Batch Scheduling on XT3

Chad Vizino <vizino@psc.edu>
Pittsburgh Supercomputing Center

CUG 2005

Overview

- Simon Scheduler
  - Design
  - Features
- XT3 Scheduling at PSC
  - Past
  - Present
  - Future

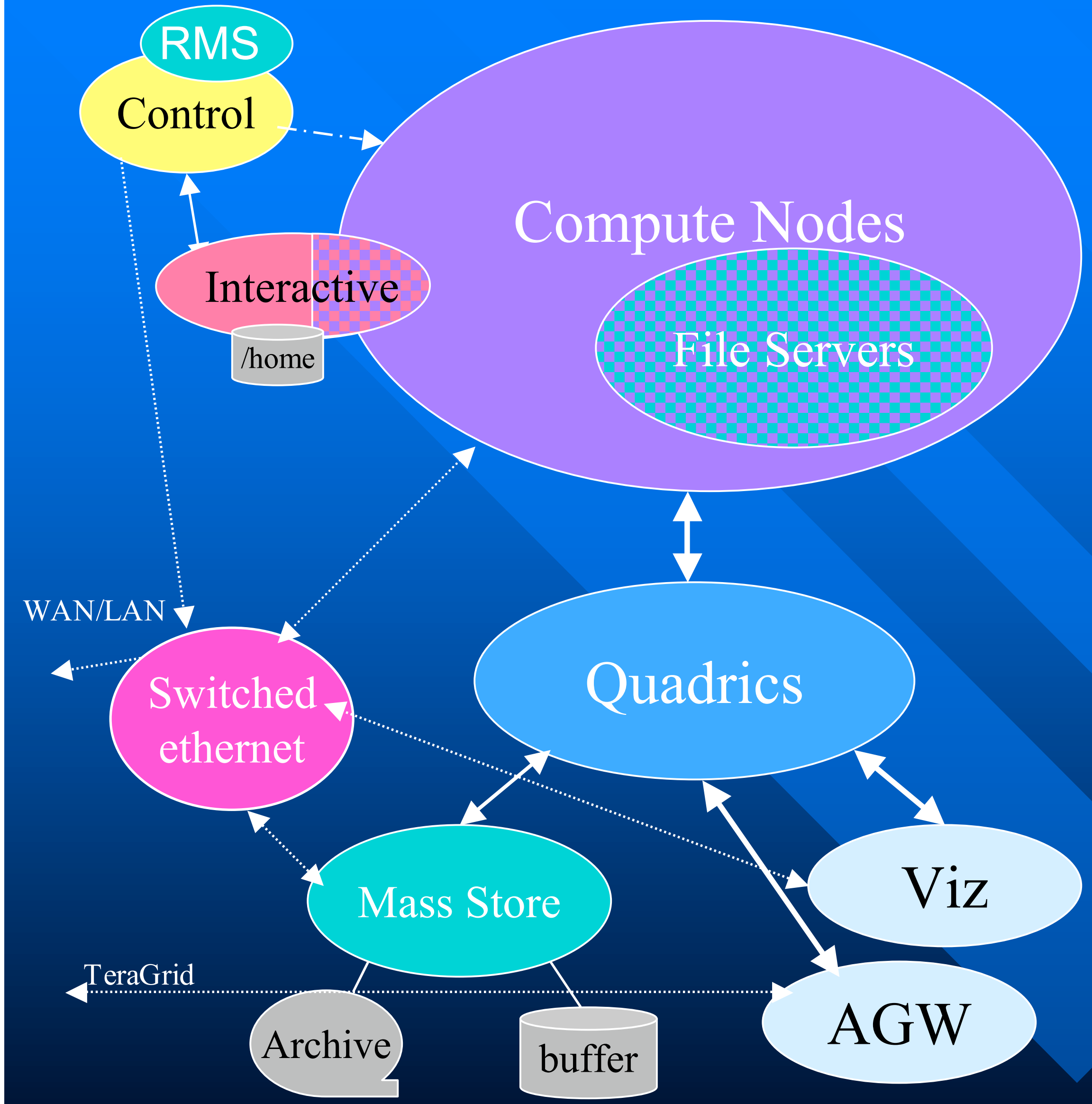
Back…

…to the Future!

Scheduler Design Goals

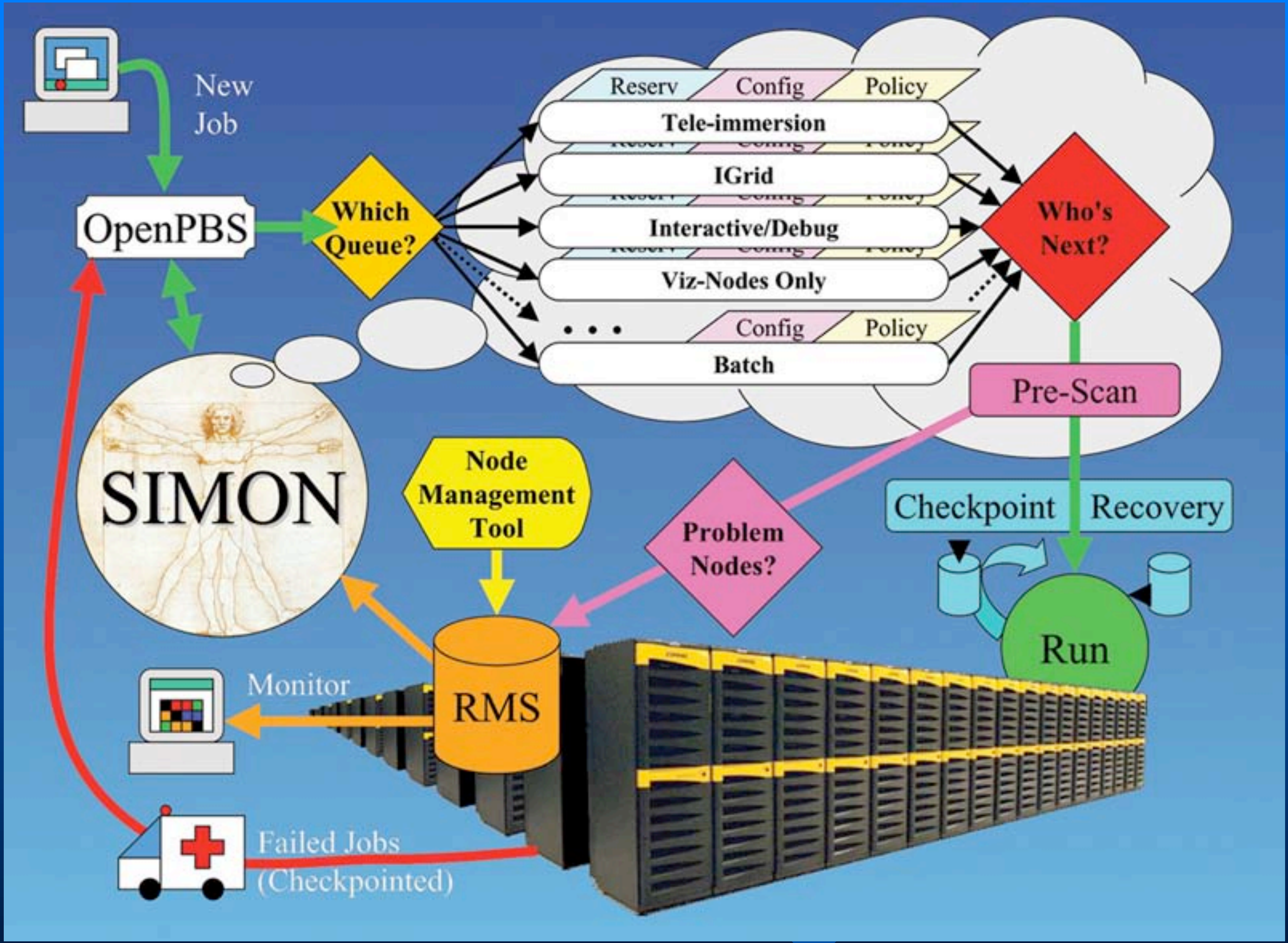
- Support PSC Scheduling Goals
  - Encourage Large Jobs
  - Foster Parallel Development
  - Fair
- Maintain good utilization

Terascale Computing System (LeMieux)



