MONITORING AT THE RCF LINUX FARMS

Large-Scale Cluster Computing Workshop

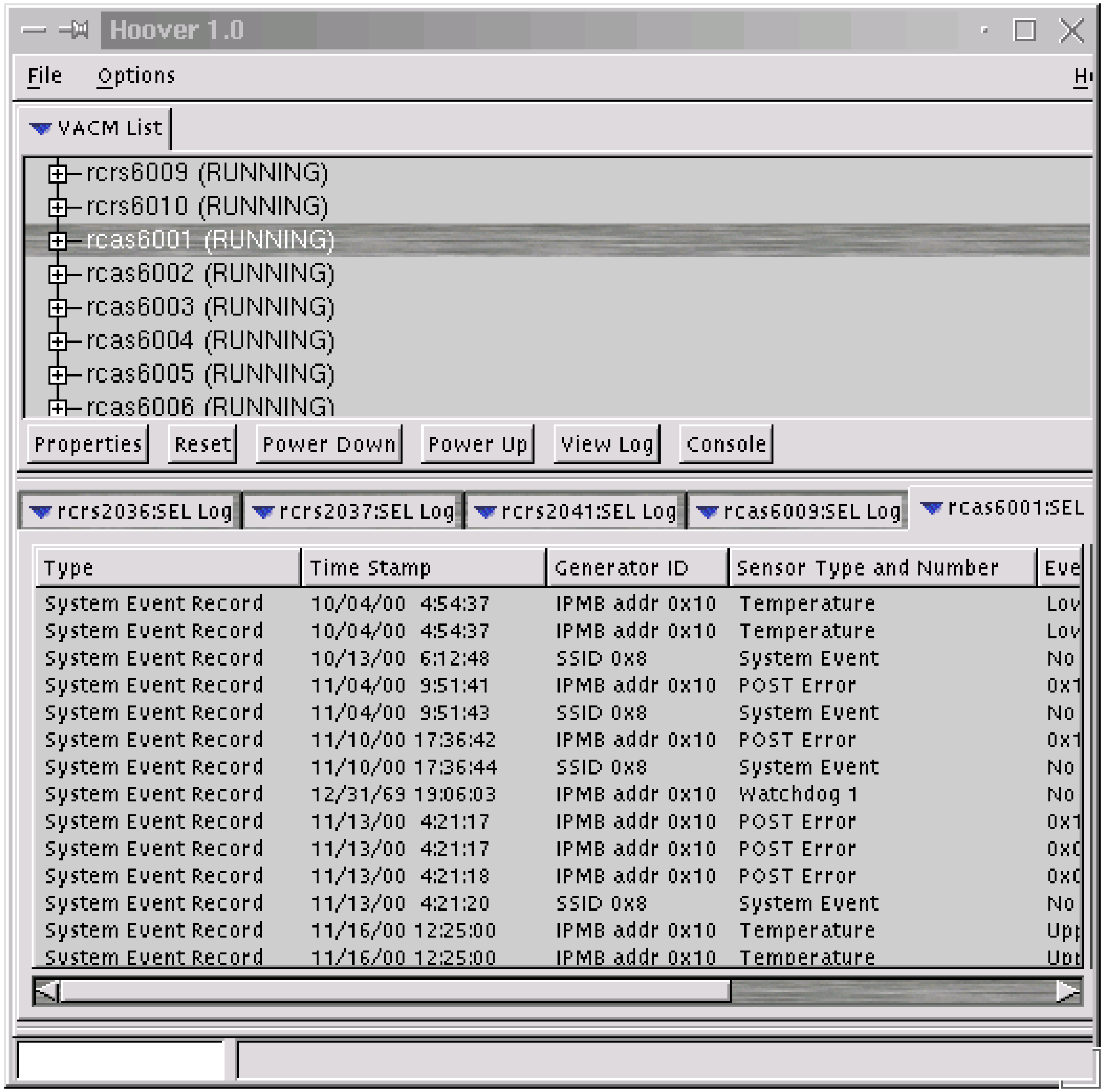
May 24, 2001

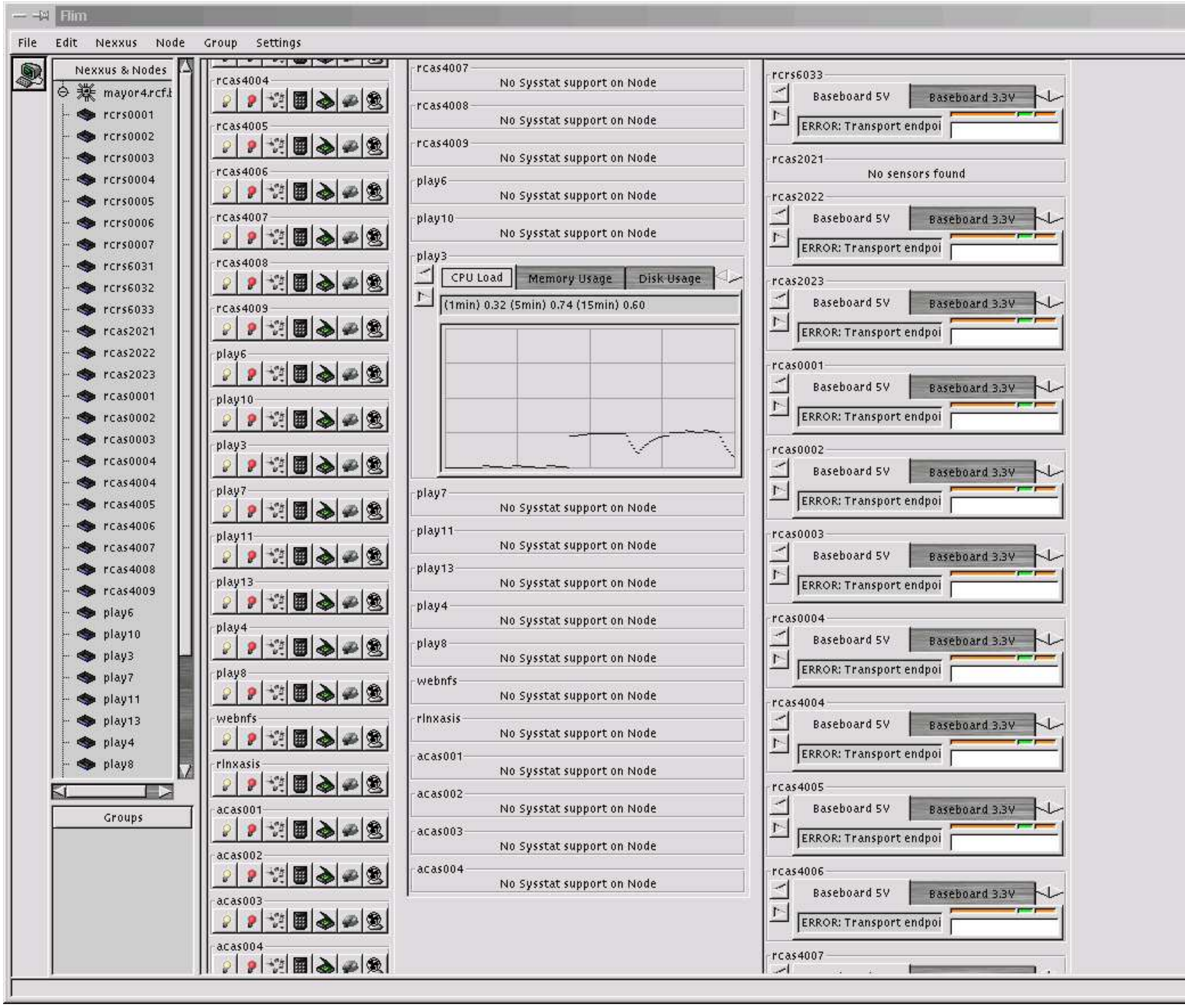
Tony Chan

BROOKHAVEN NATIONAL LABORATORY

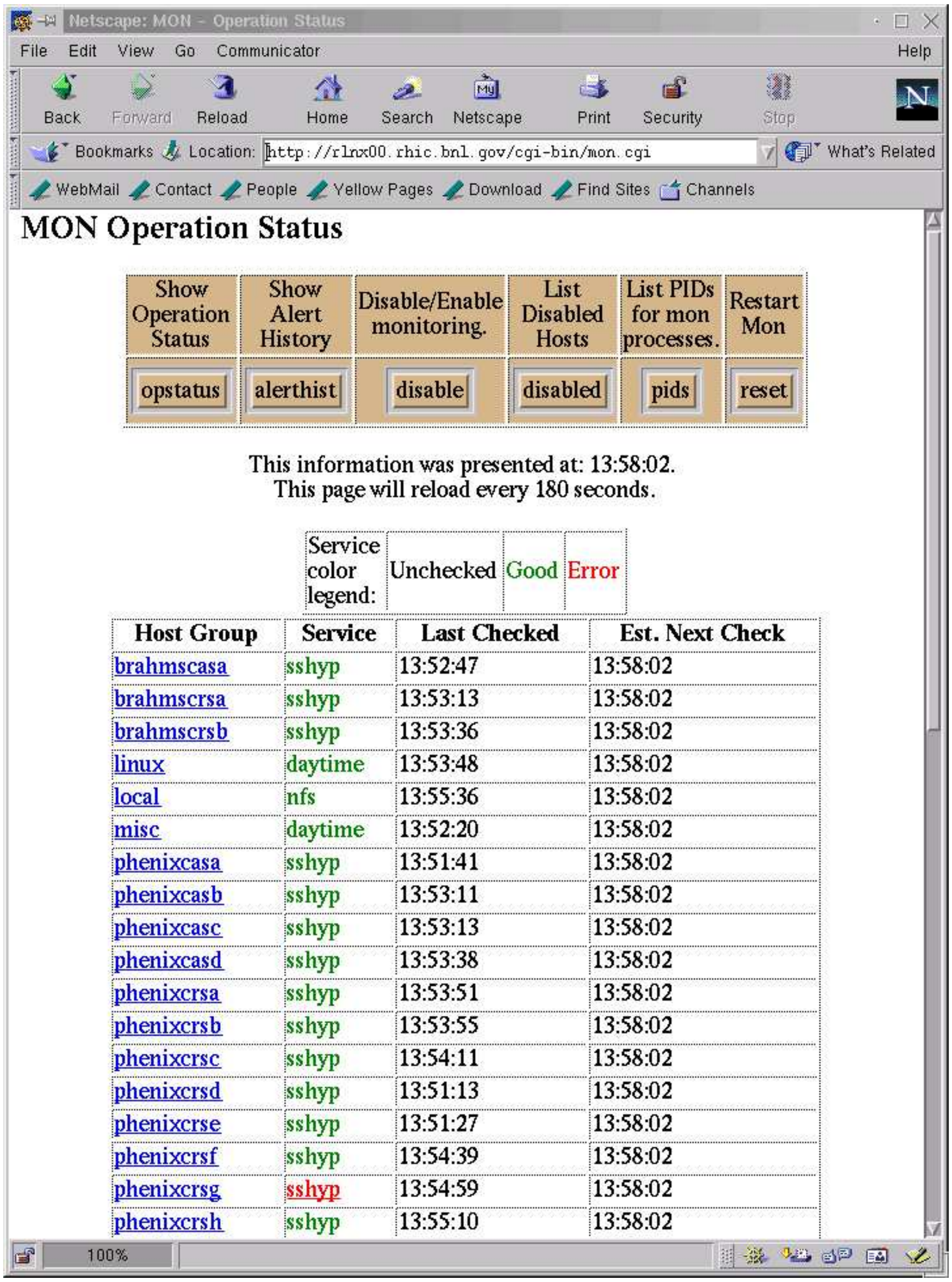
OUTLINE

- System hardware
- System software
    1. Operating System
    2. Databases
- Services available to users
    1. AFS
    2. NFS
    3. LSF
- Infrastructure
    1. Power (UPS)
    2. Cooling
    3. Others
- Alert mechanism
    1. Beeper
    2. E-mail


Hoover 1.0
File
Options
VACM List
rcrs6009 (RUNNING)
rcrs6010 (RUNNING)
rcas6001 (RUNNING)
rcas6002 (RUNNING)
rcas6003 (RUNNING)
rcas6004 (RUNNING)
rcas6005 (RUNNING)
rcas6006 (RUNNING)
Properties
Reset
Power Down
Power Up
View Log
Console
rcrs2036:SEL Log
rcrs2037:SEL Log
rcrs2041:SEL Log
rcas6009:SEL Log
rcas6001:SEL
Type
Time Stamp
Generator ID
Sensor Type and Number
System Event Record 10/04/00 4:54:37 IPMB addr 0x10 Temperature
System Event Record 10/04/00 4:54:37 IPMB addr 0x10 Temperature
System Event Record 10/13/00 6:12:48 SSID 0x8 System Event
System Event Record 11/04/00 9:51:41 IPMB addr 0x10 POST Error
System Event Record 11/04/00 9:51:43 SSID 0x8 System Event
System Event Record 11/10/00 17:36:42 IPMB addr 0x10 POST Error
System Event Record 11/10/00 17:36:44 SSID 0x8 System Event
System Event Record 12/31/69 19:06:03 IPMB addr 0x10 Watchdog 1
System Event Record 11/13/00 4:21:17 IPMB addr 0x10 POST Error
System Event Record 11/13/00 4:21:17 IPMB addr 0x10 POST Error
System Event Record 11/13/00 4:21:18 IPMB addr 0x10 POST Error
System Event Record 11/13/00 4:21:20 SSID 0x8 System Event
System Event Record 11/16/00 12:25:00 IPMB addr 0x10 Temperature
System Event Record 11/16/00 12:25:00 IPMB addr 0x10 Temperature



