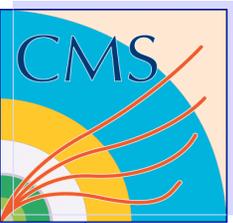


# Facility Operations and Upgrades, Procurements

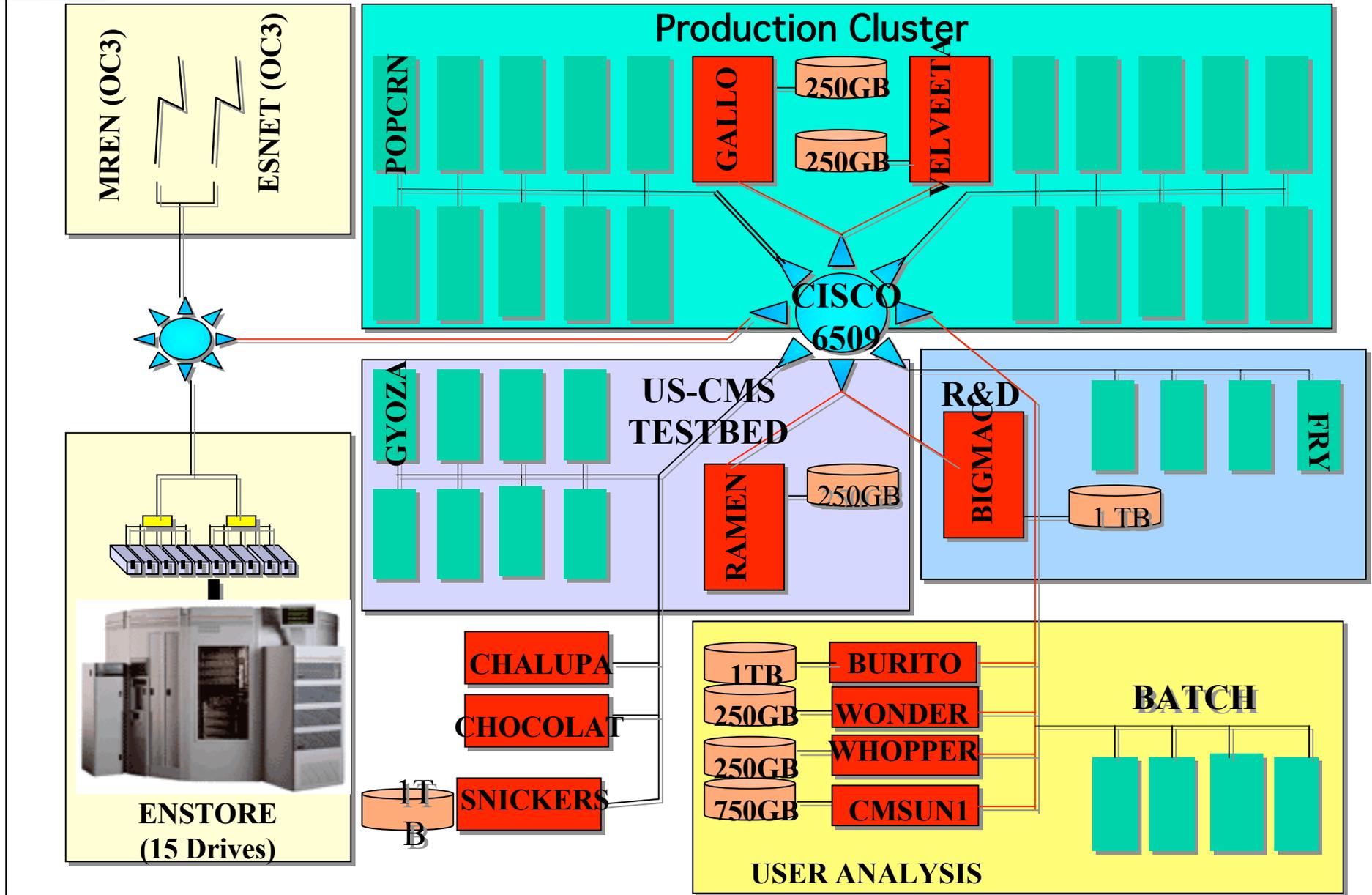
Hans Wenzel  
Fermilab

- ❑ The project is tracked by a WBS which was just recently completely revised. Emphasize on rolling prototype, closer integration of Tier 2 sites (system), distributed production, simplified more realistic hardware planning.
  
