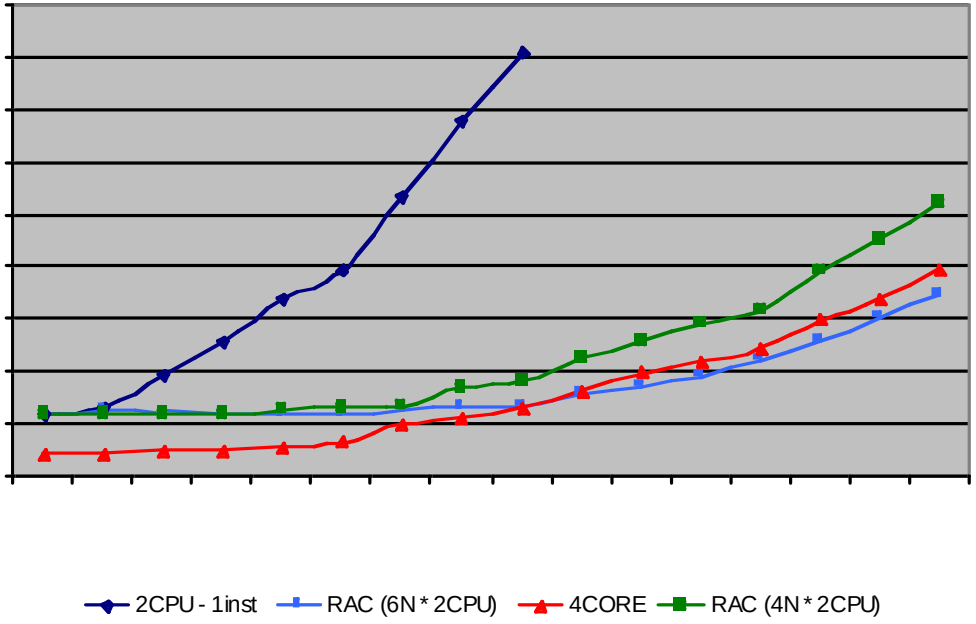
Quad-core/64 bit tests

- We have tested a 2x Intel quad-core CPUs and 16GB of RAM and compared with current Tier0 hardware (2x P4 CPUs)
- Natural point to introduce 64-bit Oracle servers
- All results are available at https://twiki.cern.ch/twiki/bin/viewauth/PSSGroup/QuadCoreTests

CERN - IT Department
CH-1211 Genève 23
Switzerland
maria.girone@cern.ch

Response time measurements

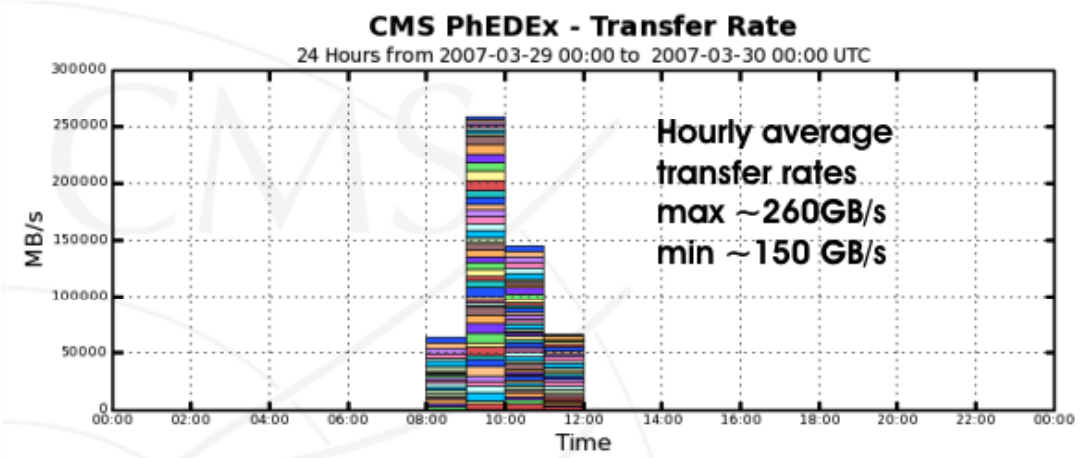
- Measurements of CPU to memory access were performed using a ORACLE workload (Jonathan Lewis Oracle Computing Index query JLOCI)



