Graphical Monitoring and Management of Cray Supercomputers

Jim Robanske
Cray Inc.

CRAY
THE SUPERCOMPUTER COMPANY

XTGUI

- Client/Server application
- Server runs on SMW
- Client runs on SMW or any network attached computer system

Software Distribution

- The xtgui client and server software is distributed with the SMW software release. Both the xtgui server and client are normally installed on the SMW.
- The xtgui server automatically starts with the rest of the CRMS daemons. The xtgui client may be run on the SMW, or downloaded and operated from a remote system.
- To package the xtgui for remote download, run the script: /opt/cray/bin/packagextgui

XTGUI Server

- Started with other CRMS daemons
- Obtains status of all XT components from CRMS state manager
- Listens for CRMS events
- Sends each client a copy of XT component status upon client startup

XTGUI Server

- Forwards selected CRMS events to clients
- Responds to client requests
- Monitors selected log files on the SMW, sending new entries to clients
- Runs xtcli commands on behalf of clients

Log Files

- Server startup/error log:
  - /opt/craylog/xtguiserver.out
  - Contains server operational status messages.
- Server application log:
  - /opt/craylog/RsmsJServer.log.mmddyy
  - Contains copy of significant xtgui client related events, such as xtcli commands executed on behalf of clients

Properties Files

- /opt/cray/etc/RsmsJServerProperties.txt
- Contains entries to configure various server options
- Example:
  - # The buffer size (number of lines to read) used when “tailing" server
  - #log files.
  - server.watchedFileBufferSize=200
  - # The frequency with which log files are examined for new data
  - server.watchedFileLatencyMilliseconds=5000

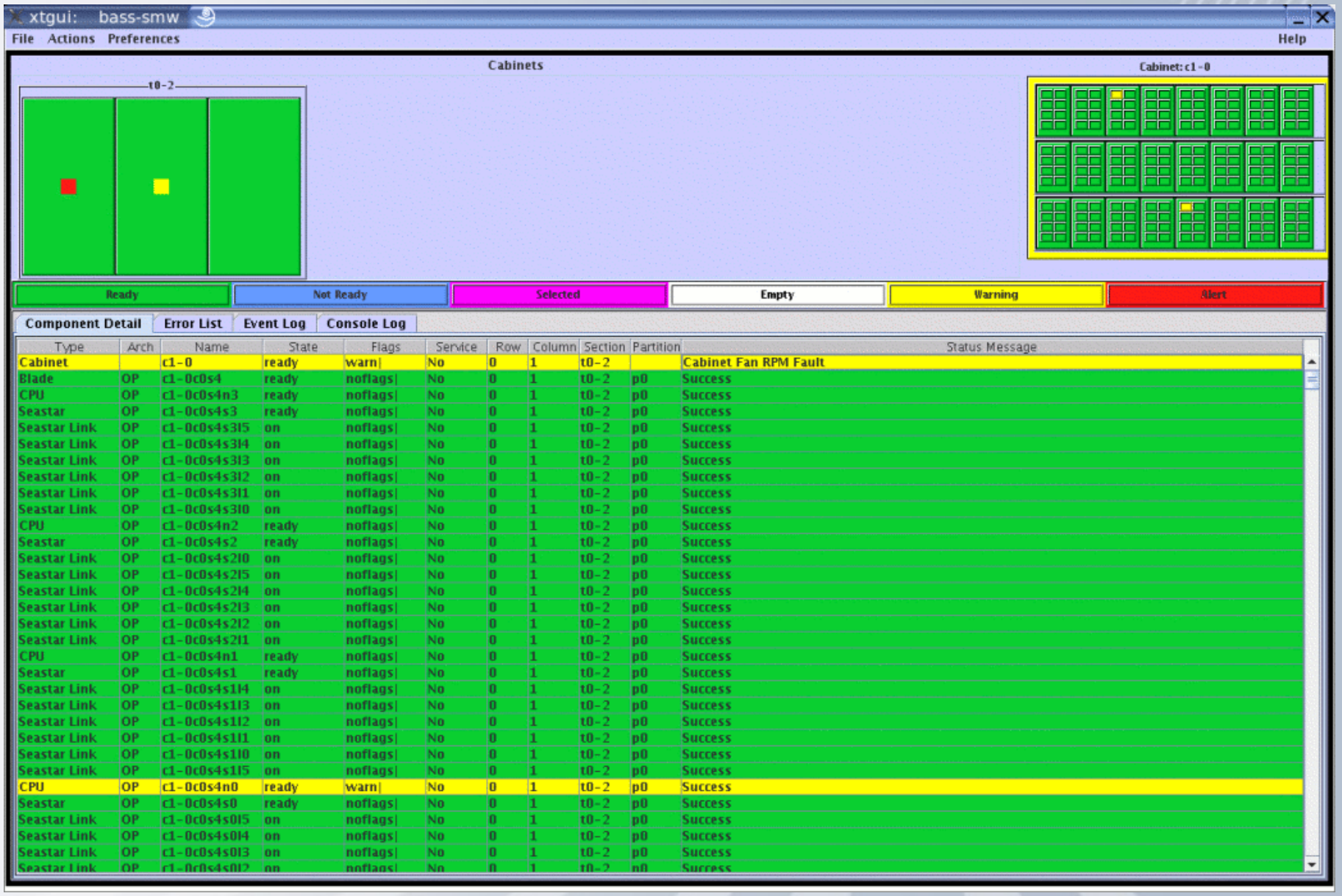
Security

- Two modes of server/client interaction:
  - Standard mode:
    - Any network attached xtgui client may connect
  - SMW only mode:
    - Connections from xtgui clients running on SMW
- Controlled by “SMWOnly” option in server startup script

Security

- Password level access control and encrypted data communications from remote sites my be accomplished by using a combination of ssh and vncserver plus vncviewer

1. Run vncserver on secure workstation or SMW
2. Ssh into secure site from remote location
3. Use vncviewer to run the xtgui on secure platform
4. Xtgui application will display on the remote system