Flim
File
Edit
Nexxus
Node
Group
Settings
Nexxus & Nodes
mayor4.rcf.
rcrs0001
rcrs0002
rcrs0003
rcrs0004
rcrs0005
rcrs0006
rcrs0007
rcrs6031
rcrs6032
rcrs6033
rcas2021
rcas2022
rcas2023
rcas0001
rcas0002
rcas0003
rcas0004
rcas4004
rcas4005
rcas4006
rcas4007
rcas4008
rcas4009
play6
play10
play3
play7
play11
play13
play4
play8
Groups
rcas4004
rcas4005
rcas4006
rcas4007
rcas4008
rcas4009
play6
play10
play3
play7
play11
play13
play4
play8
webnfs
rlnxasis
acas001
acas002
acas003
acas004
rcas4007
No Sysstat support on Node
rcas4008
No Sysstat support on Node
rcas4009
No Sysstat support on Node
play6
No Sysstat support on Node
play10
No Sysstat support on Node
play3
CPU Load
Memory Usage
Disk Usage
(1min) 0.32 (5min) 0.74 (15min) 0.60
play7
No Sysstat support on Node
play11
No Sysstat support on Node
play13
No Sysstat support on Node
play4
No Sysstat support on Node
play8
No Sysstat support on Node
webnfs
No Sysstat support on Node
rlnxasis
No Sysstat support on Node
acas001
No Sysstat support on Node
acas002
No Sysstat support on Node
acas003
No Sysstat support on Node
acas004
No Sysstat support on Node
rcrs6033
Baseboard 5V
Baseboard 3.3V
ERROR: Transport endpoi
rcas2021
No sensors found
rcas2022
Baseboard 5V
Baseboard 3.3V
ERROR: Transport endpoi
rcas2023
Baseboard 5V
Baseboard 3.3V
ERROR: Transport endpoi
rcas0001
Baseboard 5V
Baseboard 3.3V
ERROR: Transport endpoi
rcas0002
Baseboard 5V
Baseboard 3.3V
ERROR: Transport endpoi
rcas0003
Baseboard 5V
Baseboard 3.3V
ERROR: Transport endpoi
rcas0004
Baseboard 5V
Baseboard 3.3V
ERROR: Transport endpoi
rcas4004
Baseboard 5V
Baseboard 3.3V
ERROR: Transport endpoi
rcas4005
Baseboard 5V
Baseboard 3.3V
ERROR: Transport endpoi
rcas4006
Baseboard 5V
Baseboard 3.3V
ERROR: Transport endpoi
rcas4007



Netscape: MON - Operation Status

File Edit View Go Communicator Help

Back Forward Reload Home Search Netscape Print Security Stop

Bookmarks Location: http://rlnx00.rhic.bnl.gov/cgi-bin/mon.cgi What's Related

WebMail Contact People Yellow Pages Download Find Sites Channels

MON Operation Status

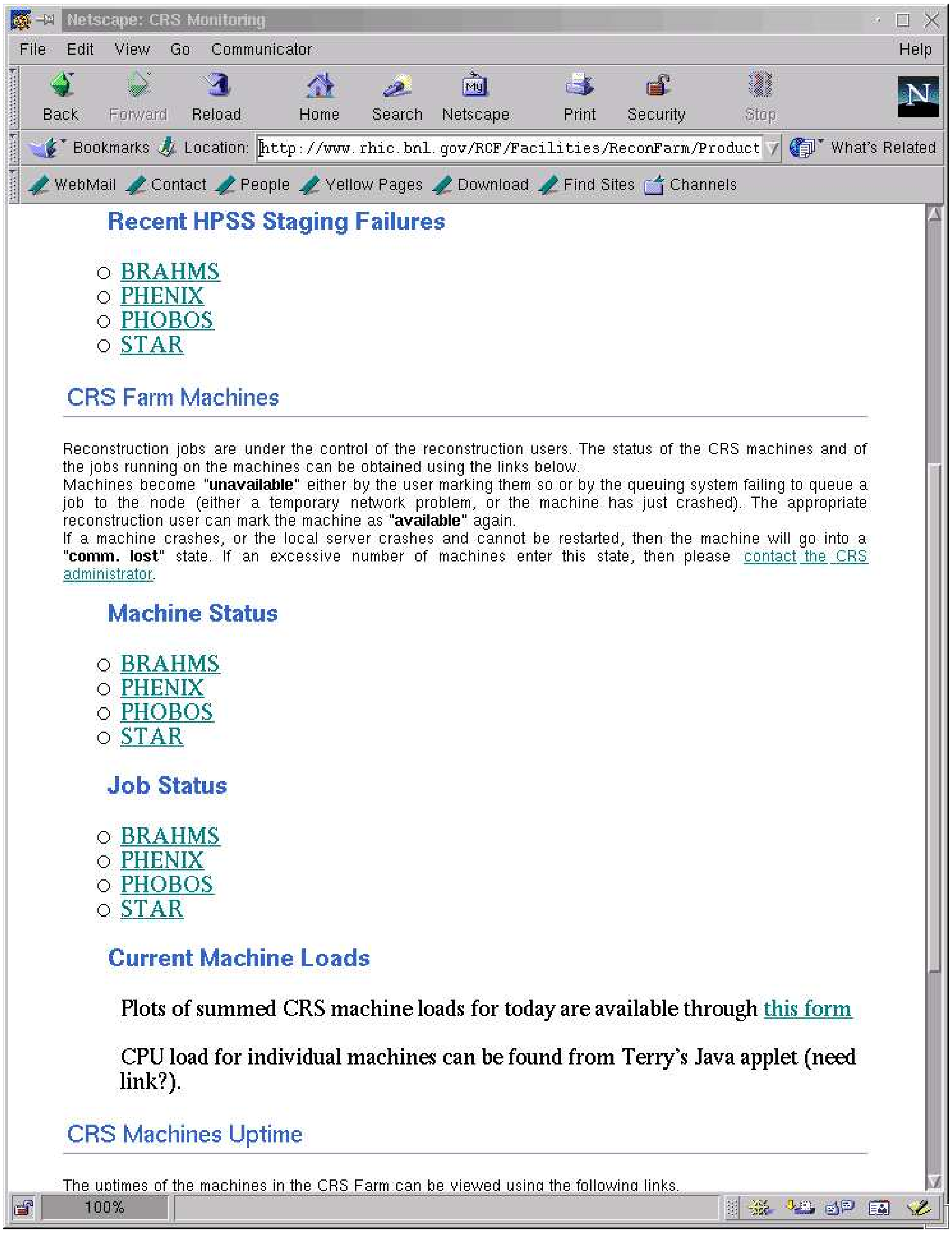
Show Operation Status	Show Alert History	Disable/Enable monitoring.	List Disabled Hosts	List PIDs for mon processes.	Restart Mon
opstatus	alerthist	disable	disabled	pids	reset

This information was presented at: 13:58:02.
This page will reload every 180 seconds.