- ❑ Activities and major upcoming milestones:
  - ❑ be operational as part of “LCG Production Grid” in June 2003
  - ❑ participation in the CMS “5% data challenge” DC04
  - ❑ Support user computing e.g. test-beam activities (user interactive, user batch, host data, user support, tutorials ....)
  
  - ❑ Active R&D



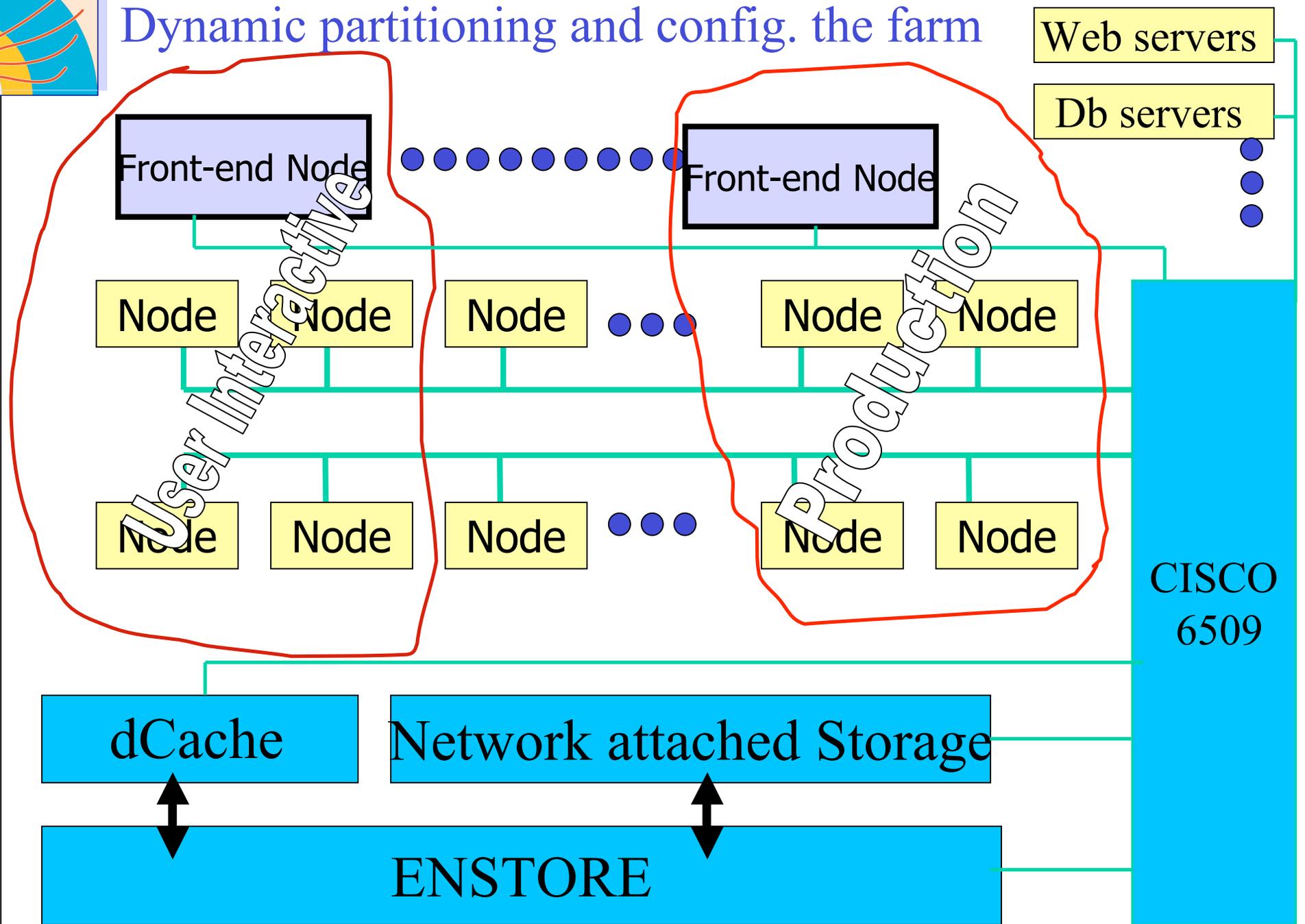


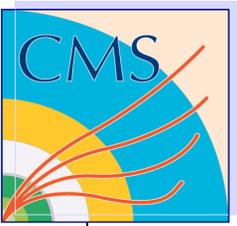
# Status of T1 in spring 2002





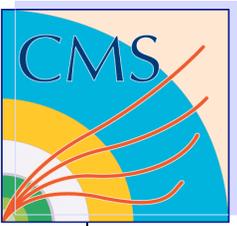
# Dynamic partitioning and config. the farm