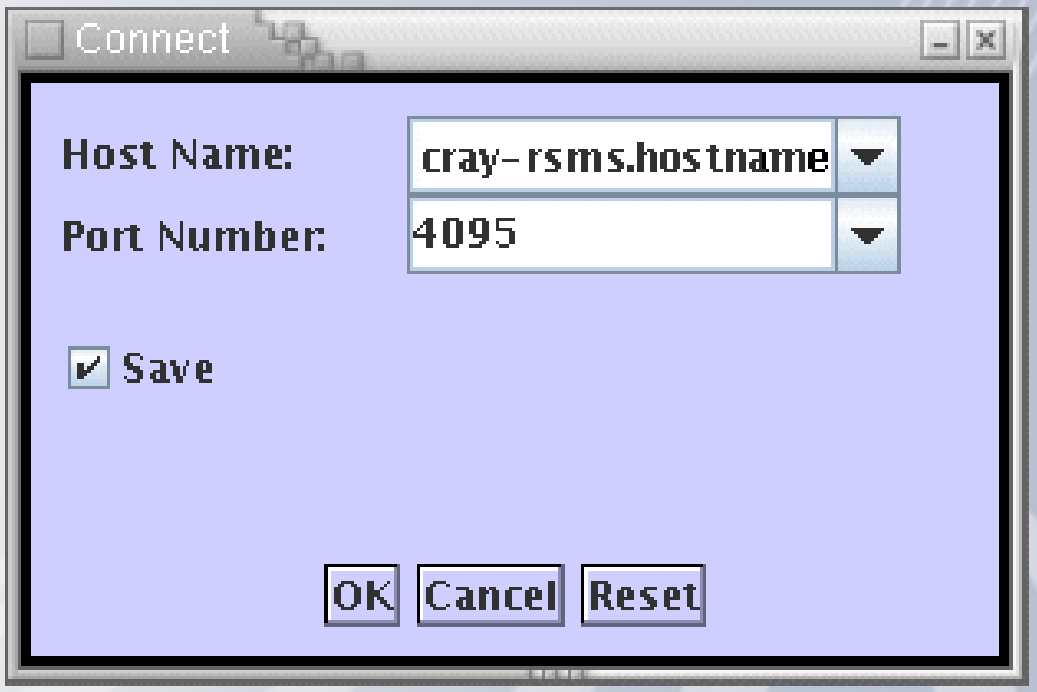

xtgui: bass-smw
File Actions Preferences
Help
Cabinets
t0-2
Cabinet: c1-0
Ready
Not Ready
Selected
Empty
Warning
Alert
Component Detail
Error List
Event Log
Console Log
Type Arch Name State Flags Service Row Column Section Partition Status Message
Cabinet c1-0 ready warn| No 0 1 t0-2 Cabinet Fan RPM Fault
Blade OP c1-0c0s4 ready noflags| No 0 1 t0-2 p0 Success
CPU OP c1-0c0s4n3 ready noflags| No 0 1 t0-2 p0 Success
Seastar OP c1-0c0s4s3 ready noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s3l5 on noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s3l4 on noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s3l3 on noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s3l2 on noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s3l1 on noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s3l0 on noflags| No 0 1 t0-2 p0 Success
CPU OP c1-0c0s4n2 ready noflags| No 0 1 t0-2 p0 Success
Seastar OP c1-0c0s4s2 ready noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s2l0 on noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s2l5 on noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s2l4 on noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s2l3 on noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s2l2 on noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s2l1 on noflags| No 0 1 t0-2 p0 Success
CPU OP c1-0c0s4n1 ready noflags| No 0 1 t0-2 p0 Success
Seastar OP c1-0c0s4s1 ready noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s1l4 on noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s1l3 on noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s1l2 on noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s1l1 on noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s1l0 on noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s1l5 on noflags| No 0 1 t0-2 p0 Success
CPU OP c1-0c0s4n0 ready warn| No 0 1 t0-2 p0 Success
Seastar OP c1-0c0s4s0 ready noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s0l5 on noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s0l4 on noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s0l3 on noflags| No 0 1 t0-2 p0 Success
Seastar Link OP c1-0c0s4s0l2 on noflags| No 0 1 t0-2 p0 Success

XTGUI Client

- Color-coded visual representation of all major XT system components, clearly illustrating their current status
- Supports all xtcli commands, except “boot”
- Runs on any workstation with network access to the SMW
- May be operated in “view only” mode, where no XT configuration actions are allowed
- Scales from small to very large systems

Starting The Application

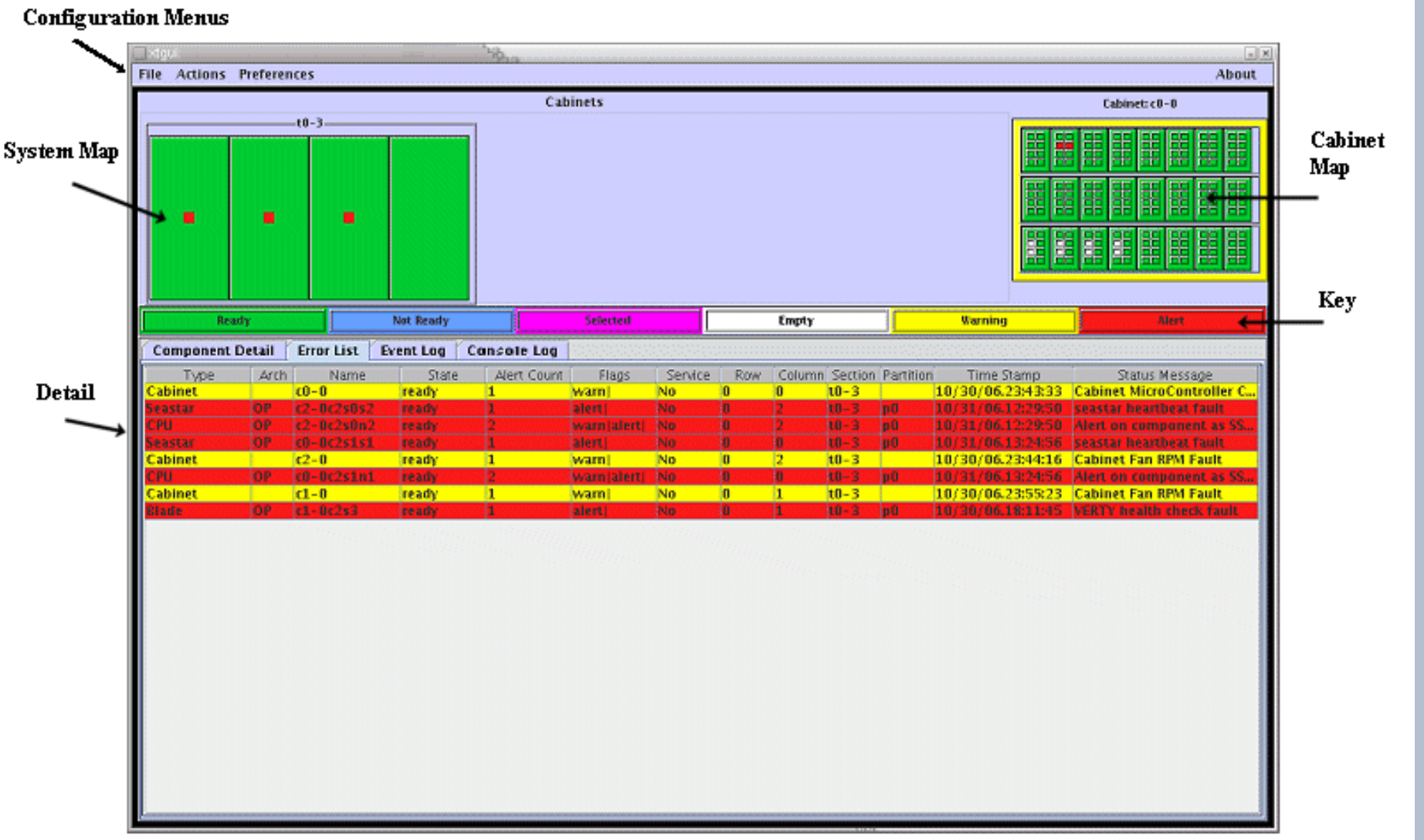
- If the application is started with a host name argument, then it will attempt to contact the xtgui server on that host, at the default port (4095)
- If no host name argument is provided, a connection screen is shown:



Six Major Views

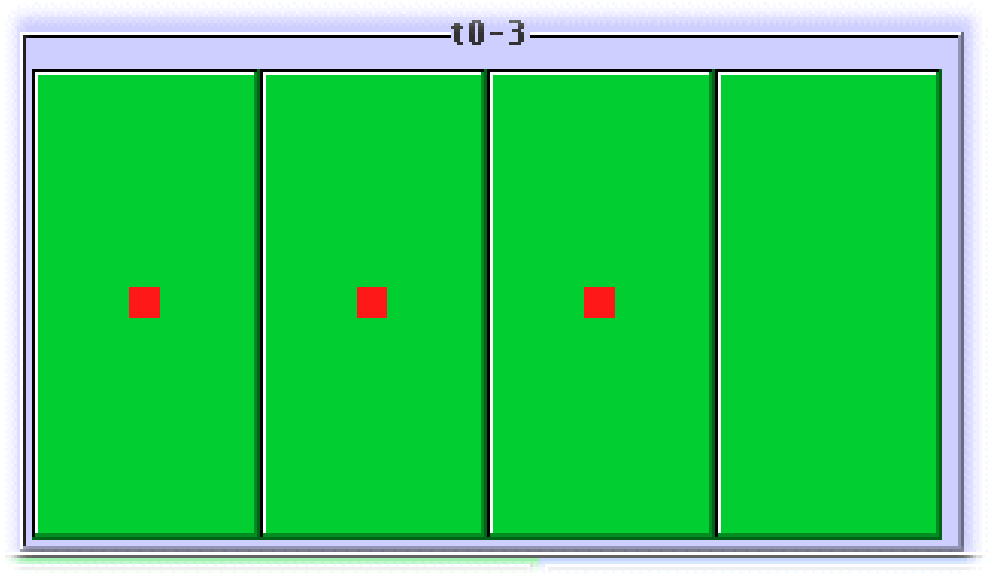
- System Map
- Cabinet Detail
- Component Detail
- Error List
- Event Log
- Console Log

Major Views


Configuration Menus
System Map
Detail
Cabinet Map
Key
File
Actions
Preferences
About
Cabinets
t0-3
Cabinet: c0-0
Ready
Not Ready
Selected
Empty
Warning
Alert
Component Detail
Error List
Event Log
Console Log
Type
Arch
Name
State
Alert Count
Flags
Service
Row
Column
Section
Partition
Time Stamp
Status Message
Cabinet c0-0 ready 1 warn| No 0 0 t0-3 10/30/06.23:43:33 Cabinet MicroController C...
Seastar OP c2-0c2s0s2 ready 1 alert| No 0 2 t0-3 p0 10/31/06.12:29:50 seastar heartbeat fault
CPU OP c2-0c2s0n2 ready 2 warn|alert| No 0 2 t0-3 p0 10/31/06.12:29:50 Alert on component as SS...
Seastar OP c0-0c2s1s1 ready 1 alert| No 0 0 t0-3 p0 10/31/06.13:24:56 seastar heartbeat fault
Cabinet c2-0 ready 1 warn| No 0 2 t0-3 10/30/06.23:44:16 Cabinet Fan RPM Fault
CPU OP c0-0c2s1n1 ready 2 warn|alert| No 0 0 t0-3 p0 10/31/06.13:24:56 Alert on component as SS...
Cabinet c1-0 ready 1 warn| No 0 1 t0-3 10/30/06.23:55:23 Cabinet Fan RPM Fault
Blade OP c1-0c2s3 ready 1 alert| No 0 1 t0-3 p0 10/30/06.18:11:45 VERTY health check fault

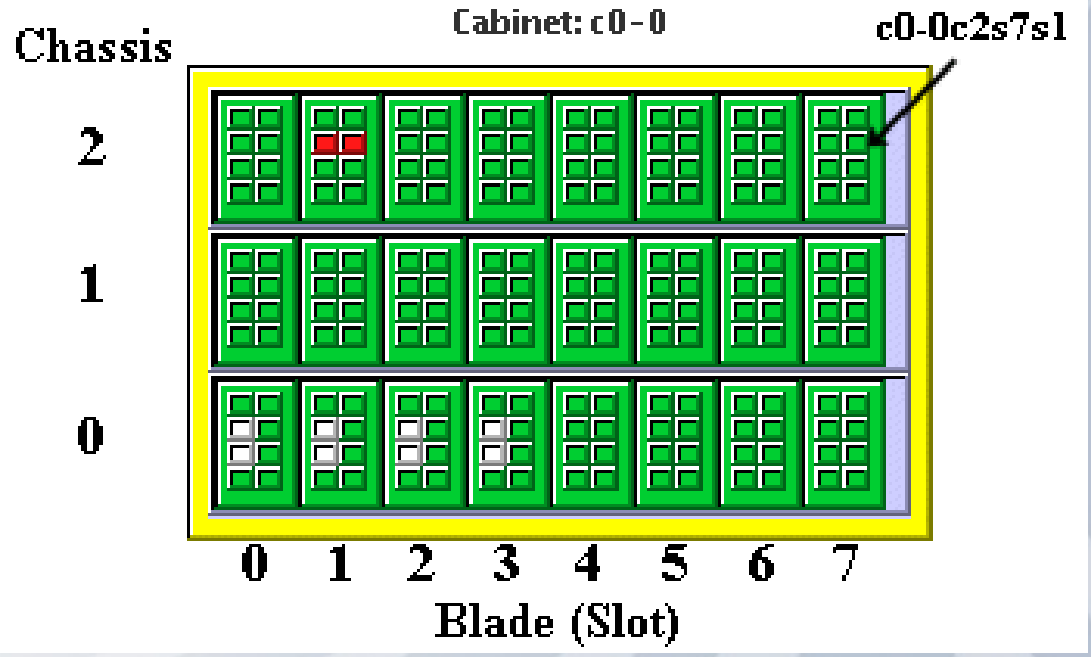
System Map

- The system map is a two dimensional representation of the Cray XT cabinet layout
- Each rectangle represents one cabinet
- A warning or error condition on any component within a cabinet is indicated by a differentially colored rectangle within the cabinet representation