Service color legend:	Unchecked	Good	Error

Host Group	Service	Last Checked	Est. Next Check
brahmscasa	sshyp	13:52:47	13:58:02
brahmscrsa	sshyp	13:53:13	13:58:02
brahmscrsb	sshyp	13:53:36	13:58:02
linux	daytime	13:53:48	13:58:02
local	nfs	13:55:36	13:58:02
misc	daytime	13:52:20	13:58:02
phenixcasa	sshyp	13:51:41	13:58:02
phenixcasb	sshyp	13:53:11	13:58:02
phenixcasc	sshyp	13:53:13	13:58:02
phenixcasd	sshyp	13:53:38	13:58:02
phenixcrsa	sshyp	13:53:51	13:58:02
phenixcrsb	sshyp	13:53:55	13:58:02
phenixcrsc	sshyp	13:54:11	13:58:02
phenixcrsd	sshyp	13:51:13	13:58:02
phenixcrse	sshyp	13:51:27	13:58:02
phenixcrsf	sshyp	13:54:39	13:58:02
phenixcrsg	sshyp	13:54:59	13:58:02
phenixcrsh	sshyp	13:55:10	13:58:02

100%



Recent HPSS Staging Failures

- BRAHMS
- PHENIX
- PHOBOS
- STAR

CRS Farm Machines

Reconstruction jobs are under the control of the reconstruction users. The status of the CRS machines and of the jobs running on the machines can be obtained using the links below.
Machines become "**unavailable**" either by the user marking them so or by the queuing system failing to queue a job to the node (either a temporary network problem, or the machine has just crashed). The appropriate reconstruction user can mark the machine as "**available**" again.
If a machine crashes, or the local server crashes and cannot be restarted, then the machine will go into a "**comm. lost**" state. If an excessive number of machines enter this state, then please contact the CRS administrator.