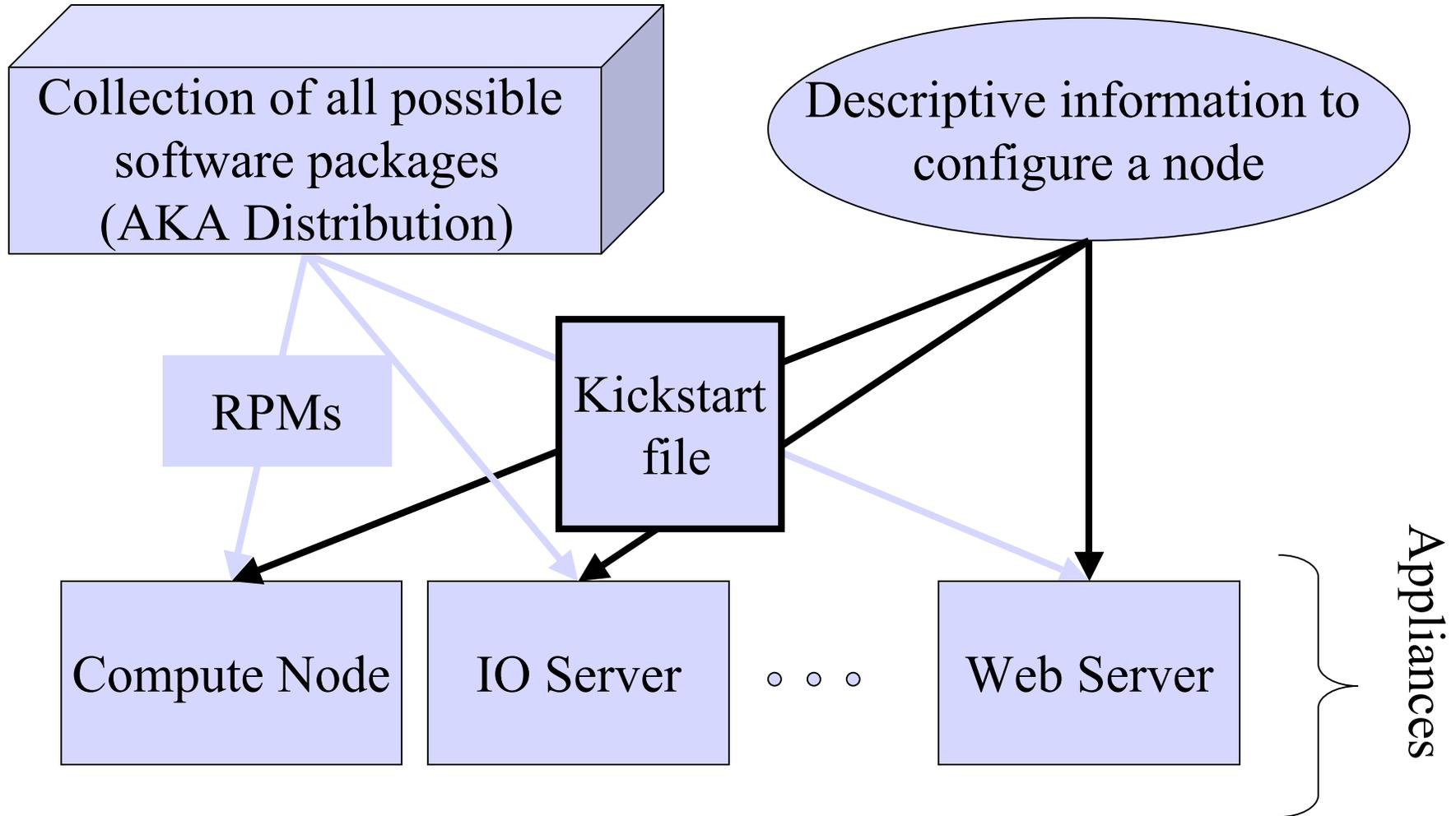


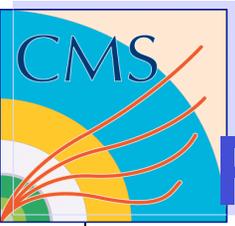
## R&D efforts concentrate on

- ❑ Continue Evaluating disk systems (commercial: zambeel, panasas, exanet, spinnacker HEP: dCache, dfarm (local disk on farm nodes).....).
- ❑ We think it is absolutely essential to procure a NAS system this year! (provide work space for users, allow for dynamic partitioning...)
- ❑ Load balancing scheduling: Configure and deploy farm based interactive/batch user computing (FBSNG/LVS) a la Ixplus cluster at CERN.
- ❑ OS and farm configuration: Continue to evaluate ROCKS, YUM, autorpm for security patches....
- ❑ Software configuration and deployment: cfengine, LDAP, rsync