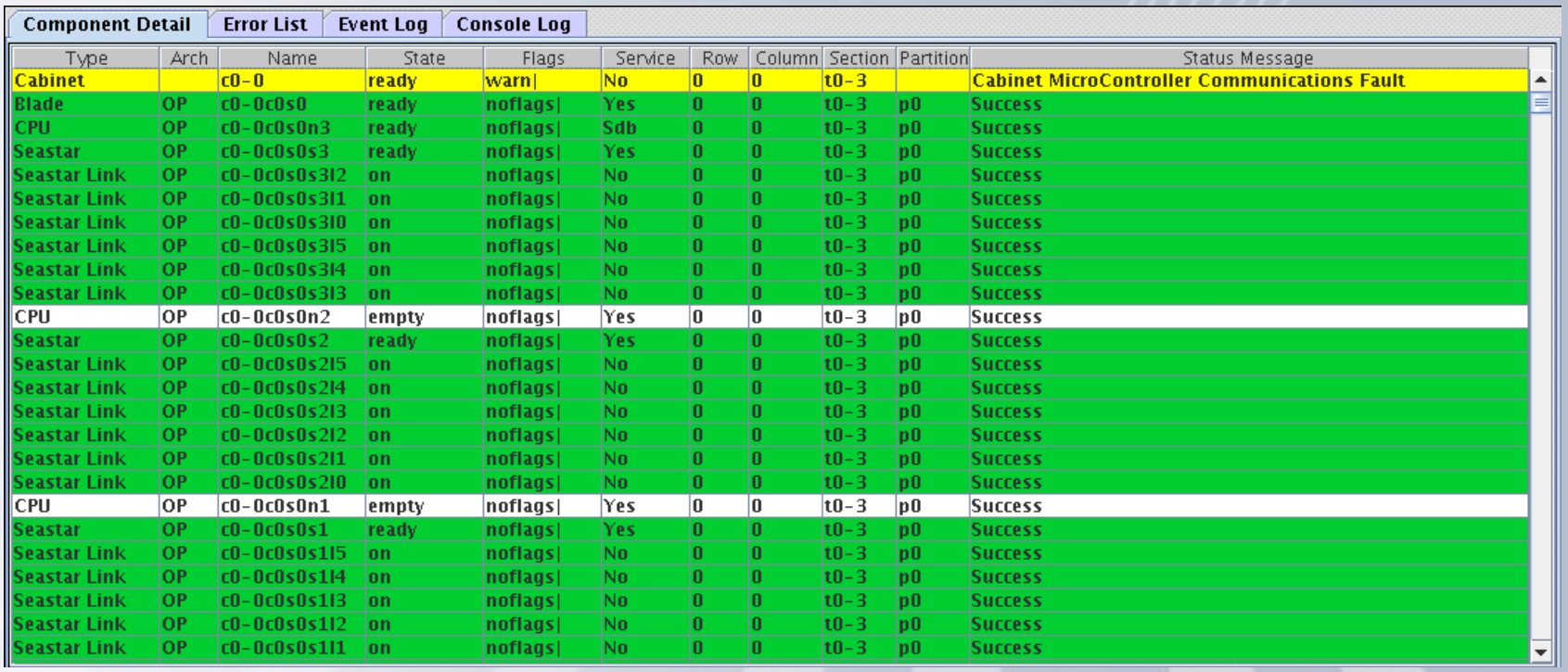
Cabinet Detail Map

- The cabinet detail map shows contents of the cabinet selected from the system map
- The status of the cabinet, it's blades, CPUs and Cray Seastars are shown



Component Detail

- Selecting a cabinet from the system map, or a blade from the cabinet detail map, causes the component detail tab to be filled with information on the selected object
- Selecting table column headers will cause the table to be sorted on the contents of that column



Type	Arch	Name	State	Flags	Service	Row	Column	Section	Partition	Status Message
Cabinet		c0-0	ready	warn|	No	0	0	t0-3		Cabinet MicroController Communications Fault
Blade	OP	c0-0c0s0	ready	noflags|	Yes	0	0	t0-3	p0	Success
CPU	OP	c0-0c0s0n3	ready	noflags|	Sdb	0	0	t0-3	p0	Success
Seastar	OP	c0-0c0s0s3	ready	noflags|	Yes	0	0	t0-3	p0	Success
Seastar Link	OP	c0-0c0s0s3l2	on	noflags|	No	0	0	t0-3	p0	Success
Seastar Link	OP	c0-0c0s0s3l1	on	noflags|	No	0	0	t0-3	p0	Success
Seastar Link	OP	c0-0c0s0s3l0	on	noflags|	No	0	0	t0-3	p0	Success
Seastar Link	OP	c0-0c0s0s3l5	on	noflags|	No	0	0	t0-3	p0	Success
Seastar Link	OP	c0-0c0s0s3l4	on	noflags|	No	0	0	t0-3	p0	Success
Seastar Link	OP	c0-0c0s0s3l3	on	noflags|	No	0	0	t0-3	p0	Success
CPU	OP	c0-0c0s0n2	empty	noflags|	Yes	0	0	t0-3	p0	Success
Seastar	OP	c0-0c0s0s2	ready	noflags|	Yes	0	0	t0-3	p0	Success
Seastar Link	OP	c0-0c0s0s2l5	on	noflags|	No	0	0	t0-3	p0	Success
Seastar Link	OP	c0-0c0s0s2l4	on	noflags|	No	0	0	t0-3	p0	Success
Seastar Link	OP	c0-0c0s0s2l3	on	noflags|	No	0	0	t0-3	p0	Success
Seastar Link	OP	c0-0c0s0s2l2	on	noflags|	No	0	0	t0-3	p0	Success
Seastar Link	OP	c0-0c0s0s2l1	on	noflags|	No	0	0	t0-3	p0	Success
Seastar Link	OP	c0-0c0s0s2l0	on	noflags|	No	0	0	t0-3	p0	Success
CPU	OP	c0-0c0s0n1	empty	noflags|	Yes	0	0	t0-3	p0	Success
Seastar	OP	c0-0c0s0s1	ready	noflags|	Yes	0	0	t0-3	p0	Success
Seastar Link	OP	c0-0c0s0s1l5	on	noflags|	No	0	0	t0-3	p0	Success
Seastar Link	OP	c0-0c0s0s1l4	on	noflags|	No	0	0	t0-3	p0	Success
Seastar Link	OP	c0-0c0s0s1l3	on	noflags|	No	0	0	t0-3	p0	Success
Seastar Link	OP	c0-0c0s0s1l2	on	noflags|	No	0	0	t0-3	p0	Success
Seastar Link	OP	c0-0c0s0s1l1	on	noflags|	No	0	0	t0-3	p0	Success