Machine Status

- BRAHMS
- PHENIX
- PHOBOS
- STAR

Job Status

- BRAHMS
- PHENIX
- PHOBOS
- STAR

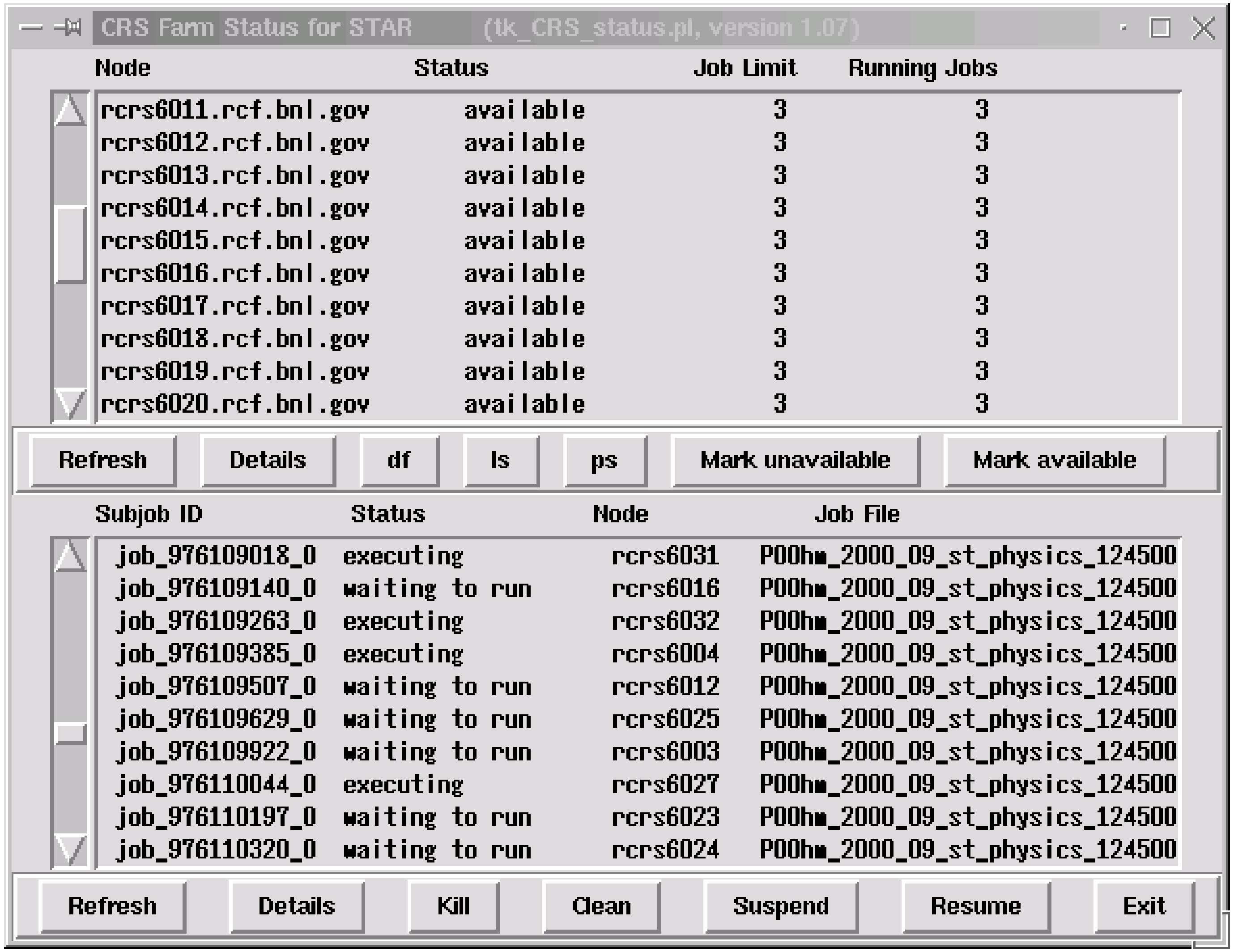
Current Machine Loads

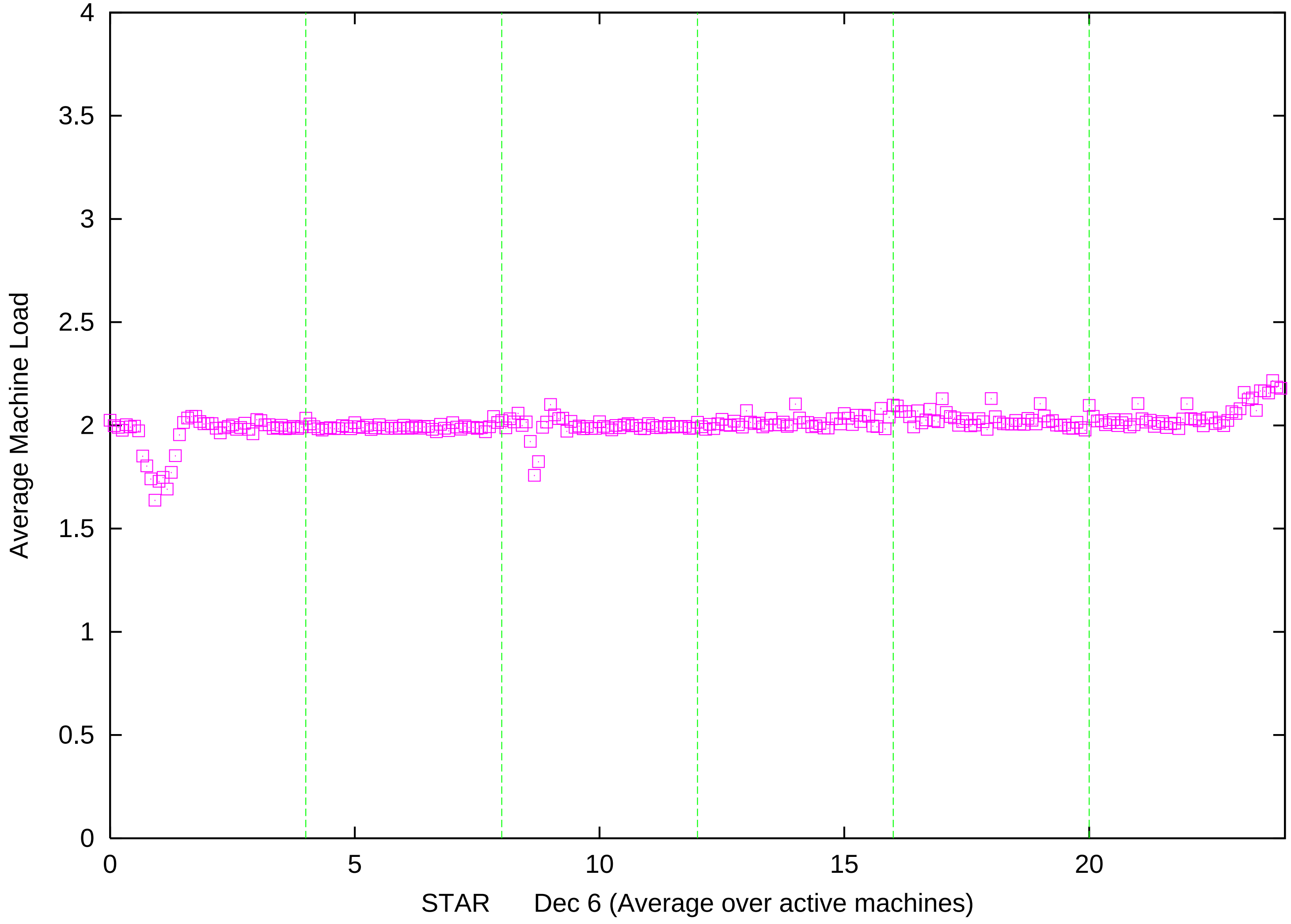
Plots of summed CRS machine loads for today are available through this form

CPU load for individual machines can be found from Terry's Java applet (need link?).