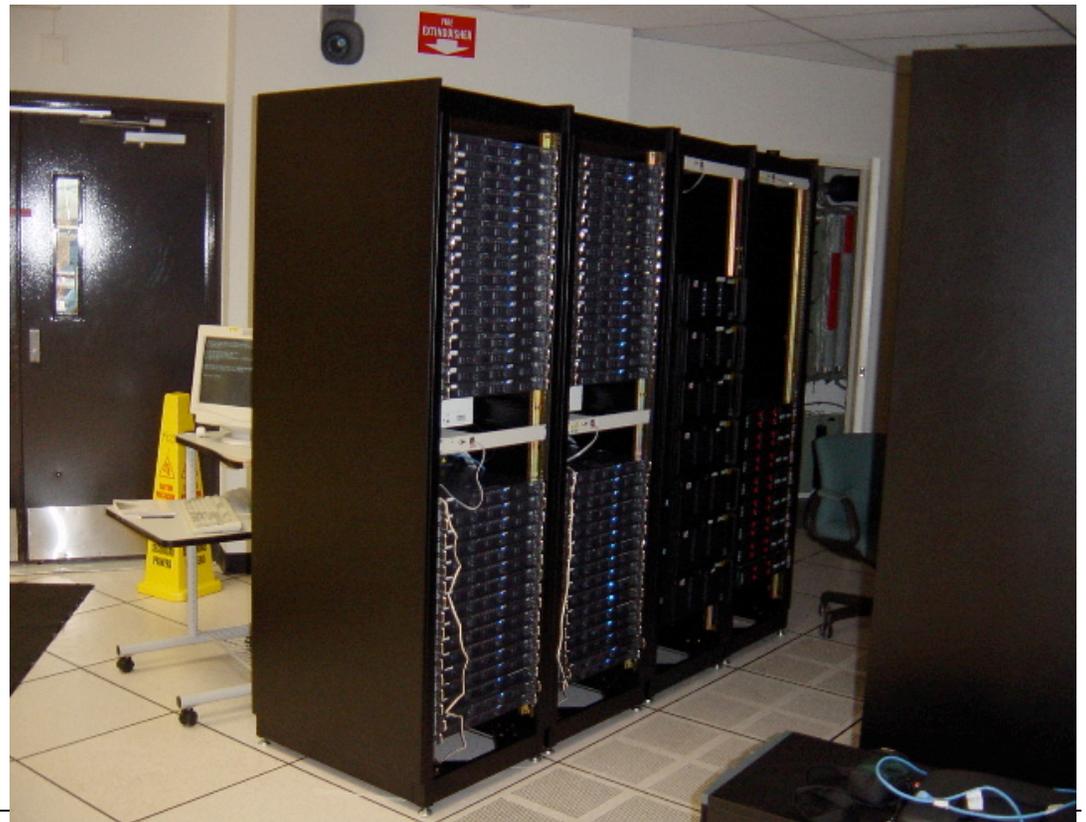
# NPACI ROCKS

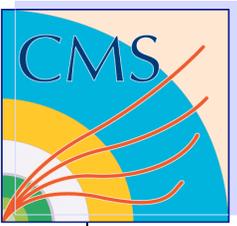




# Recently completed hardware upgrade of facility

- ❑ 65 dual farm nodes (amd 1.9 Athlon 1U form factor) were installed and finished acceptance (>98 % uptime for 4 weeks) testing before thanksgiving.
- ❑ Installed 7 dCache servers (2 reserved for R&D).
- ❑ Completed remote power automation project, which required complete rewiring and racking of nodes.
- ❑ Better connectivity between cms computing and central switch.
- ❑ Now FNAL has OC12 network connection to outside world.
- ❑ Installation and configuration of private network for ROCKS.