Summary

- 750 Compute Nodes
- 3000 EV68 processors
- 6 Tf (peak,
  est >4Tf on LSMS)
- 3. TB memory
- 27 TB local disk
- Multi-rail fat-tree network
- Redundant monitor/ctrl
- WAN/LAN accessible
- Parallel visualization
- File servers:
  30TB, ~32 GB/s
- Resource Management System (RMS)


New
Job
OpenPBS
Which Queue?
Reserv
Config
Policy
Tele-immersion
IGrid
Interactive/Debug
Viz-Nodes Only
• • •
Batch
Who's Next?
Pre-Scan
Checkpoint
Recovery
Run
SIMON
Node Management Tool
Problem Nodes?
RMS
Monitor
Failed Jobs
(Checkpointed)

Simon Batch Scheduler

- OpenPBS
- 6,500 lines of TCL
  - Rapid prototyping in TCL (pbs_tclsh)
  - Easy to add new modules
  - Only scripting language supported by OpenPBS
- Over four years in production
- Backfill
- Reservations
- Node Maintenance

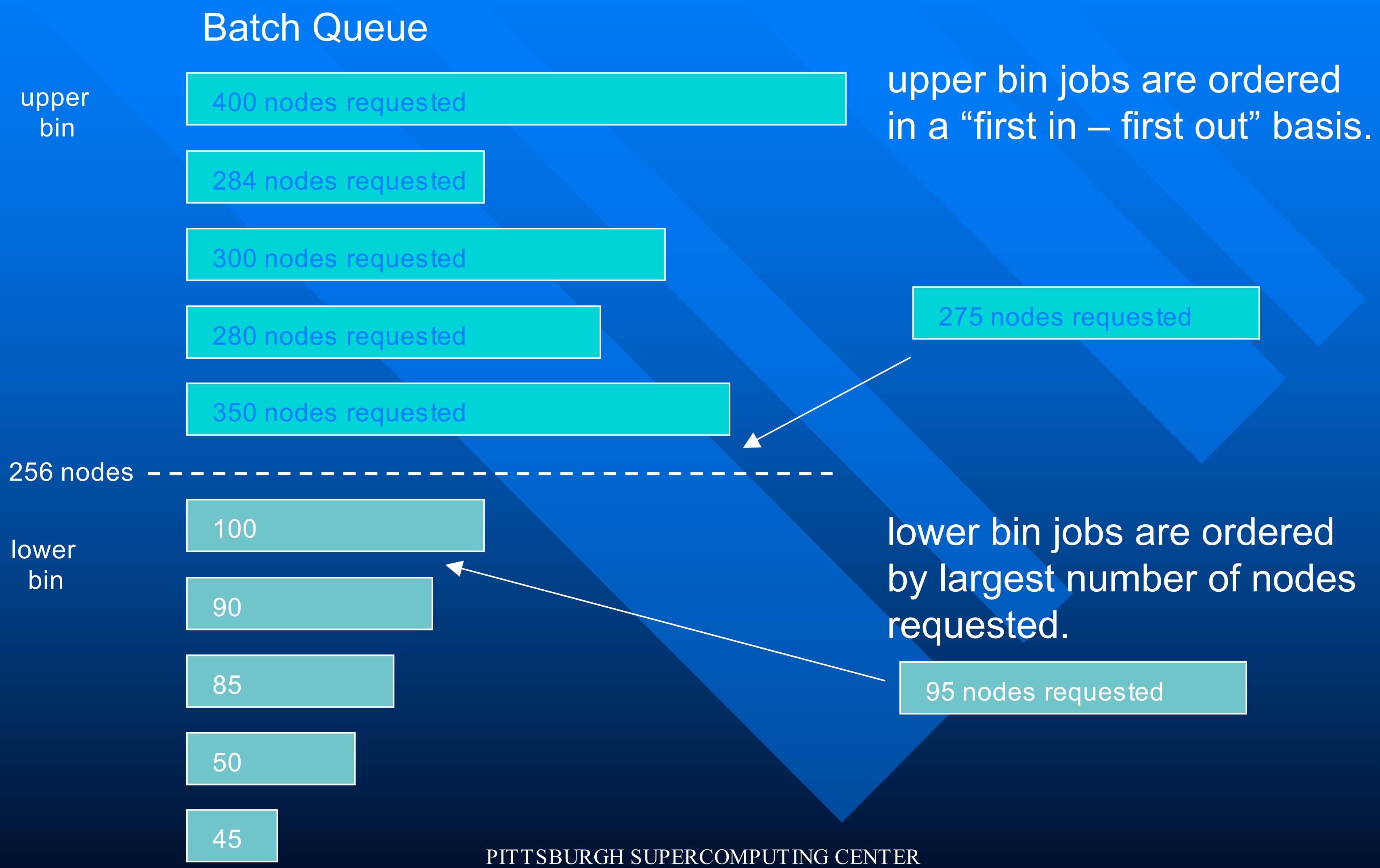
Simon Batch Scheduler

- Co-scheduling
  - Visualization cluster
  - Application GateWay nodes
- Pre-job scanning
- Ties to Usage Accounting
  - Queue statistics
  - Reservation charges

Simon Namesake

- "Simon" is named for Dr. Herbert Simon (1916-2001), University Professor of Computer Science and Psychology at Carnegie Mellon University, and winner of the 1978 Nobel Prize in Economics.
- Argued that inevitable limits on knowledge and analytical ability force people to choose the first option that "satisfices" or is good enough for them.
- Scheduling a large computing system often requires making choices with limited knowledge.

PSC Scheduling



Batch Queue

upper
bin

400 nodes requested

284 nodes requested

300 nodes requested

280 nodes requested

350 nodes requested

upper bin jobs are ordered in a “first in – first out” basis.

275 nodes requested

256 nodes

lower
bin

100

90

85

50

45

lower bin jobs are ordered by largest number of nodes requested.