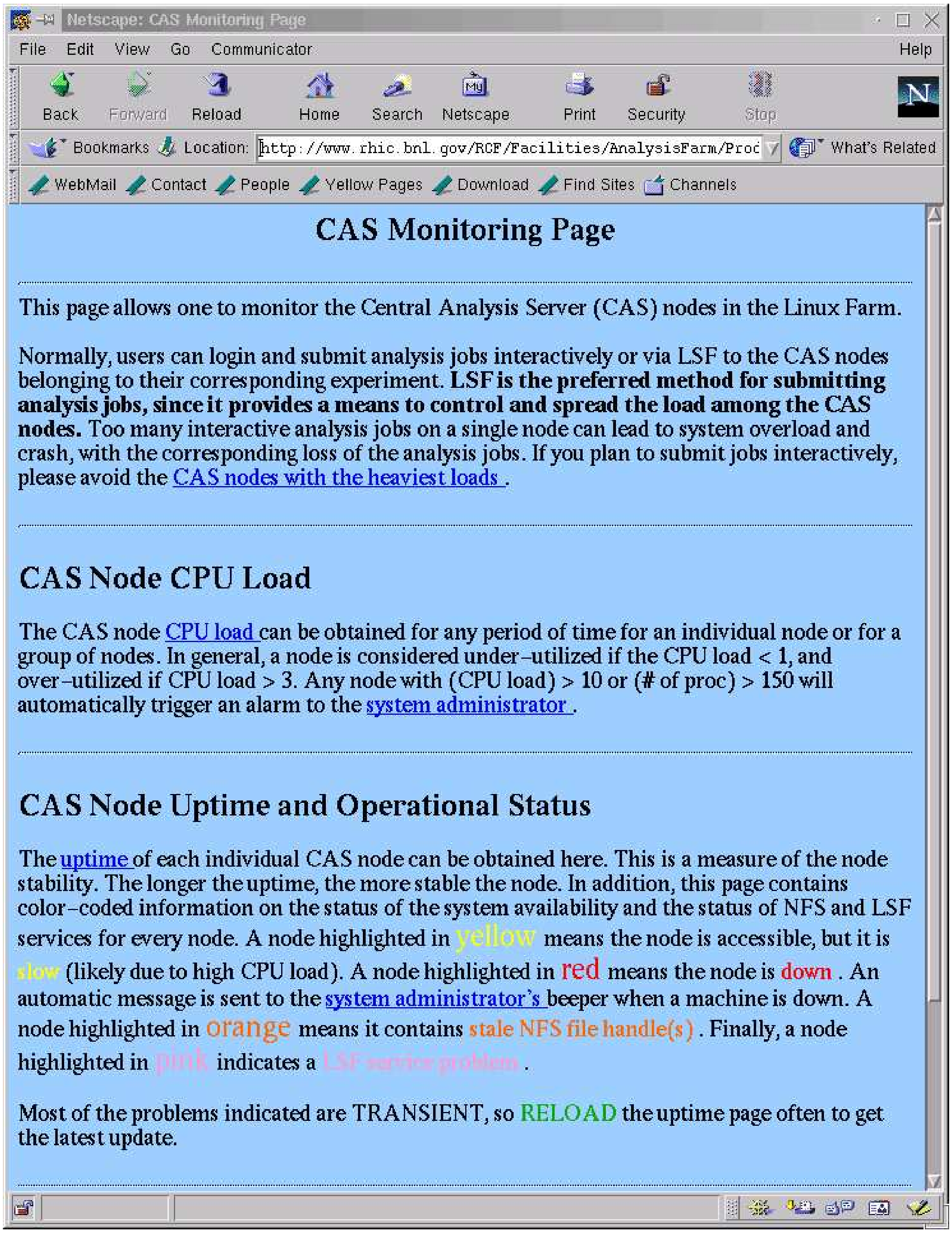
CRS Machines Uptime

The uptimes of the machines in the CRS Farm can be viewed using the following links.


CRS Farm Status for STAR (tk_CRS_status.pl, version 1.07)
Node Status Job Limit Running Jobs
rcrs6011.rcf.bnl.gov available 3 3
rcrs6012.rcf.bnl.gov available 3 3
rcrs6013.rcf.bnl.gov available 3 3
rcrs6014.rcf.bnl.gov available 3 3
rcrs6015.rcf.bnl.gov available 3 3
rcrs6016.rcf.bnl.gov available 3 3
rcrs6017.rcf.bnl.gov available 3 3
rcrs6018.rcf.bnl.gov available 3 3
rcrs6019.rcf.bnl.gov available 3 3
rcrs6020.rcf.bnl.gov available 3 3
Refresh
Details
df
ls
ps
Mark unavailable
Mark available
Subjob ID Status Node Job File
job_976109018_0 executing rcrs6031 P00hm_2000_09_st_physics_124500
job_976109140_0 waiting to run rcrs6016 P00hm_2000_09_st_physics_124500
job_976109263_0 executing rcrs6032 P00hm_2000_09_st_physics_124500
job_976109385_0 executing rcrs6004 P00hm_2000_09_st_physics_124500
job_976109507_0 waiting to run rcrs6012 P00hm_2000_09_st_physics_124500
job_976109629_0 waiting to run rcrs6025 P00hm_2000_09_st_physics_124500
job_976109922_0 waiting to run rcrs6003 P00hm_2000_09_st_physics_124500
job_976110044_0 executing rcrs6027 P00hm_2000_09_st_physics_124500
job_976110197_0 waiting to run rcrs6023 P00hm_2000_09_st_physics_124500
job_976110320_0 waiting to run rcrs6024 P00hm_2000_09_st_physics_124500
Refresh
Details
Kill
Clean
Suspend
Resume
Exit



Average Machine Load
0
0.5
1
1.5
2
2.5
3
3.5
4
0
5
10
15
20
STAR Dec 6 (Average over active machines)



CAS Monitoring Page

This page allows one to monitor the Central Analysis Server (CAS) nodes in the Linux Farm.

Normally, users can login and submit analysis jobs interactively or via LSF to the CAS nodes belonging to their corresponding experiment. **LSF is the preferred method for submitting analysis jobs, since it provides a means to control and spread the load among the CAS nodes.** Too many interactive analysis jobs on a single node can lead to system overload and crash, with the corresponding loss of the analysis jobs. If you plan to submit jobs interactively, please avoid the CAS nodes with the heaviest loads .