Component Detail

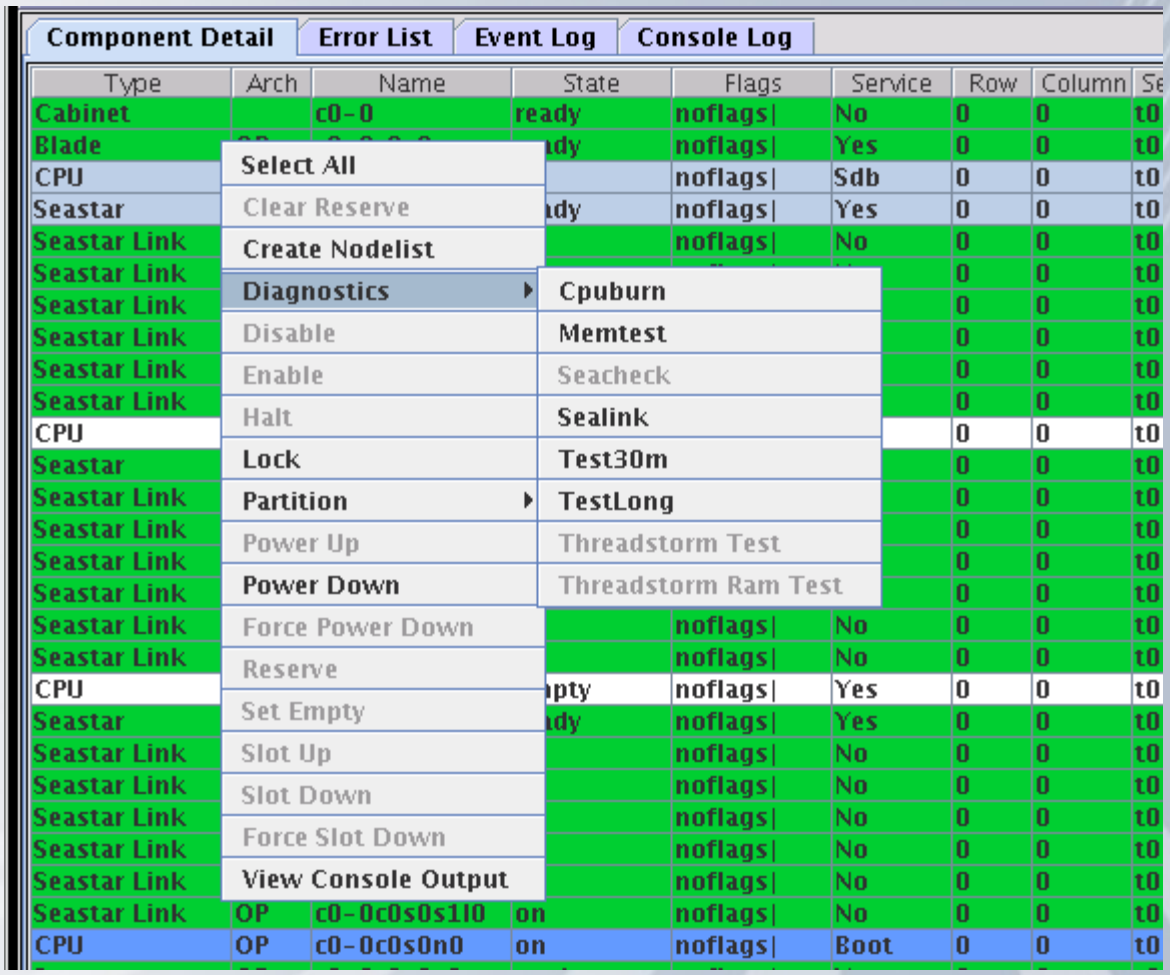
- Components may be selected with the left mouse button
- The right mouse button brings up a menu with the following options:

Clear Reserve	Create Nodelist	Diagnostics
Disable	Enable	Halt
Lock	Partition	Power Up
Power Down	Force Power Down	Reserve
Set Empty	Slot Up	Slot Down
Force Slot Down	View Console Output	

All menu options are context sensitive, so the chosen option will effect only those components on which such an operation is possible, given the type and state of the selected components.

Component Detail Menu Options

- When a menu option is chosen, the user is always shown a confirmation dialog prior to the action being performed

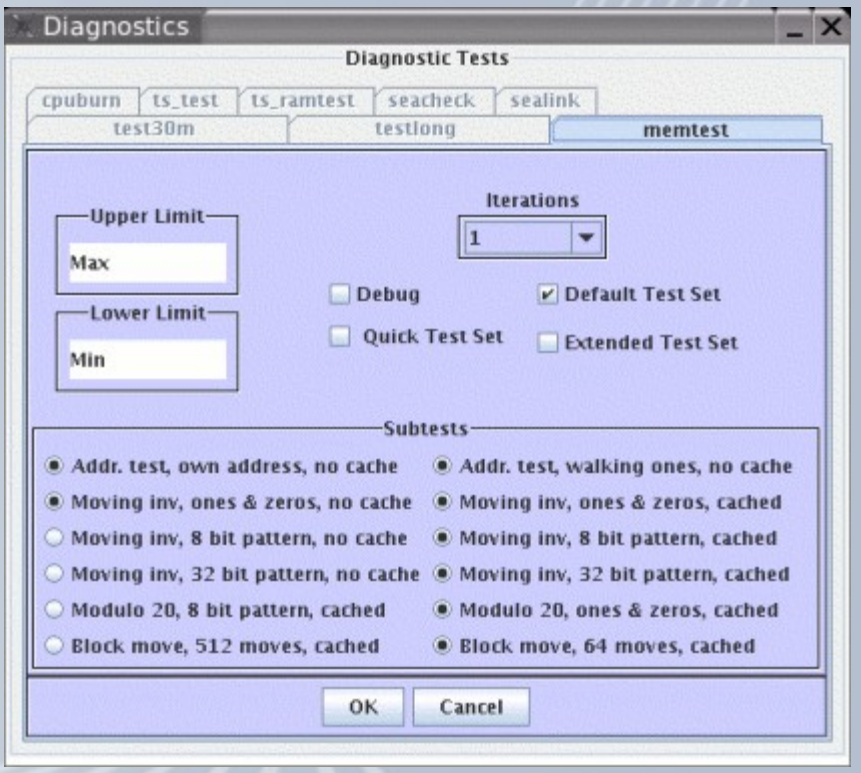


Component Detail Menu Options

- Clear Reserve:
  - To release a reserved component and return it to normal operation
  - Once a component is reserved, it will not accept new jobs, but any jobs already running on the component are completed in an orderly fashion
- Create Nodelist
  - This option is similar to the Save Node List option on the File menu
  - The difference is that the File menu Save Node List option creates a list of all nodes in all cabinets in the system, while this option creates a list of only the selected nodes
  - Selecting a cabinet or blade adds all of their component nodes to the nodelist, while selecting individual nodes adds only the selected nodes to the nodelist
  - By default, this file is written to your home directory and has the name RsmsNodeList-MMDD-HHMM.SS.txt