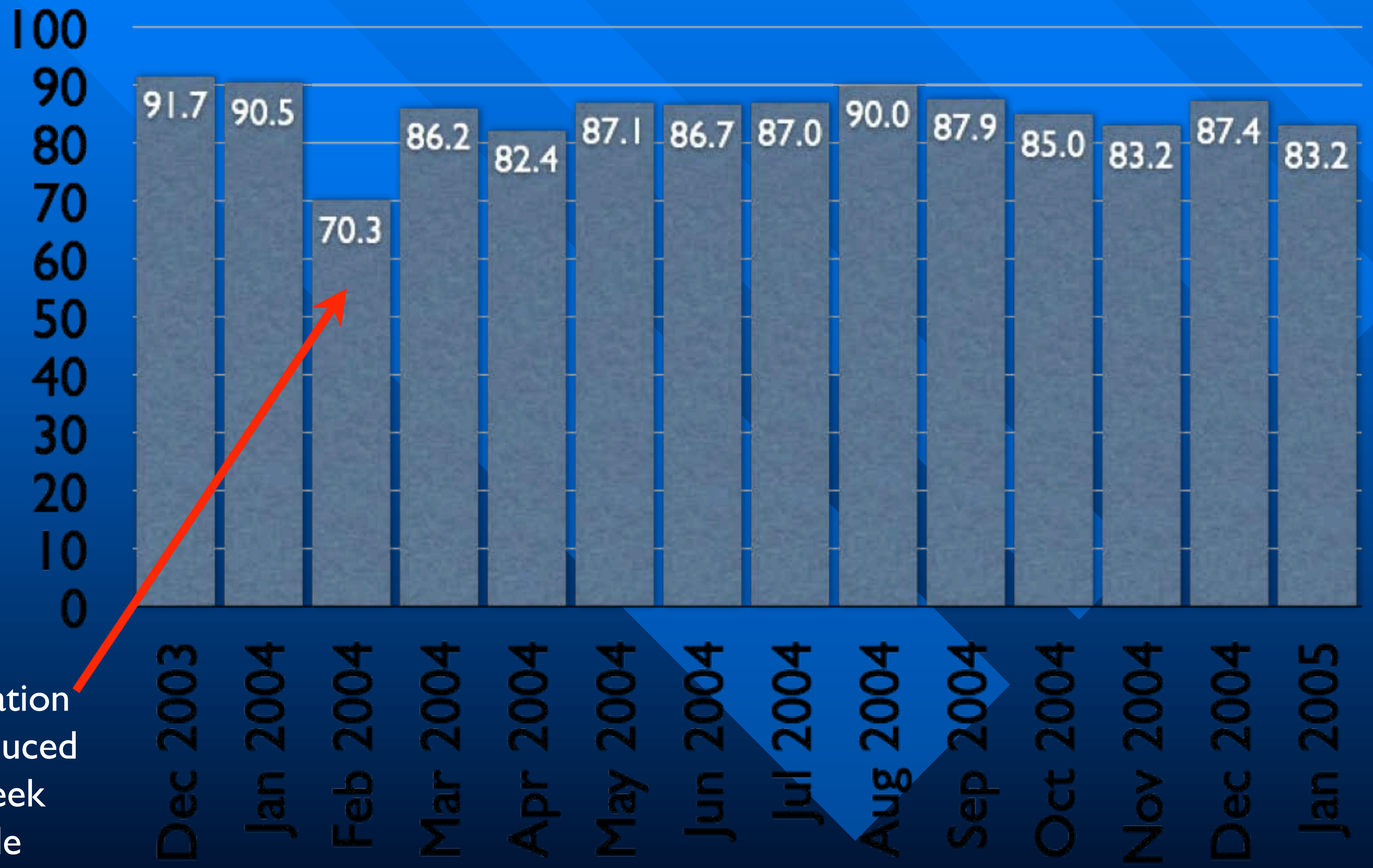
95 nodes requested

LeMieux >= 1024 Processor Usage

Year	% Processor*Hours
2002	6.5
2003	28.6
2004	46.7

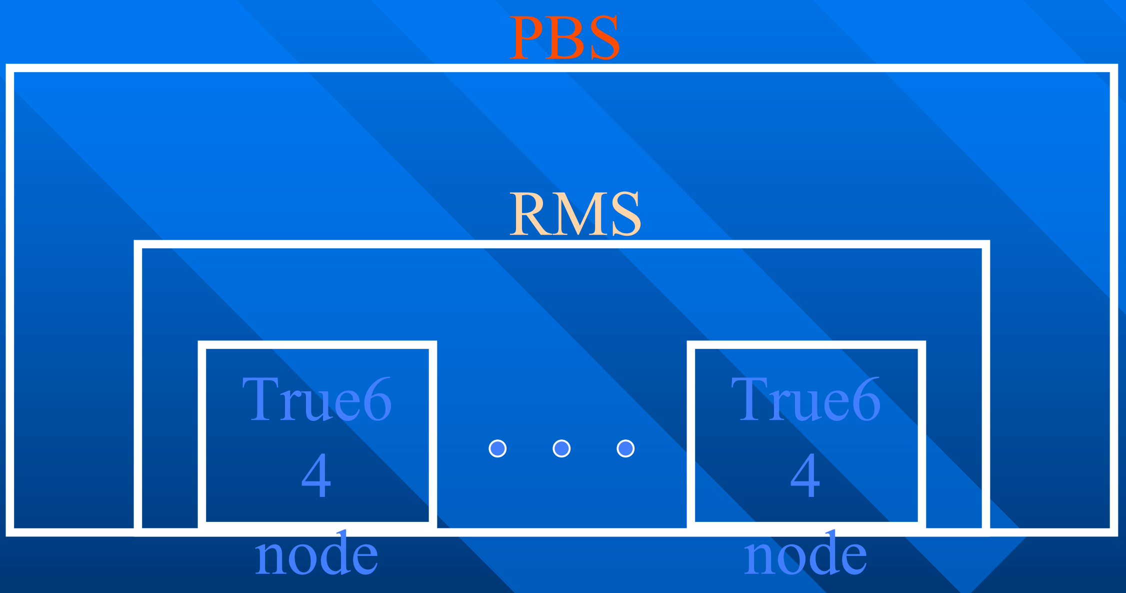
There are no spare cycles...

Average Daily % Utilization of TCS (lemieux.psc.edu)
750x4p Alpha EV68



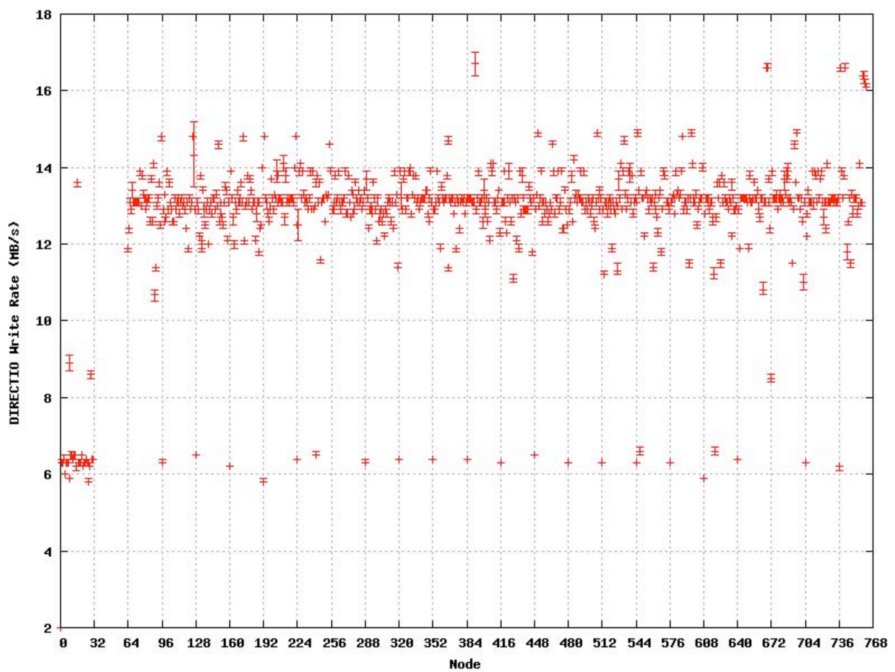
Average utilization for month reduced due to one week system upgrade downtime

PBS-RMS Relationship



Scheduling Support for Job Success

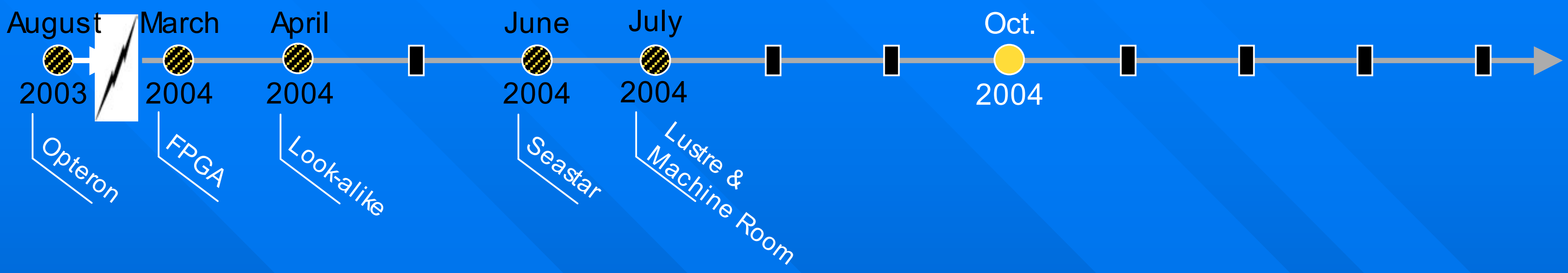
- Pre-Scan
  - File Systems
  - Orphan/busy processes
  - CPU availability/accuracy
  - Interconnect availability and performance
  - Failure removes the node from scheduling
- Job Profiling
  - Disk Performance
  - Live network connections
  - Application Fingerprints