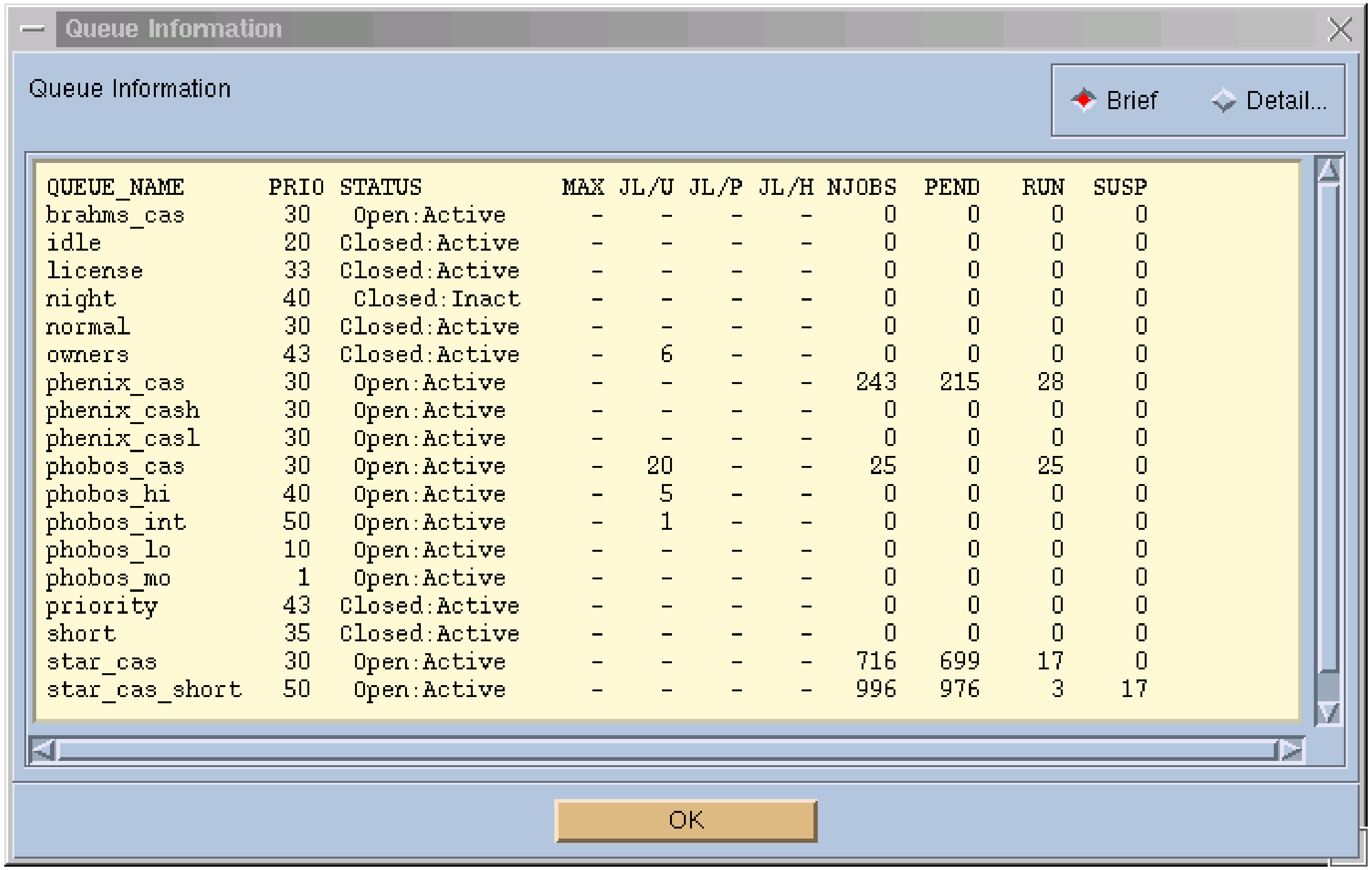
CAS Node CPU Load

The CAS node CPU load can be obtained for any period of time for an individual node or for a group of nodes. In general, a node is considered under-utilized if the CPU load < 1, and over-utilized if CPU load > 3. Any node with (CPU load) > 10 or (# of proc) > 150 will automatically trigger an alarm to the system administrator .