Diagnostics Menu Options

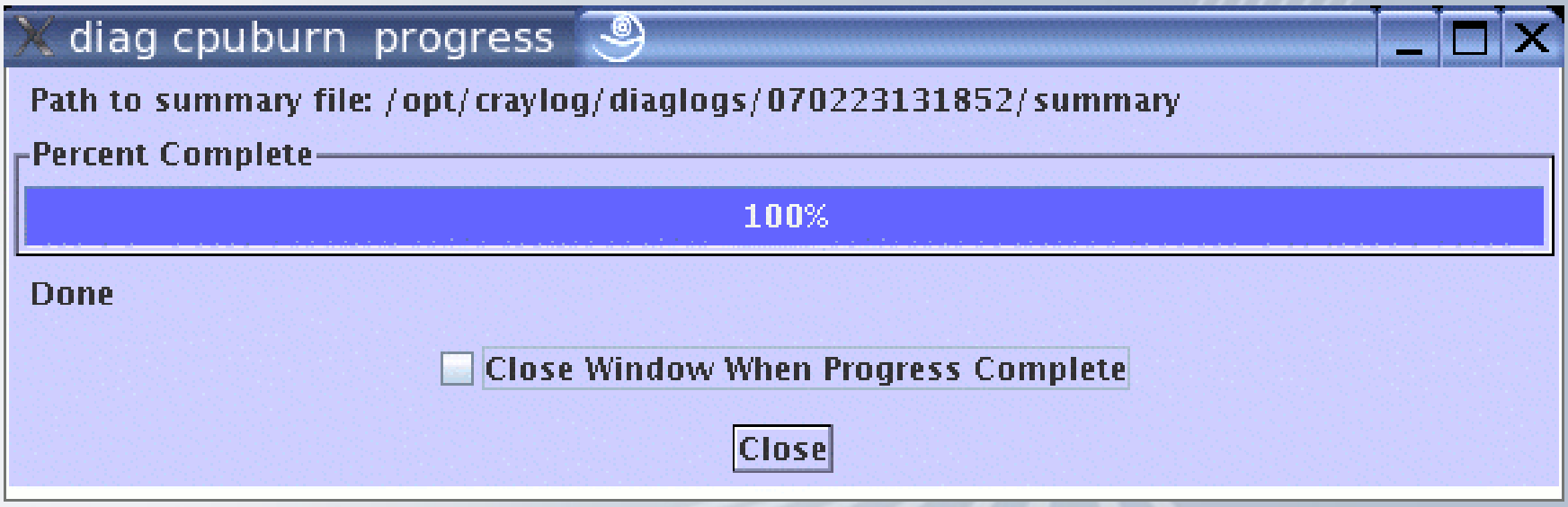
- The list of diagnostic options includes:
  - cpuburn
  - memtest
  - seacheck
  - sealink
  - test30m
  - testlong
  - threadstorm test
  - threadstorm ram test



- The diagnostic options offered will be dependent upon the type of selected components, as well as their current state

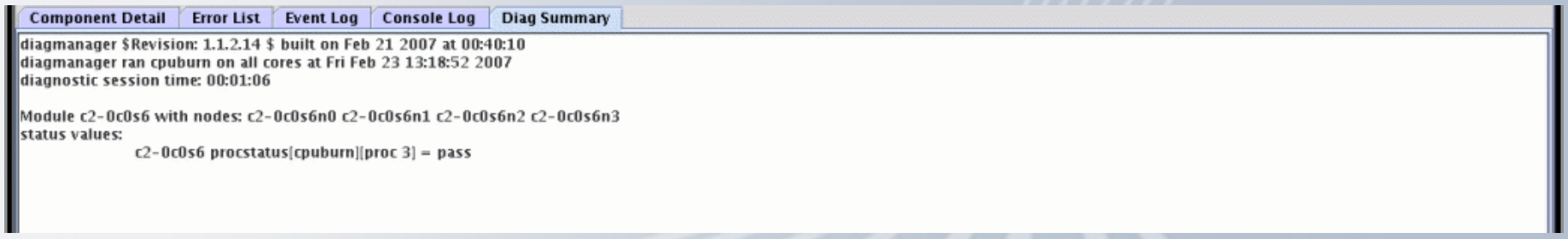
Diagnostics Operation

- A progress dialog window is shown during the course of the diagnostic test



Diagnostics Operation

- When the diagnostic test completes, a diagnostic summary tab is added to the xtgui display
- A right mouse click on this window will bring up a menu with the options to:
  - close tab
  - close all diagnostics tabs



Detail Menu Options

Disable	If links, nodes, or Cray Seastar chips have hardware problems, you can mark the component as downed by the administrator and prevent its being allocated.
Enable	Re-enable a component and return it to normal operation.
Halt	Stop a component immediately. The component immediately ceases operation and any data or processes running on the component are lost.
Lock	Locks a component manually. Components are locked automatically when a command that can change their state is running. As the command is started, the state manager locks the component so that nothing else can affect the component's state while the command is running. When a manager is finished with a command, it unlocks the component automatically. If the manager for some reason fails to unlock the component, it can be unlocked manually using the Actions menu Show Locks option.

Detail Menu Options

Power Up	Power up a component. Power commands are hierarchical; that is, there are a number of ways to power up or power down a lower-level component. For example, to power up a node, you can power it up directly or power up a component of which it is a part.
Power Down	Power down a component. Powering down a cabinet powers down all components within the cabinet, including the L0 controllers.
Force Power Down	Force a power down of a component. If you choose this option, the power manager ignores the operational state of the components that are being acted upon.
Reserve	Reserve a component. Once a component is reserved, it will not accept new jobs, but any jobs already running on the component are completed in an orderly fashion.

Detail Menu Options

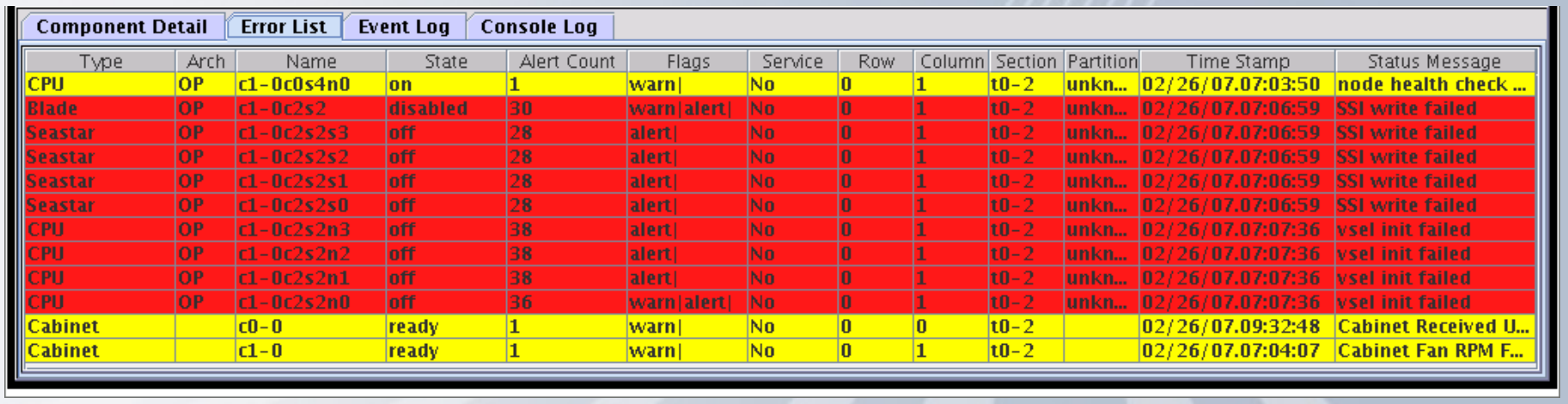
Set Empty	Set a selected component to the empty state. This is typically done when a component, usually a blade, is physically removed from the system. When a component is set to empty, the system ignores it and routes around it. To return a component to normal operation after it has been marked as empty -- for example, after a replacement blade has been installed , select the Enable option from the menu.
Slot Up	Power up all components on a blade.
Slot Down	Power down all components on a blade including the L0 controller.
Force Slot Down	Force the blade to power down regardless of operational state.