# Linux dCache node:

3 ware 7850 based disk servers made of off the shell components are in very wide use in the community, CMS T1 and T2 Sites, CDF, SDSS...lots of experience how to set them up available, list of parameters that have to be tweaked for optimum performance Important Kernel 2.4.18 to avoid memory management problems XFS file system: we found it's the only one that scales still Delivers performance when File system is full

Add SCSI system disk

Need server specific

Linux distribution!!!

Need to tweak

Many parameters to

Achieve optimum

Performance-> feed back to vendors

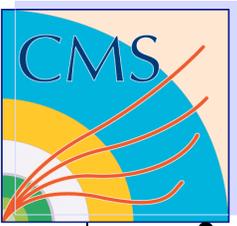
Next generation:

Xeon based, PCIX bus

Large capacity disks

Dual System disk (raid1)





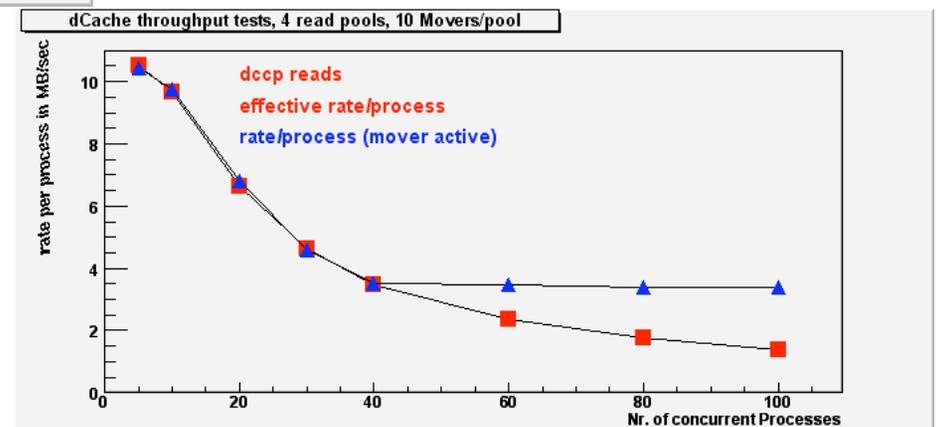
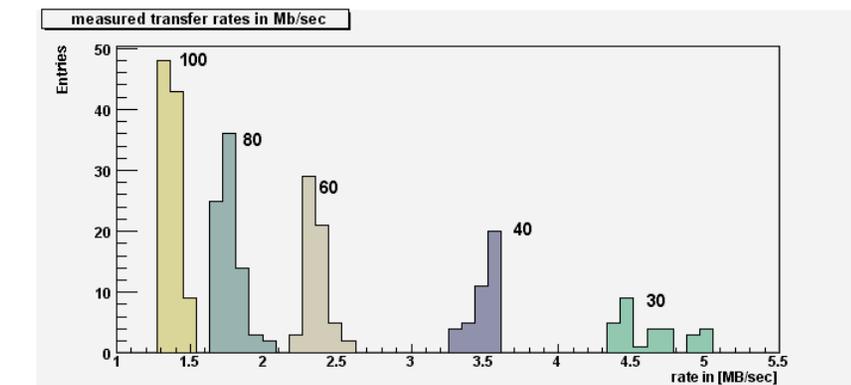
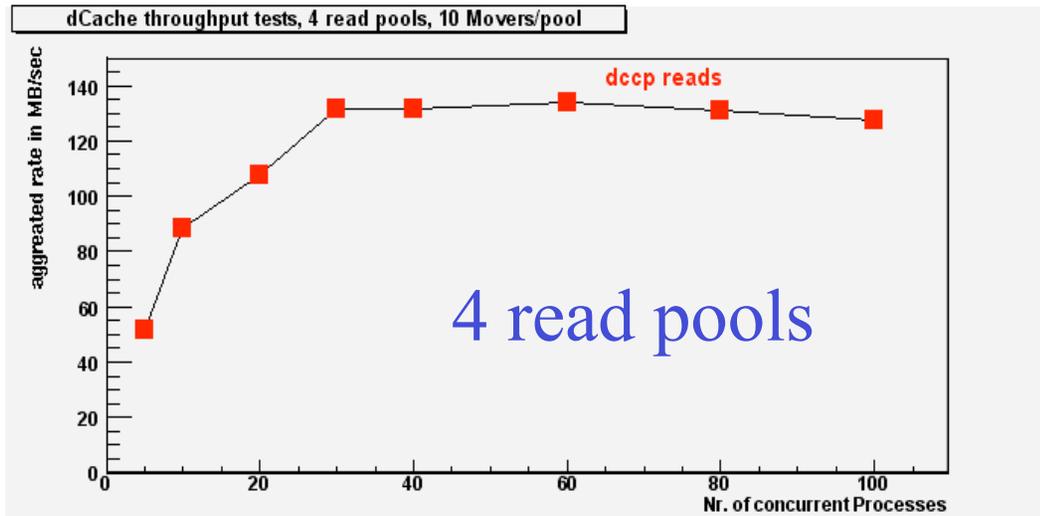
# results with dCache system