Move to the recent past…

…the XT3

Single Cabinet System



August 2003 – Opteron

March 2004 – FPGA

April 2004 – Look-alike

June 2004 – Seastar

July 2004 – Lustre & Machine Room

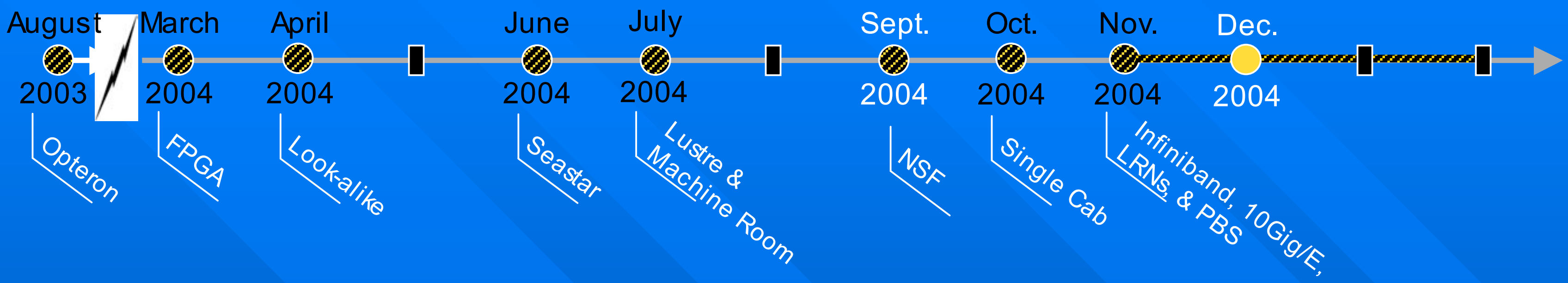
Oct. 2004

Single cabinet

October 2004

Demonstrated Running Applications at SC04

XT3 First Row Installed



August 2003 – Opteron

March 2004 – FPGA

April 2004 – Look-alike

June 2004 – Seastar

July 2004 – Lustre & Machine Room

Sept. 2004 – NSF

Oct. 2004 – Single Cab

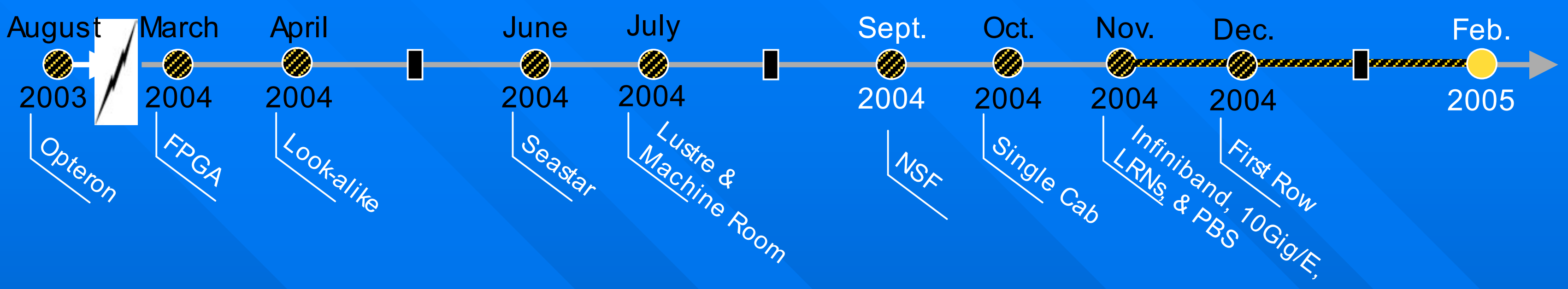
Nov. 2004 – Infiniband, 10Gig/E, LRNs, & PBS

Dec. 2004

Row 1

December 2004

XT3 Second Row Installed



August 2003 Opteron

March 2004 FPGA

April 2004 Look-alike

June 2004 Seastar

July 2004 Lustre & Machine Room

Sept. 2004 NSF

Oct. 2004 Single Cab

Nov. 2004 Infiniband, 10Gig/E, LRNs, & PBS

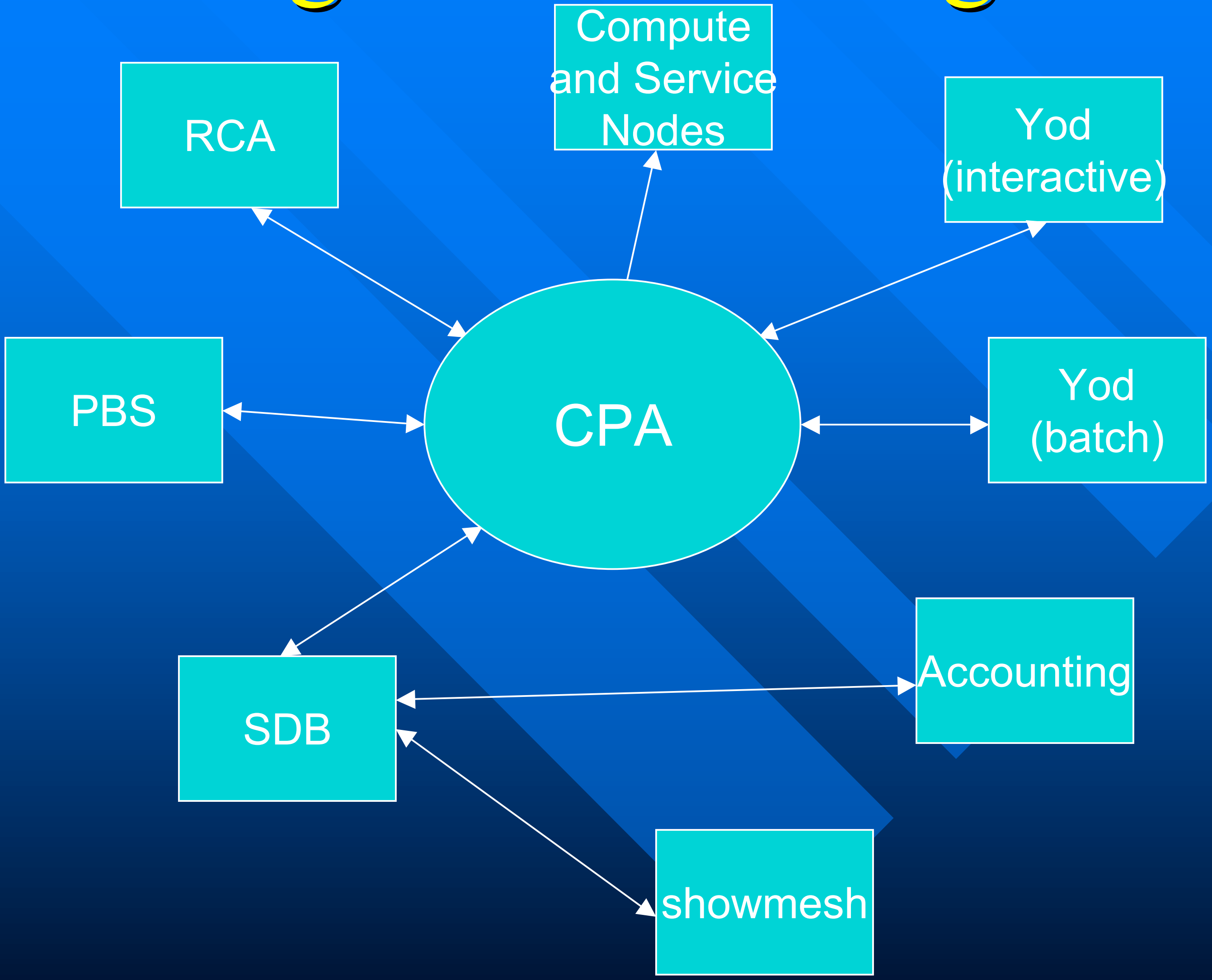
Dec. 2004 First Row

Feb. 2005

Row 2

February 2005

Original XT3 Design


RCA

Compute and Service Nodes

Yod (interactive)

PBS

CPA

Yod (batch)

SDB

Accounting

showmesh

Challenges

- System ran dev harness
  - – No SDB
  - – No CPA
  - – No functional batch system
    - » Major setback for efficient applications work
  - – Boot login nodes to clear bad nids
  - – Yod –list 10..20,30,50,90..100 ....
- Lack of unique, writable file system on login nodes
- System partitioned into separate dev harness systems