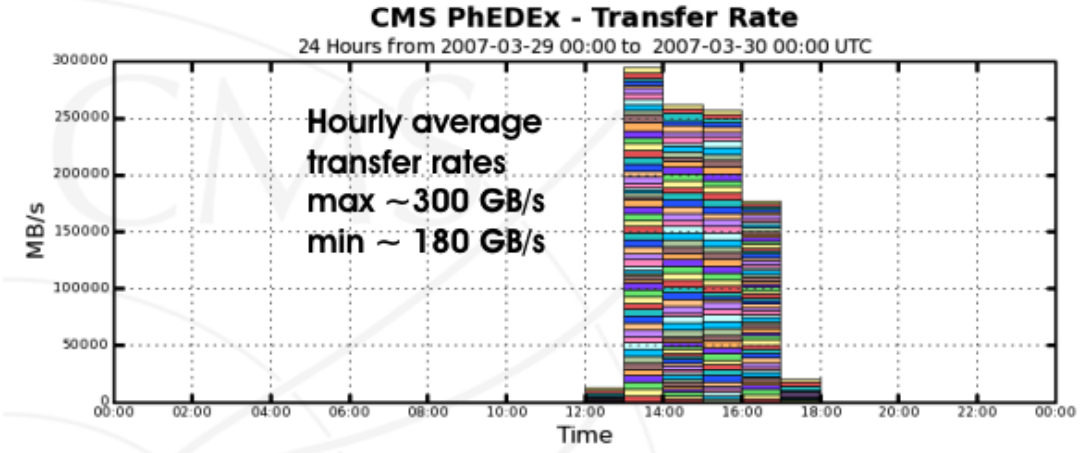
- The quad-core server can sustain the same logical IO workload as 5 of the servers currently deployed in production

CERN - IT Department
CH-1211 Genève 23
Switzerland
maria.girone@cern.ch

Performance testing with PhEDEx



6-node RAC



Quad-core server

A single quad core server is able to handle PhEDEx-like workload (a transaction oriented application) even more efficiently then a 6-node RAC

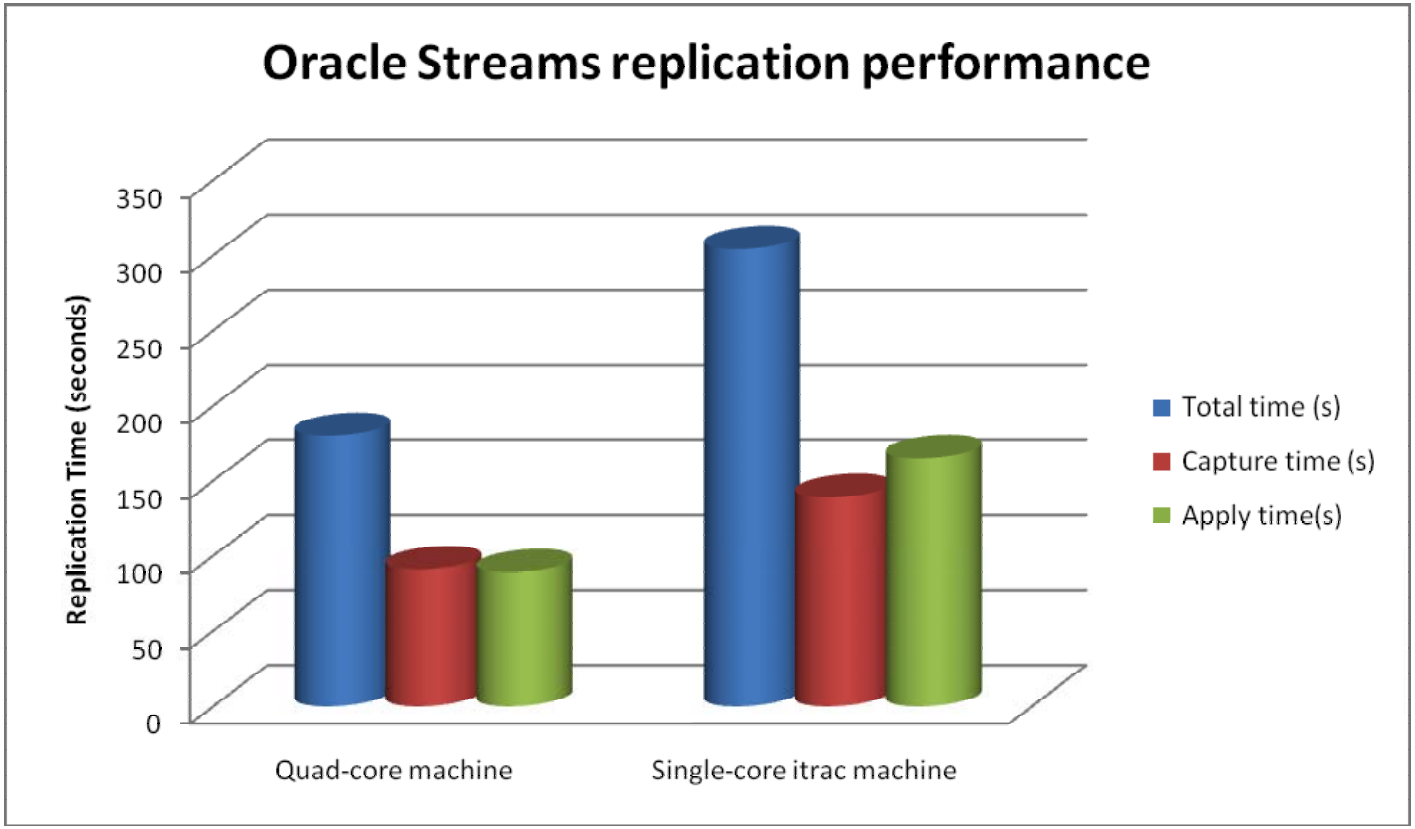
CERN - IT Department
CH-1211 Genève 23
Switzerland
maria.girone@cern.ch