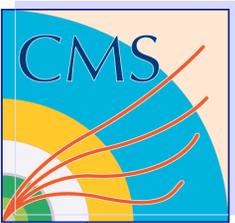
Developed test suite to measure performance and analyze behavior of system. The average file size is ~1.2 GBytes residing on disk the reads are equally distributed over all read pools. Reads with dcp from nodes into /dev/null. TDCacheFile -> transparent access to

MSS

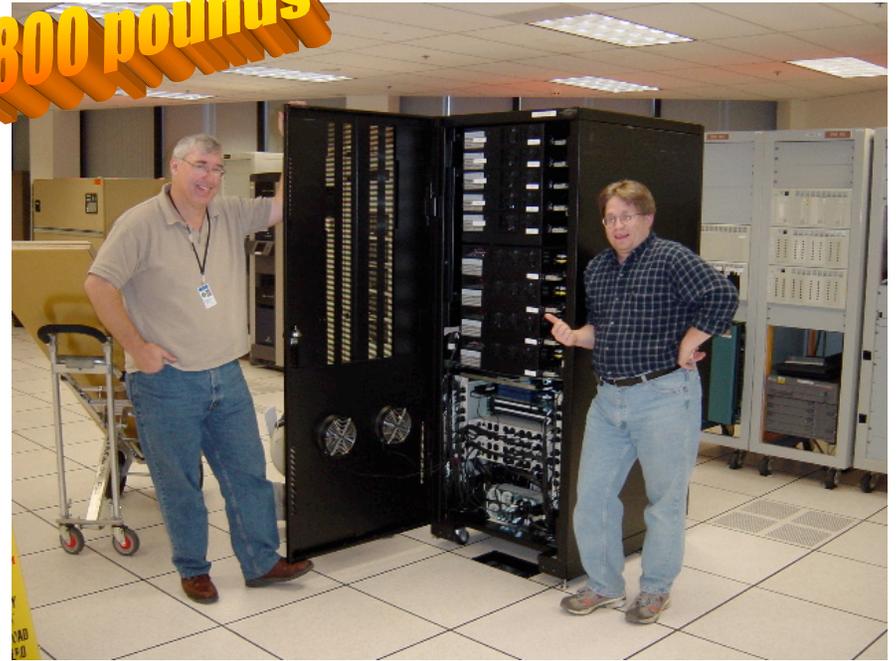
dCache Status:

<http://gyoza7.fnal.gov:443/>



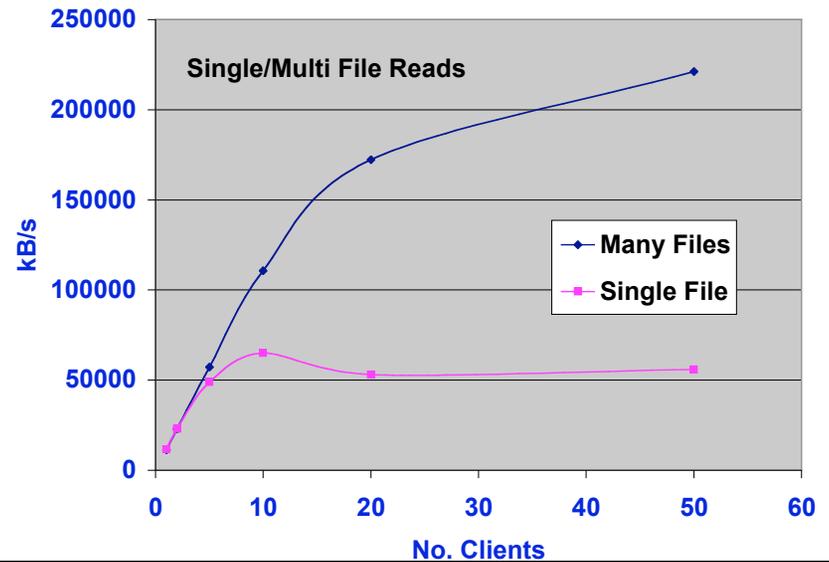
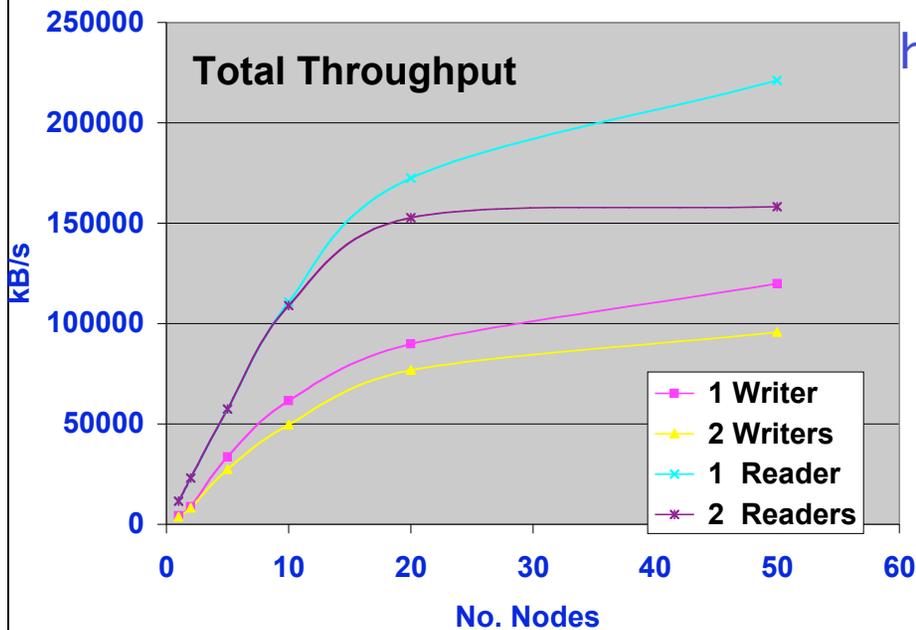


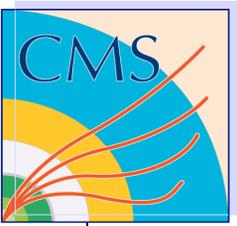
1800 pounds



## Aztera results (after performance tuning)

<http://www-oss.fnal.gov/projects/disksuite/>

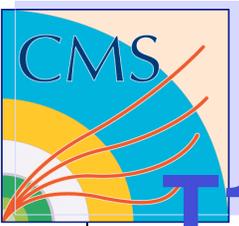




# Facilities procurements in FY 2002

Tape drives + Movers	31 K\$
Tape media (800 x 60GB)	40 K\$
Networking Equipment	74 K\$
7 File servers	63.5 K\$
Extend hardware warranties and software licenses	16.6 K\$
P4 evaluation model	1.2 K\$
<b>Total (without Fermilab overhead)</b>	<b>235.8 K\$</b>

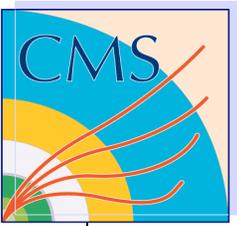
Did not buy all the equipment originally planned



# T1 UF procurements in FY 2003 (so far)

## Fermilab overhead not included

65 farm nodes (dual 1900+ AMS) +racks+ console servers	125 K\$
2x 9940b tape drive	60 K\$
Tape media for 9940 300 (60GB 9940A, 200GB 9940B)	18 K\$
System upgrades	3.1 K\$



# FNAL vendor evaluation

CMS takes actively part in FNAL Linux vendor qualification taskforce (which has representatives from all experiments at FNAL + CD).

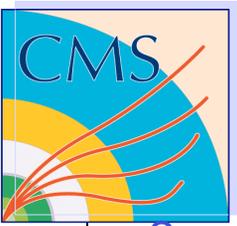
Currently concentrate on farm nodes. Would be great to do it for Linux servers and desktops (laptops hopeless).