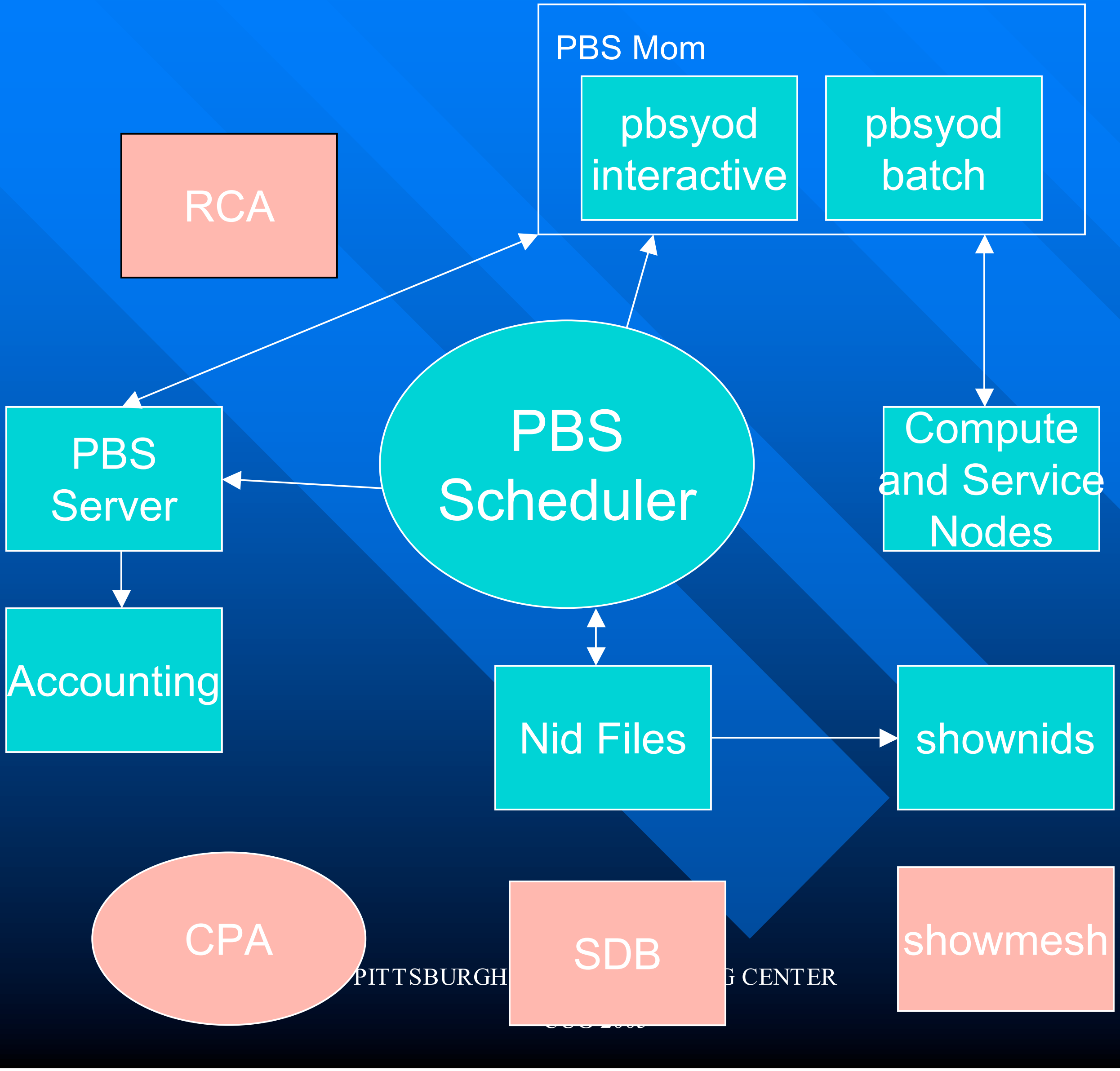
Solutions

- Use Torque (OpenPBS)
  - Builds on 64-bit platforms
  - Open source
  - Free
  - Have lots of experience with it
  - Can write custom scheduler
- Use RAM file system and load /var/spool/torque loaded from /usr/users (NFS mounted)
- Let flat files act as the SDB
- Let scheduler be the CPA
- Batch and interactive can be handled

SDB/CPA Replacement

- SDB processor table
  - /opt/harness/default/ssconfig/sysN/node_list
    - Cray managed (HW list)
  - /usr/users/torque/nids_list_loginN
    - Scheduler managed
    - Holds state of nids (enabled/disabled/allocated)
    - Query via shownids
- Yod
  - pbsyod reads YOD_NIDLIST from scheduler
  - -size, -base options to stack in job

Harness Layout



PBS Mom

pbsyod interactive

pbsyod batch

RCA

PBS Server

PBS Scheduler

Compute and Service Nodes

Accounting

Nid Files

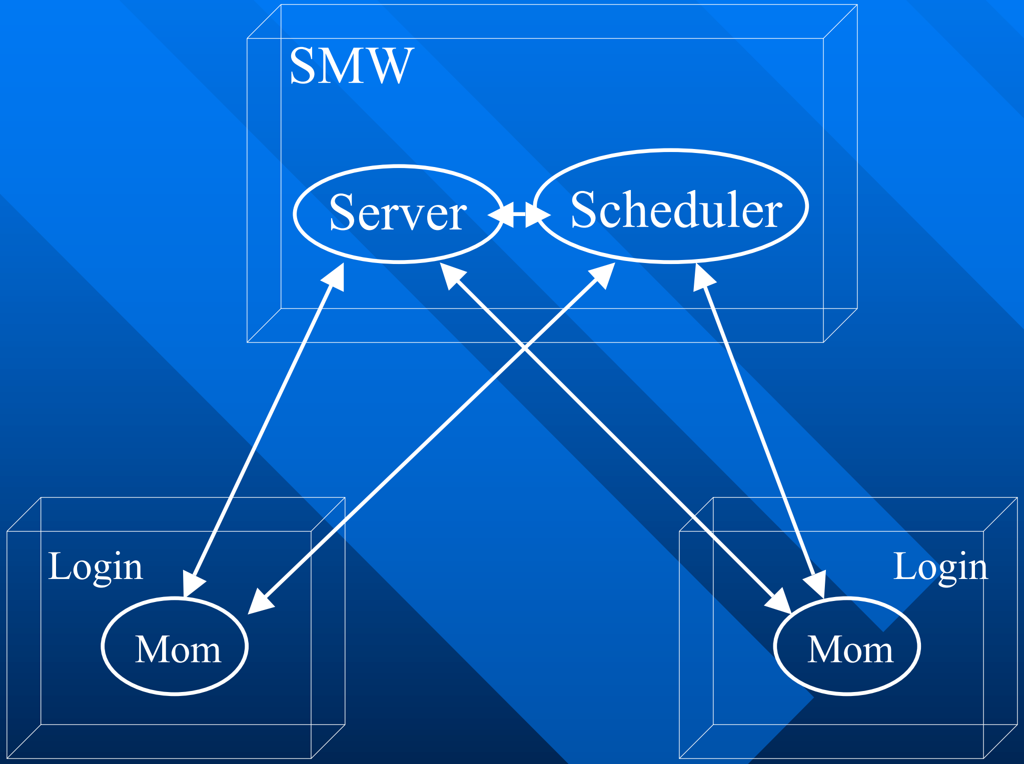
shownids

CPA

SDB

showmesh

Harness PBS Configuration



Early Scheduler Features

- Adaptation to harness
  - Schedule to multi-cab arrangement
- Backfilling
  - System drains
  - Use aggressive backfilling (EASY)
    - » Switchable (can just use First Fit)
  - Top job (largest) gets reservation
    - » Use FIFO ordering
  - Other jobs can run as long as they don’t delay the start of the top job (backfill)
- Pre-scan

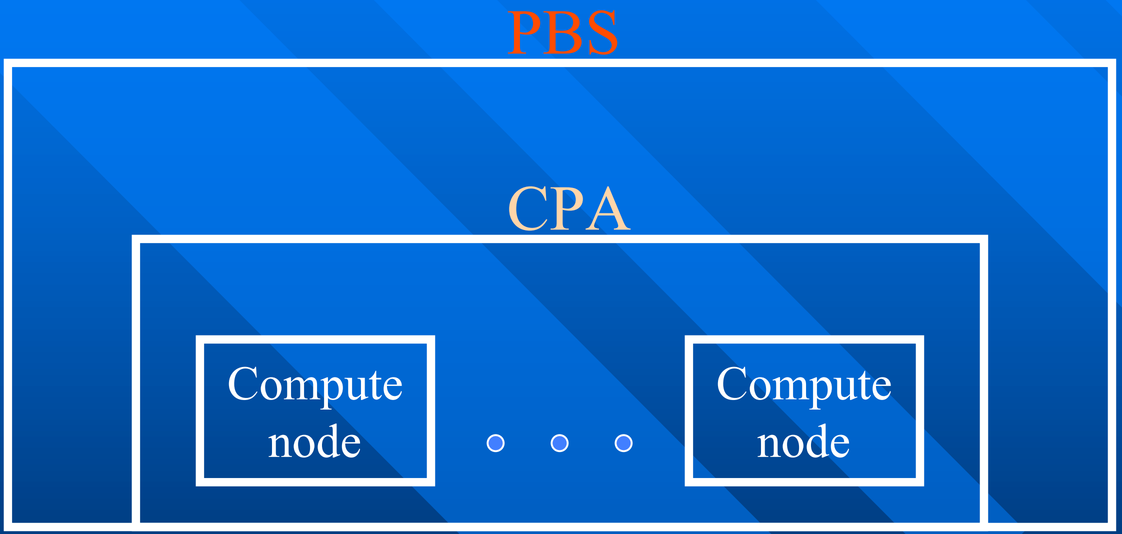
Defensive Mechanisms

- Maintain lists of node categories
  - Checking
  - Bad
- Started calling ping_node
  - After job finished
  - Slow
- Call ping_list before job starts
  - Fail nodes
- Develop new, faster ping_node
  - ping_list –l 12..15,17..95,200..295
  - Returns good and bad lists
- Automated reboots
  - Let scheduler control