Quad-core performance and Oracle Streams

PSS

CERN IT Department

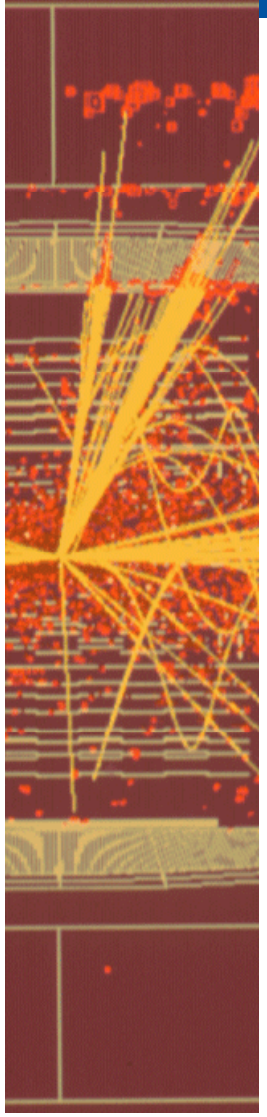
The source DB on a 2-node RAC on P4, destination (apply) machine on a single-core and quad-core DB server



The quad-core results show a 70% increase in performance compared to the current production HW

CERN - IT Department
CH-1211 Genève 23
Switzerland
maria.girone@cern.ch

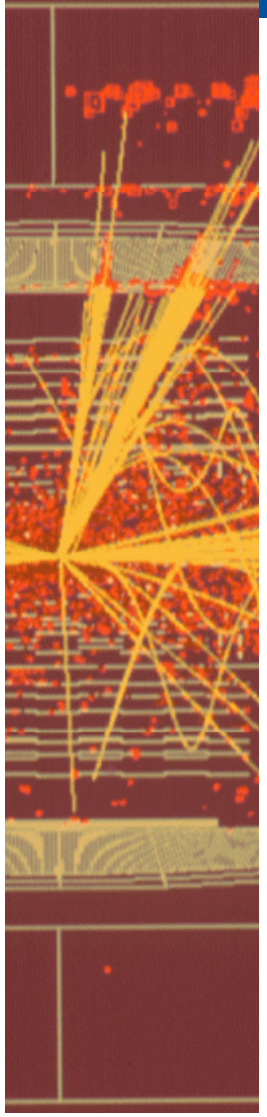
Quad core/64 bit tests (cont'd)



- Measured power consumption per core showed a gain in efficiency for quad-core compared to current RAC production of a factor 2
- Installation procedures for quad-core and RHEL 4 do not need additional effort compared to the existing (there are minor differences with current production RACs)
- SAN configuration and I/O throughput have been tested and are unchanged from the existing RAC configuration.

CERN - IT Department
CH-1211 Genève 23
Switzerland
maria.girone@cern.ch

Hardware allocation in 2008



- Production databases for LHC:
  - 3-4 node clusters built with quadcore CPU machines (24-32 cores per cluster)
  - 48-64 GB of RAM per cluster
  - Planning for >10k IOPS
  - TBs of mirrored space
- Integration and test systems:
  - Old single core CPU hardware
  - Usually 2 nodes per cluster
  - Usually 24-32 disks
- 64bit version of Linux and Oracle software

CERN - IT Department
CH-1211 Genève 23
Switzerland
maria.girone@cern.ch

Conclusions

- Tests on dual quad-core Xeon processors look very promising and were compared with dual-CPU P4 http://twiki.cern.ch/twiki/bin/viewauth/PSSgroup/QuadCoreTests
- We have decided to base our next order on this choice
- Natural move to Oracle 64 bit
- Migration tools are being prepared and tested to minimize the downtime of the production RACs

CERN - IT Department
CH-1211 Genève 23
Switzerland
maria.girone@cern.ch