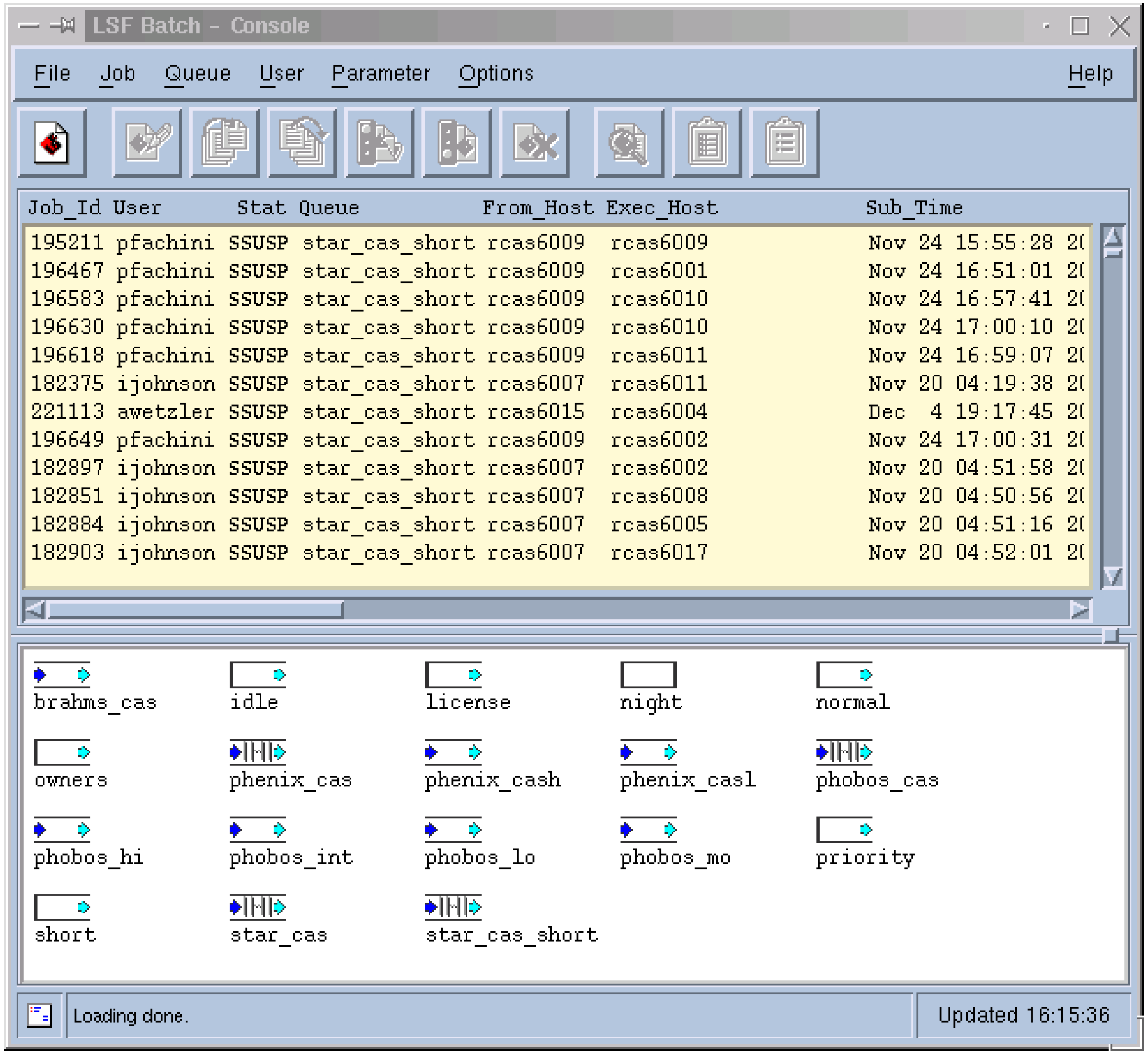
CAS Node Uptime and Operational Status

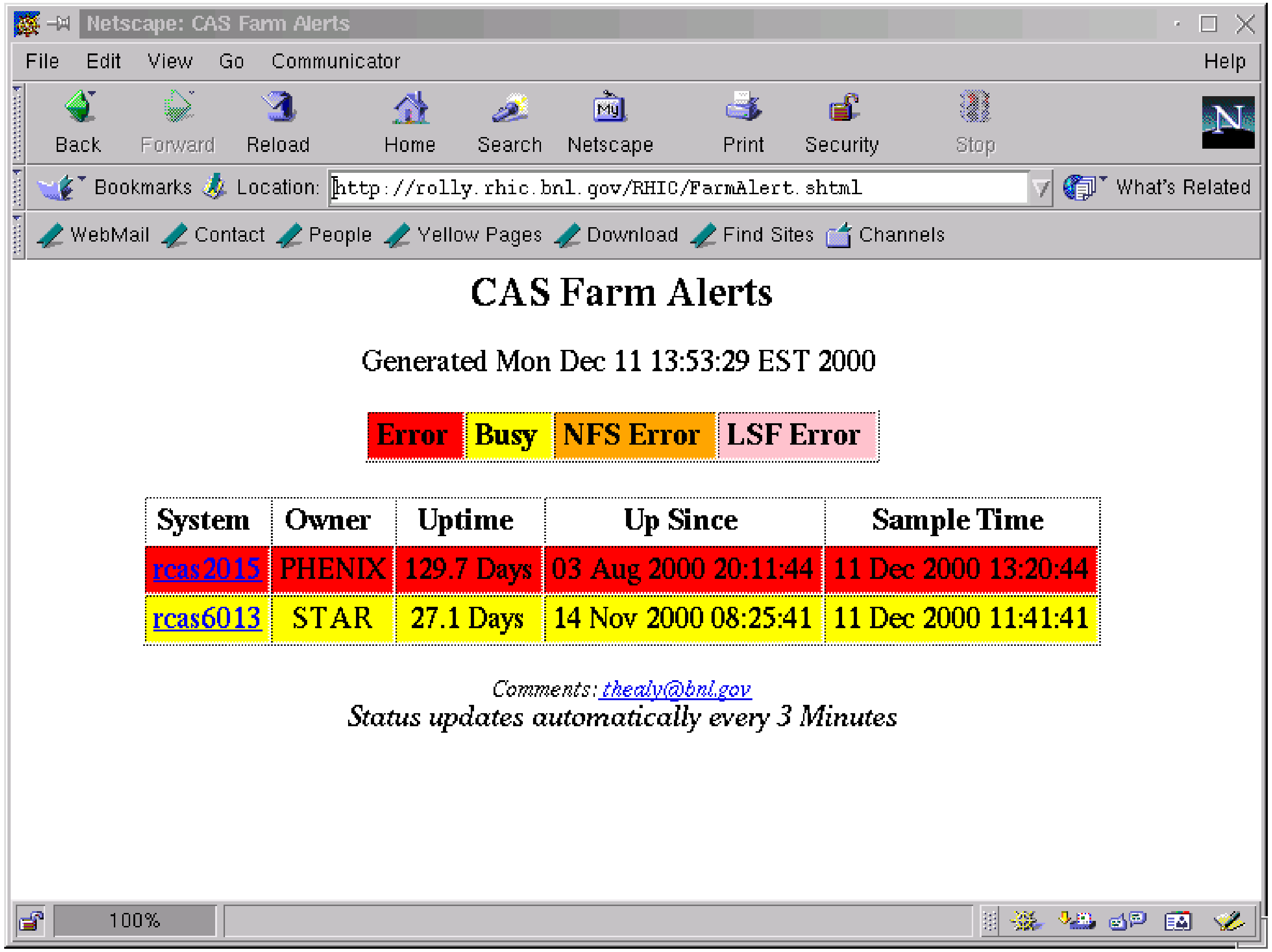
The uptime of each individual CAS node can be obtained here. This is a measure of the node stability. The longer the uptime, the more stable the node. In addition, this page contains color-coded information on the status of the system availability and the status of NFS and LSF services for every node. A node highlighted in yellow means the node is accessible, but it is slow (likely due to high CPU load). A node highlighted in red means the node is down . An automatic message is sent to the system administrator's beeper when a machine is down. A node highlighted in orange means it contains stale NFS file handle(s) . Finally, a node highlighted in pink indicates a LSF service problem .

Most of the problems indicated are TRANSIENT, so RELOAD the uptime page often to get the latest update.



QUEUE_NAME	PRIO	STATUS	MAX	JL/U	JL/P	JL/H	NJOBS	PEND	RUN	SUSP
brahms_cas	30	Open:Active	-	-	-	-	0	0	0	0
idle	20	Closed:Active	-	-	-	-	0	0	0	0
license	33	Closed:Active	-	-	-	-	0	0	0	0
night	40	Closed:Inact	-	-	-	-	0	0	0	0
normal	30	Closed:Active	-	-	-	-	0	0	0	0
owners	43	Closed:Active	-	6	-	-	0	0	0	0
phenix_cas	30	Open:Active	-	-	-	-	243	215	28	0
phenix_cash	30	Open:Active	-	-	-	-	0	0	0	0
phenix_casl	30	Open:Active	-	-	-	-	0	0	0	0
phobos_cas	30	Open:Active	-	20	-	-	25	0	25	0
phobos_hi	40	Open:Active	-	5	-	-	0	0	0	0
phobos_int	50	Open:Active	-	1	-	-	0	0	0	0
phobos_lo	10	Open:Active	-	-	-	-	0	0	0	0
phobos_mo	1	Open:Active	-	-	-	-	0	0	0	0
priority	43	Closed:Active	-	-	-	-	0	0	0	0
short	35	Closed:Active	-	-	-	-	0	0	0	0
star_cas	30	Open:Active	-	-	-	-	716	699	17	0
star_cas_short	50	Open:Active	-	-	-	-	996	976	3	17


LSF Batch - Console
File Job Queue User Parameter Options Help
Job_Id User Stat Queue From_Host Exec_Host Sub_Time
195211 pfachini SSUSP star_cas_short rcas6009 rcas6009 Nov 24 15:55:28
196467 pfachini SSUSP star_cas_short rcas6009 rcas6001 Nov 24 16:51:01
196583 pfachini SSUSP star_cas_short rcas6009 rcas6010 Nov 24 16:57:41
196630 pfachini SSUSP star_cas_short rcas6009 rcas6010 Nov 24 17:00:10
196618 pfachini SSUSP star_cas_short rcas6009 rcas6011 Nov 24 16:59:07
182375 ijohnson SSUSP star_cas_short rcas6007 rcas6011 Nov 20 04:19:38
221113 awetzler SSUSP star_cas_short rcas6015 rcas6004 Dec 4 19:17:45
196649 pfachini SSUSP star_cas_short rcas6009 rcas6002 Nov 24 17:00:31
182897 ijohnson SSUSP star_cas_short rcas6007 rcas6002 Nov 20 04:51:58
182851 ijohnson SSUSP star_cas_short rcas6007 rcas6008 Nov 20 04:50:56
182884 ijohnson SSUSP star_cas_short rcas6007 rcas6005 Nov 20 04:51:16
182903 ijohnson SSUSP star_cas_short rcas6007 rcas6017 Nov 20 04:52:01
brahms_cas
idle
license
night
normal
owners
phenix_cas
phenix_cash
phenix_casl
phobos_cas
phobos_hi
phobos_int
phobos_lo
phobos_mo
priority
short
star_cas
star_cas_short
Loading done.
Updated 16:15:36


Netscape: CAS Farm Alerts
File Edit View Go Communicator Help
Back Forward Reload Home Search Netscape Print Security Stop
Bookmarks Location: http://rolly.rhic.bnl.gov/RHIC/FarmAlert.shtml
What's Related
WebMail Contact People Yellow Pages Download Find Sites Channels
CAS Farm Alerts
Generated Mon Dec 11 13:53:29 EST 2000
Error Busy NFS Error LSF Error
System Owner Uptime Up Since Sample Time
rcas2015 PHENIX 129.7 Days 03 Aug 2000 20:11:44 11 Dec 2000 13:20:44
rcas6013 STAR 27.1 Days 14 Nov 2000 08:25:41 11 Dec 2000 11:41:41
Comments: thealy@bnl.gov
Status updates automatically every 3 Minutes
100%