Detail Menu Option: View Console Output

- Display console log messages from selected nodes
- A right mouse click on this window will bring up a pop up menu with the options
  1) show filter targets
  2) close tab
  3) close all filter console

Error List

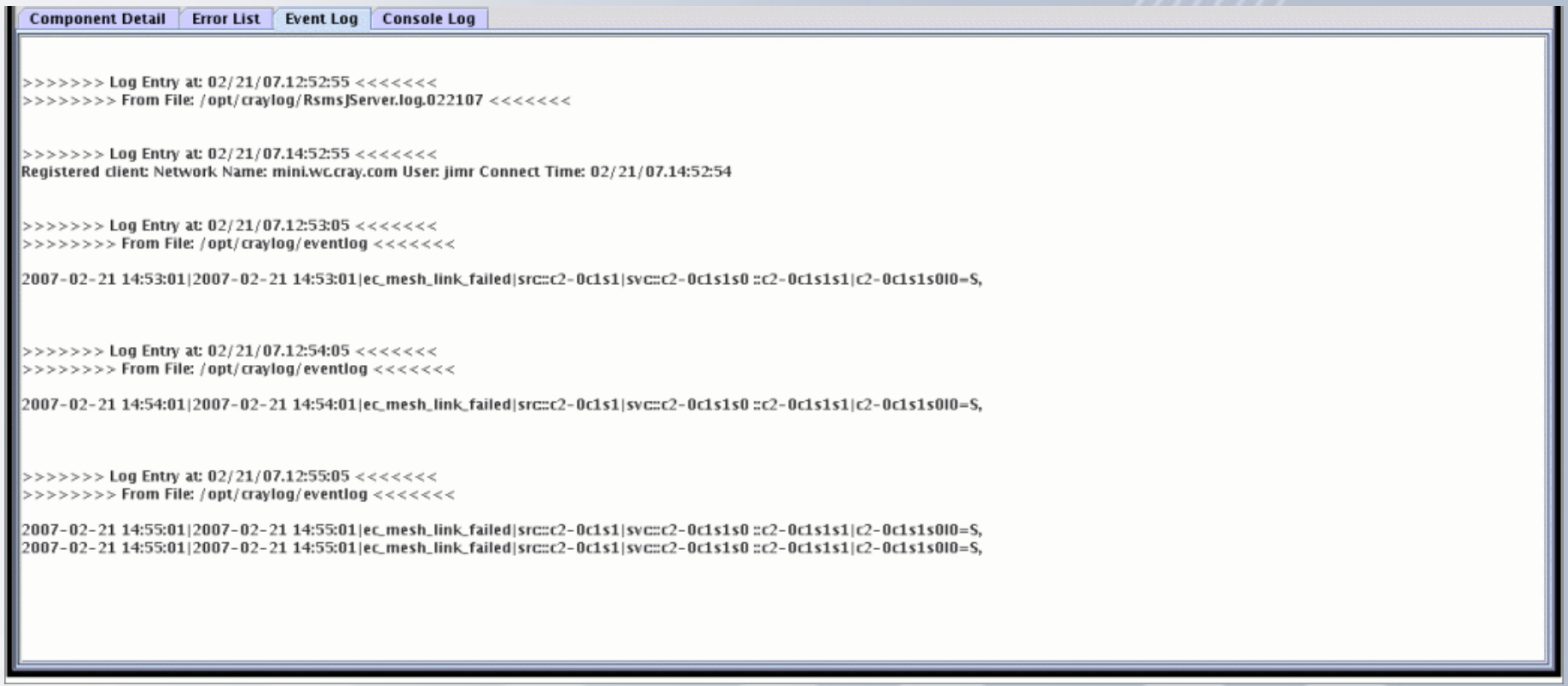
- All components with a current warning or alert status, are shown in the error list
- Selecting table column headers will cause the table to be sorted on the contents of that column
- Components may be selected with the left mouse button
- The right mouse button brings up a menu with the following options:
  - Clear Warning
  - Clear Alert



Type	Arch	Name	State	Alert Count	Flags	Service	Row	Column	Section	Partition	Time Stamp	Status Message
CPU	OP	c1-0c0s4n0	on	1	warn|	No	0	1	t0-2	unkn...	02/26/07.07:03:50	node health check ...
Blade	OP	c1-0c2s2	disabled	30	warn|alert|	No	0	1	t0-2	unkn...	02/26/07.07:06:59	SSI write failed
Seastar	OP	c1-0c2s2s3	off	28	alert|	No	0	1	t0-2	unkn...	02/26/07.07:06:59	SSI write failed
Seastar	OP	c1-0c2s2s2	off	28	alert|	No	0	1	t0-2	unkn...	02/26/07.07:06:59	SSI write failed
Seastar	OP	c1-0c2s2s1	off	28	alert|	No	0	1	t0-2	unkn...	02/26/07.07:06:59	SSI write failed
Seastar	OP	c1-0c2s2s0	off	28	alert|	No	0	1	t0-2	unkn...	02/26/07.07:06:59	SSI write failed
CPU	OP	c1-0c2s2n3	off	38	alert|	No	0	1	t0-2	unkn...	02/26/07.07:07:36	vsel init failed
CPU	OP	c1-0c2s2n2	off	38	alert|	No	0	1	t0-2	unkn...	02/26/07.07:07:36	vsel init failed
CPU	OP	c1-0c2s2n1	off	38	alert|	No	0	1	t0-2	unkn...	02/26/07.07:07:36	vsel init failed
CPU	OP	c1-0c2s2n0	off	36	warn|alert|	No	0	1	t0-2	unkn...	02/26/07.07:07:36	vsel init failed
Cabinet		c0-0	ready	1	warn|	No	0	0	t0-2		02/26/07.09:32:48	Cabinet Received U...
Cabinet		c1-0	ready	1	warn|	No	0	1	t0-2		02/26/07.07:04:07	Cabinet Fan RPM F...