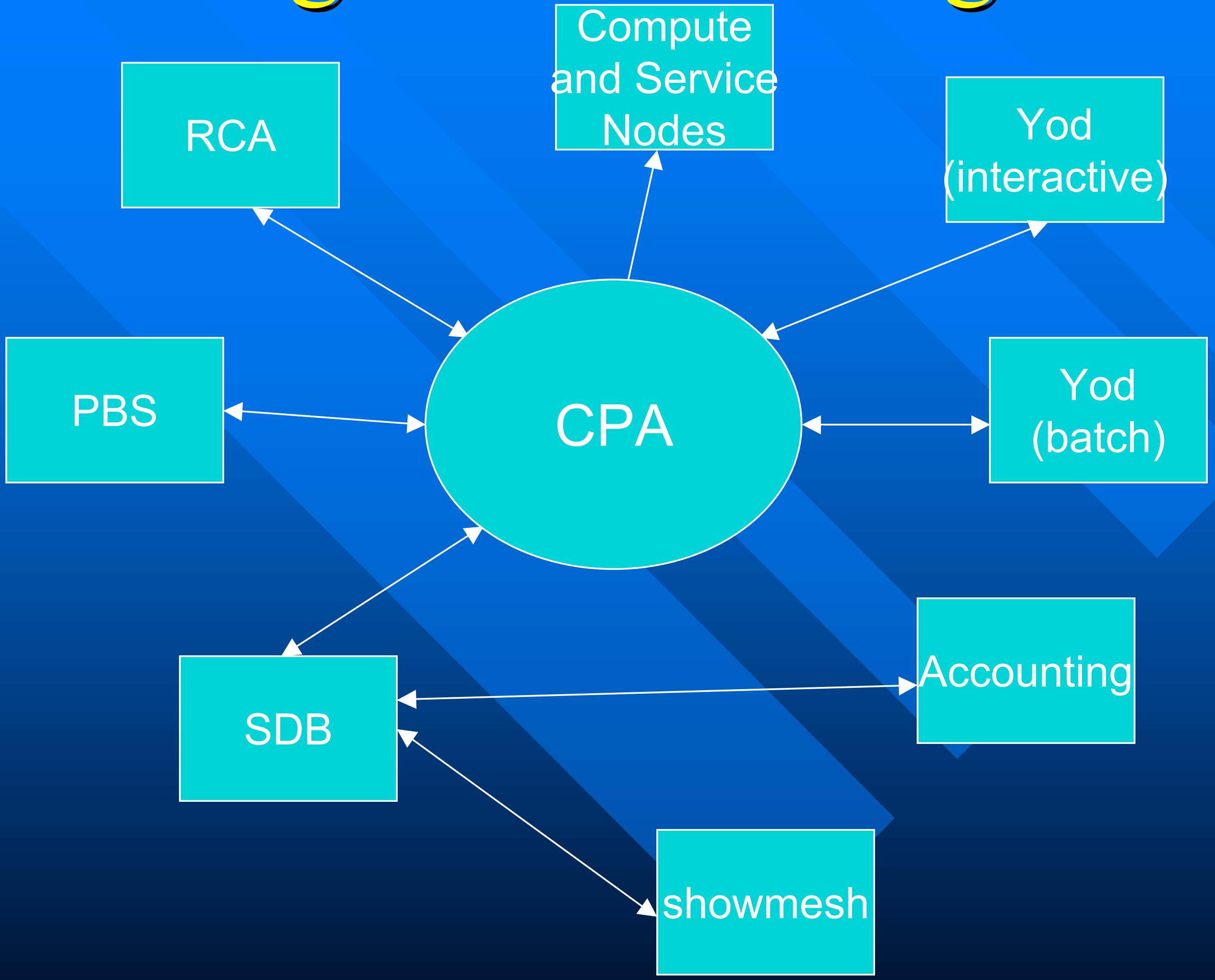
On to CRMS…

- Installed late April
- Needed to integrate harness scheduling setup with CRMS
- CPA present
- SDB present
- Booting handled differently