The schedule is the following (good match for CMS):

- Jan 14th produced a requirements document describing the qualification procedure and hardware configuration sent to vendors (> 40 vendors)

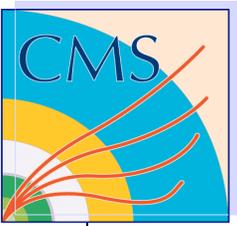
<http://www-oss.fnal.gov/scs/public/qualify2003/>

- 28 January vendor information meeting at FNAL.
- February 4th get configuration and price proposal from vendors  
(full dollar PO will be written reserving the right to return after evaluation)
- Mar 4. machine build and initial benchmarks run and e-mailed to FNAL
- Mar 11 machines physically at FNAL.
- mar 13-27 evaluation period
- one week later decision on qualified vendor list
- April request for bids for CMS purchase



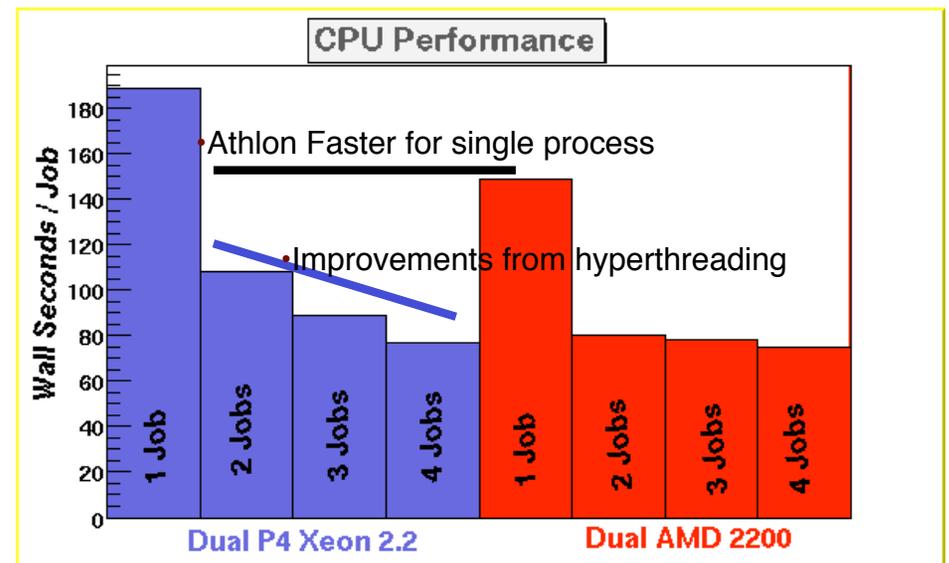
- o Vendors are judged for:
  - o Linux Competence
  - o Ability to provide quality product 1U dual xeon >2.4 Ghz (mechanics, cooling ....) in time
  - o Ability to provide competent support
- o Performance
  - o Seti-at-home CPU test
  - o Bonnie disk IO benchmark (> 20MB/sec)
  - o Memory Bandwidth using the stream benchmark
  - o Physics applications
  - o Performance measured in fermi cycles  
1GHz PIII ~ 1000fc
  - o Make info available on the web.

System	Copy (MB/sec)
i860 Chipset RAMBUS	1473
e7500 Chipset DDR	1262
GC-LE Chipset DDR	1006
AMD 760 Chipset	679



# System Tests (cont.)

- o Using a CMS standard application
  - o Simulation, reconstruction, digitization
  - o Measure the time per job
  - o Tested Xeon and Athlon based systems
    - o Athlon systems are faster on a single application as seen from the single job rate (AMD machines won bids for last farm upgrade)
    - o Gap closes when multiple applications are run with hyperthreading enabled

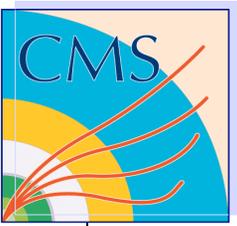




# We are considering the following procurement in FY 2003

(Some items were already planned for FY 2002, numbers are upper limits)

32 farm nodes (dual xeon > 2.4 GHz) +racks+ console servers (purchased to meet DC04 preproduction challenge)	<90K \$
Tape media	50K \$
NAS system for user data and as farm IO server	< 200K \$
Backup system	<60K \$
16 port GBit Ethernet module for CISCO Switch	15K \$
Next generation dCache Servers (5)	50 K \$
Electrical engineering to provide sufficient power	?? \$
Total:	465 K \$



# Summary

- o We successfully deployed a Tier 1 center prototype.
- o We have successfully taken part in large scale production using Grid technology.
- o Will provide strong platform for physics analysis: hosting large data sets and providing CPU power.
- o Facility upgrade geared toward being able to play a strong role in DC04.