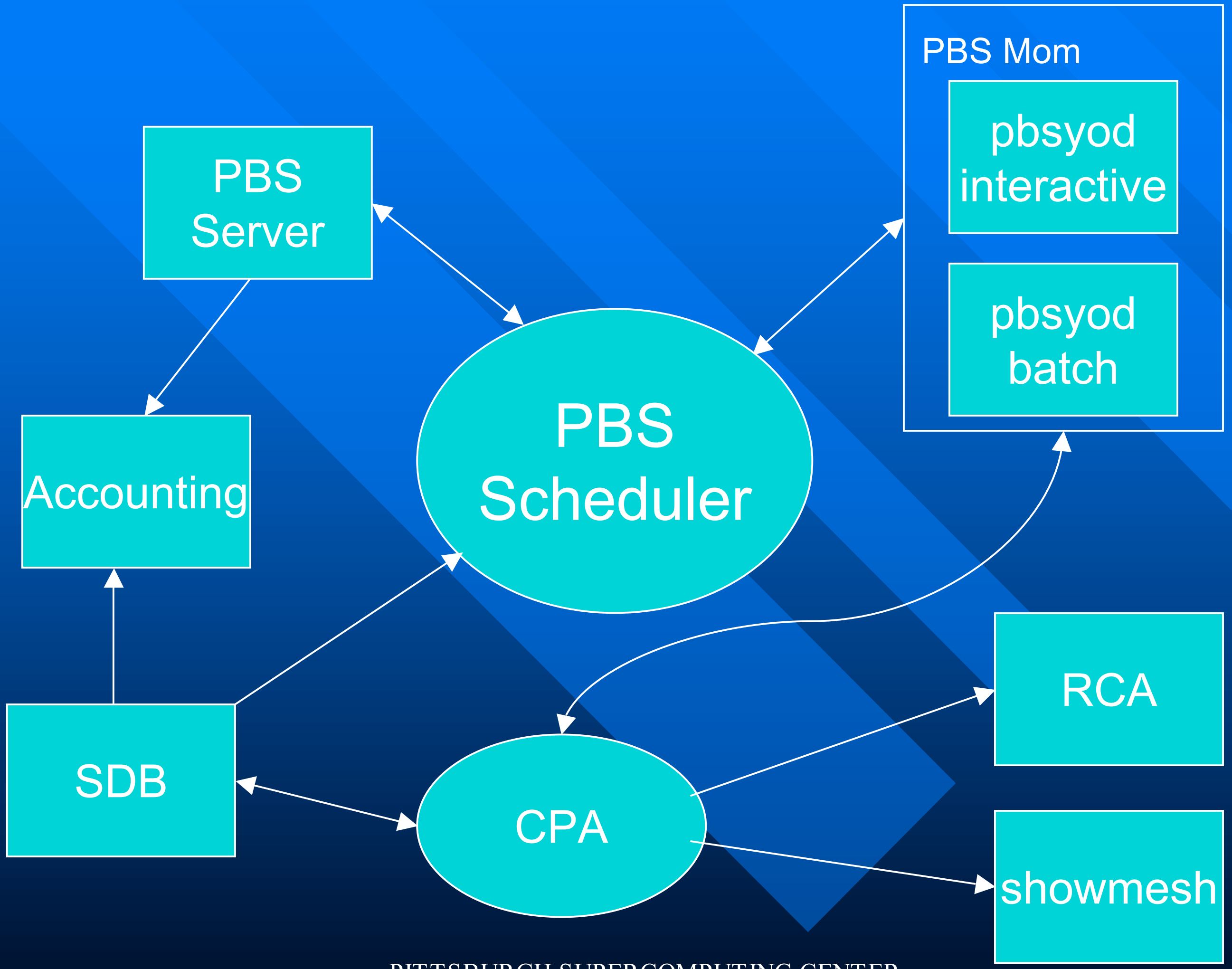
PBS-CPA Relationship



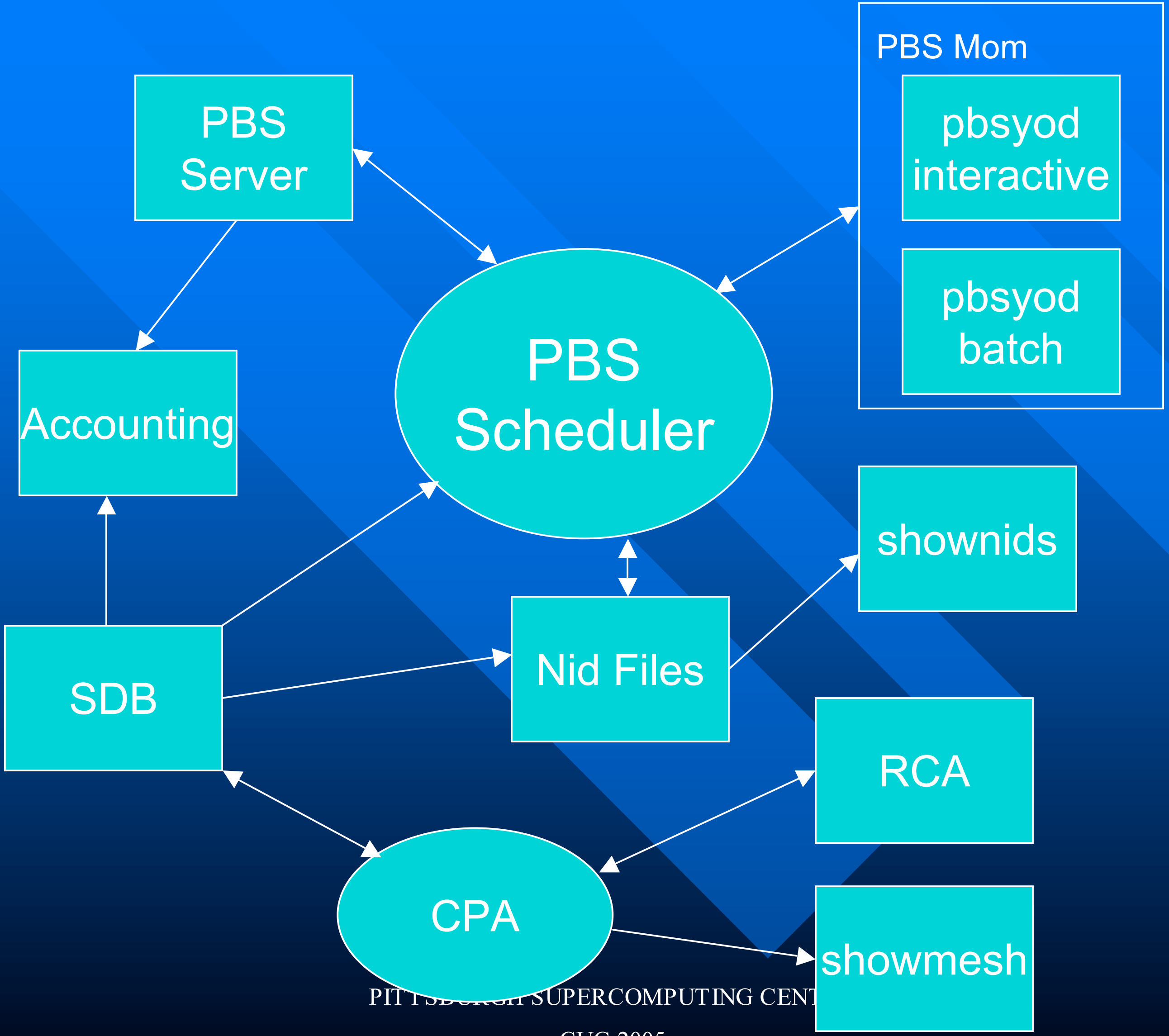
Original XT3 Design



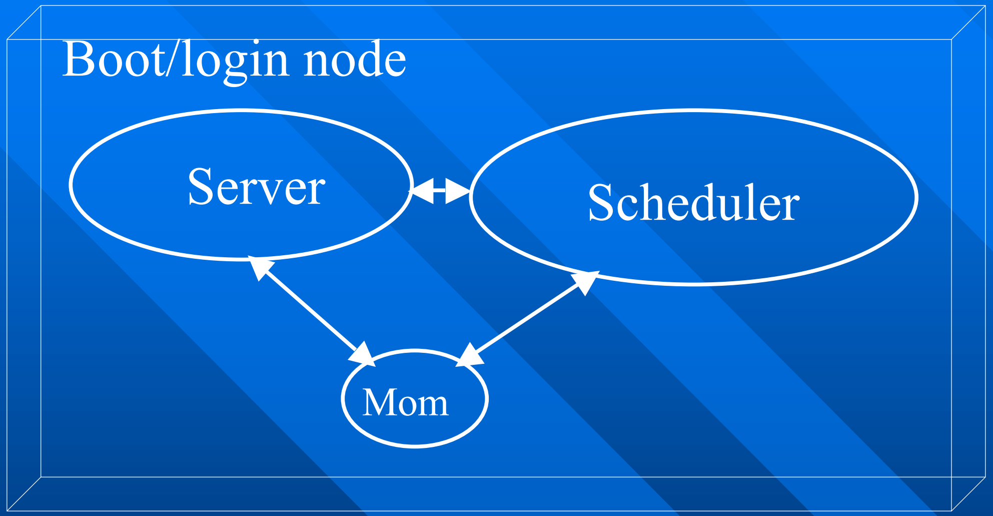
Goal


PBS Mom
pbsyod interactive
pbsyod batch
PBS Server
PBS Scheduler
Accounting
SDB
CPA
RCA
showmesh

CRMS Phase 1


PBS Mom
pbsyod interactive
pbsyod batch
PBS Server
PBS Scheduler
Accounting
shownids
SDB
Nid Files
RCA
CPA
showmesh

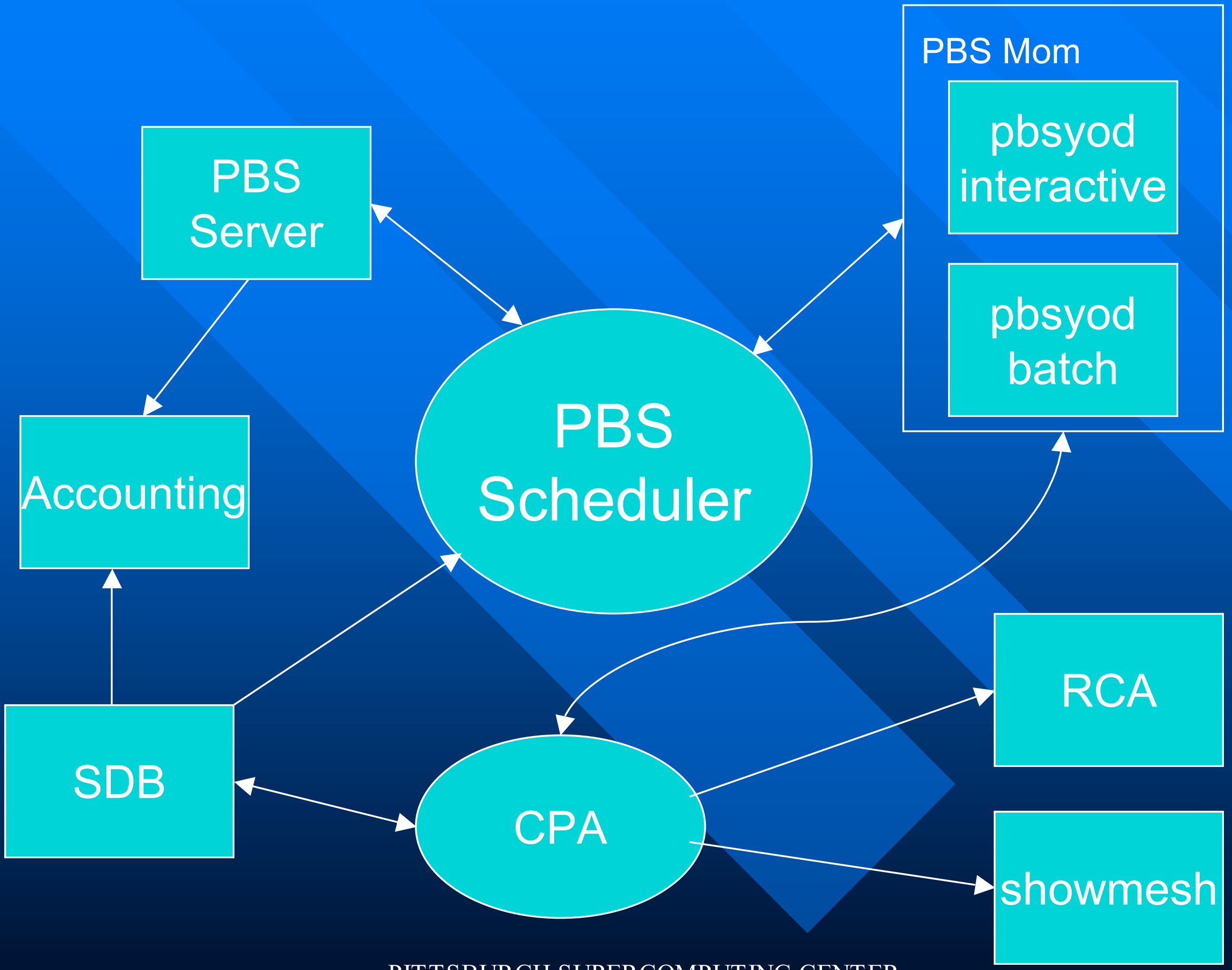
Initial CRMS PBS Configuration



Integration with CRMS

- At phase one with Torque
- Use "interactive" mode for processors
  - processor table in SDB
  - YOD_STANDALONE to bypass CPA
    - » pbsyod
- Be gentle with SDB reads
  - Scheduler synchronizes itself with processor table every 5 minutes
- Logging tools

CRMS Phase 2


PBS Mom
pbsyod interactive
pbsyod batch
PBS Server
Accounting
PBS Scheduler
SDB
CPA
RCA
showmesh

Move to Phase 2 (in test)

- Use PBS Pro (instead of Torque)
  - Build pbs_sched with TCL interpreter
  - New resources
  - Mom adaptation
    - CPA calls
- Adapt scheduler to query SDB directly
  - Most of the other code will stay the same

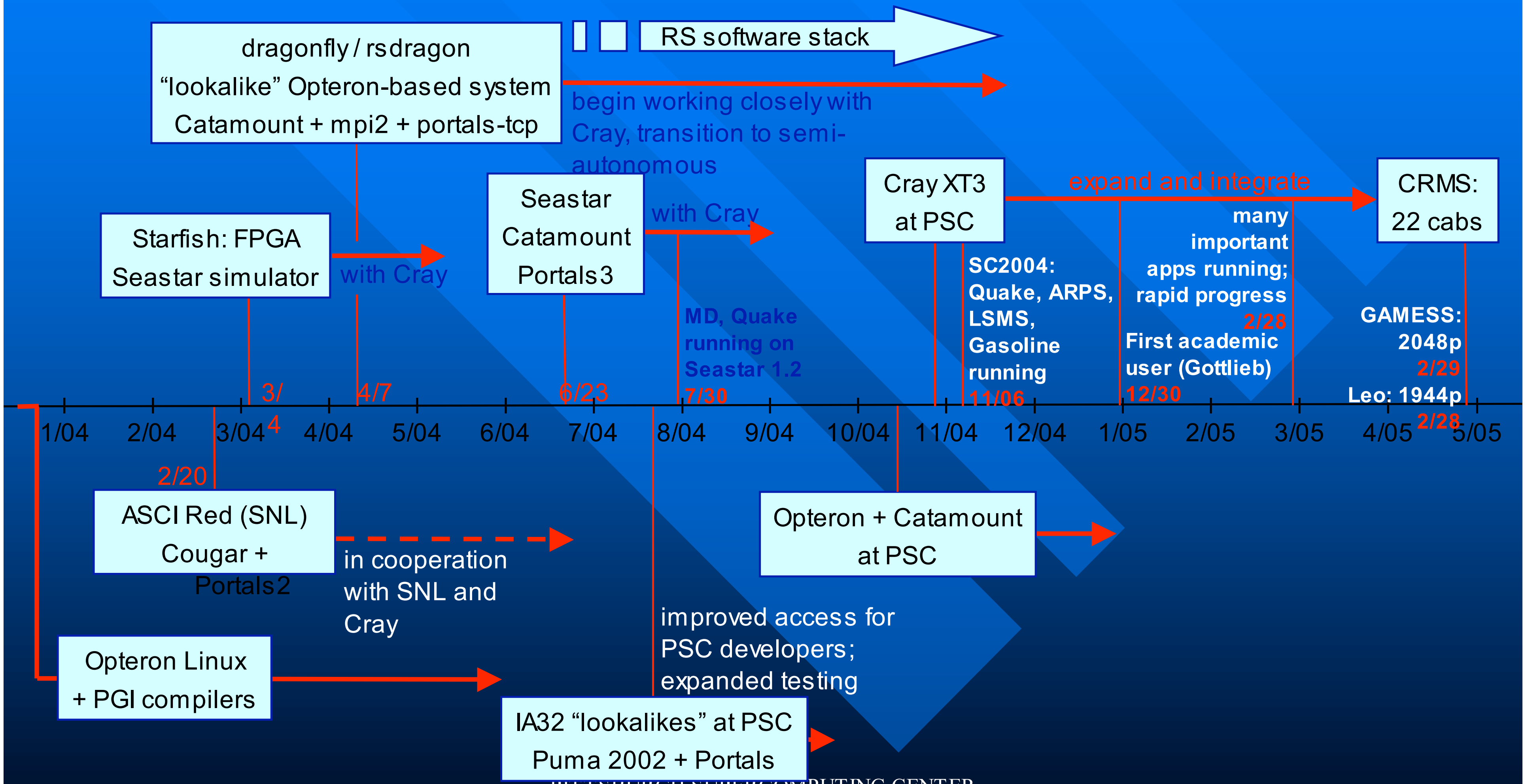
Progress on Phase 2

- Build PRO pbs_sched with TCL interpreter - Done
- New resources to PBS Pro - Done
  - nid_list
  - linux_nid_list
  - Nidmask
- Changes to pbs_mom - Done
  - Read nid_list assigned by scheduler
  - Pass to CPA
- Changes to harness scheduler – In test

Future

- More defensive node health checking
- Add more moms and schedule to these
- Integrate more features from Simon
- Investigate node allocation algorithms
  - No longer in CPA
  - Scheduler modules
  - Application specific
- Recode in Python
- Co-schedule service nodes
  - Viz
  - Data, etc.

Successes facilitated by Batch System



dragonfly / rsdragon
“lookalike” Opteron-based system
Catamount + mpi2 + portals-tcp

RS software stack

begin working closely with Cray, transition to semi-autonomous

Starfish: FPGA
Seastar simulator

with Cray

Seastar
Catamount
Portals3

with Cray

MD, Quake running on Seastar 1.2
7/30

Cray XT3
at PSC

expand and integrate

CRMS:
22 cabs

SC2004:
Quake, ARPS, LSMS, Gasoline running
11/06

many important apps running; rapid progress
2/28

First academic user (Gottlieb)
12/30

GAMESS:
2048p
2/29

Leo: 1944p
2/28

3/4

4/7

6/23

1/04 2/04 3/04 4/04 5/04 6/04 7/04 8/04 9/04 10/04 11/04 12/04 1/05 2/05 3/05 4/05 5/05

2/20

ASCI Red (SNL)
Cougar +
Portals2

in cooperation with SNL and Cray

Opteron + Catamount
at PSC

improved access for PSC developers; expanded testing

Opteron Linux
+ PGI compilers

IA32 “lookalikes” at PSC
Puma 2002 + Portals

Other PSC XT3 Talks

- *Early Applications Experience on the Cray XT3*
  - – Nick Nystrom, 2:30pm Today, Taos
- *Integrating External Storage Servers with the XT3*
  - – Jason Sommerfield, 2pm Thursday, Taos