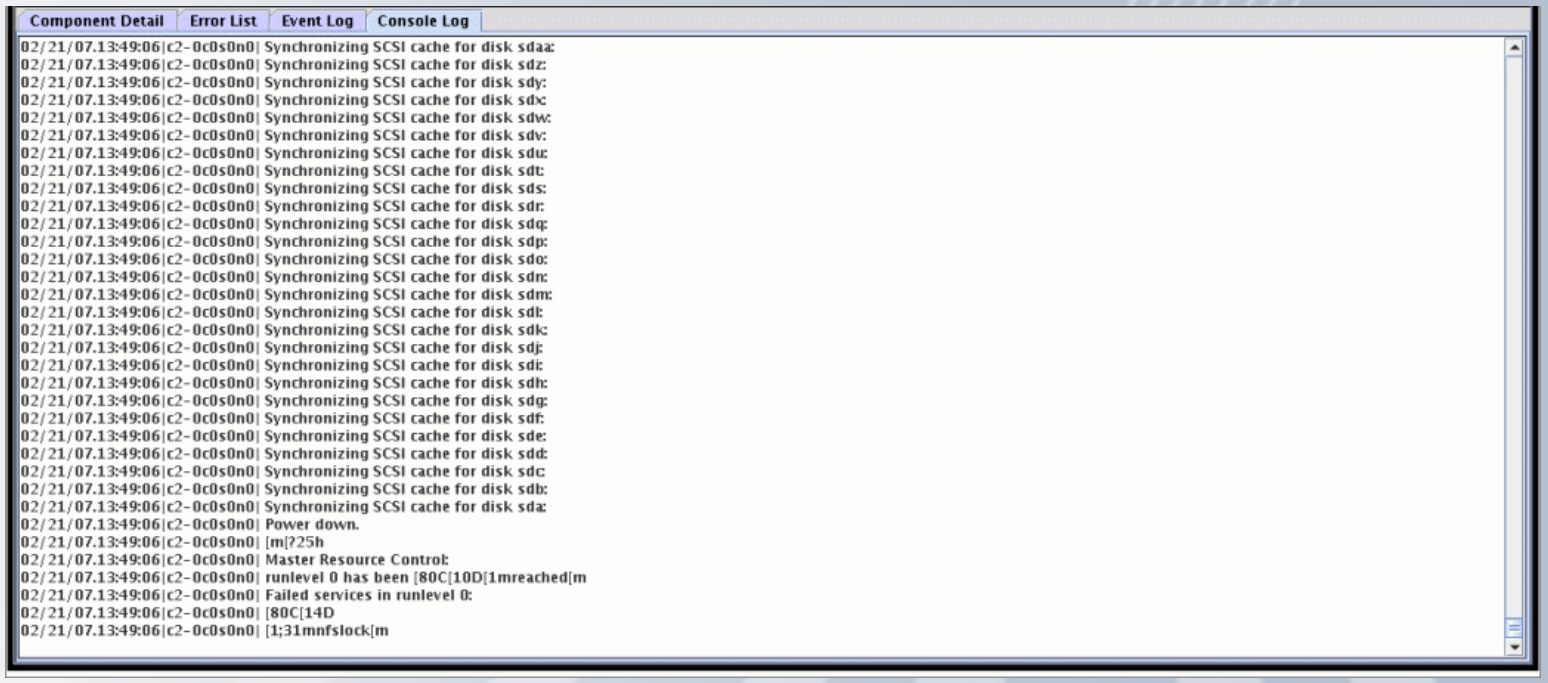
Event Log

- The event log displays recent information sent to the files that the xtgui server is configured to watch, as well as all commands executed by the xtgui server on behalf of the xtgui client



Console Log

- The console log displays console log messages from all Cray XT nodes
- Display of console output from selected nodes may be accomplished through the “view console output” menu option on the component detail popup menu



Toolbar Menus

- Various actions and configuration options my be accessed by using the options on the toolbar menus in the upper left-hand corner of the XTGUI window. These menus and their options are:

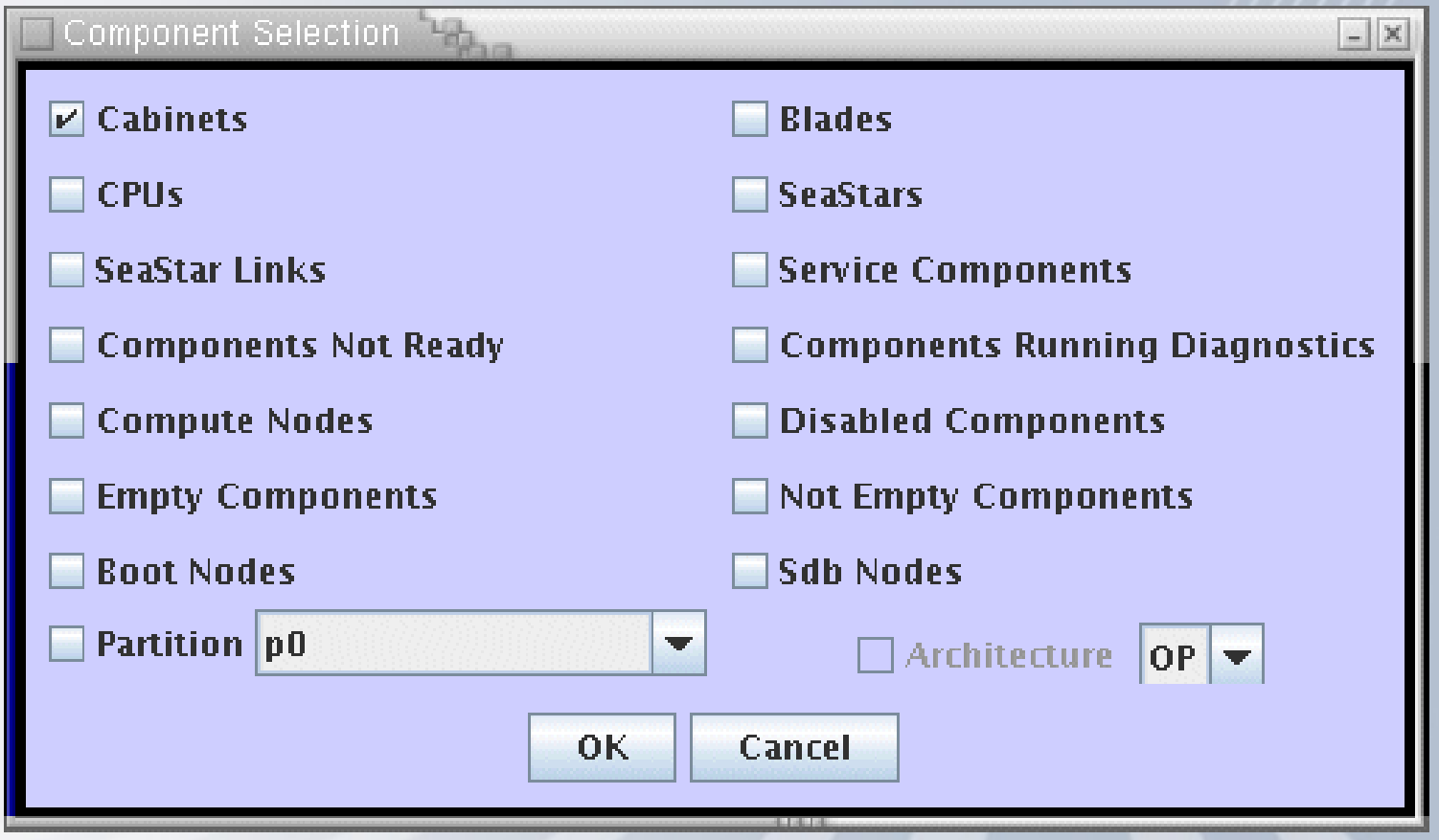
File	Actions	Preferences
Save Node List	Show Components	General
Exit	Show Active Commands	Connection
	Show Boot Configuration	Partition Configuration
	Show Server Status	
	Import/Export Sections	
	Show Locks	

File Menu

- Save Node List:
  - Create a text file listing all CPUs in the system, marking each "n" for empty or disabled, "i" for service, "c" for compute. For example:
    - c0-0c0s0n0 i
    - c0-0c0s0n1 n
    - c0-0c0s0n2 n
    - c0-0c0s0n3 i
    - c0-0c0s1n0 c
    - c0-0c0s1n1 c
  - This text file is saved in the users home directory and the naming convention is: RsmsNodeList-mmdd-hhmm.ss.txt
- Exit:
  - Exit the xtgui application.

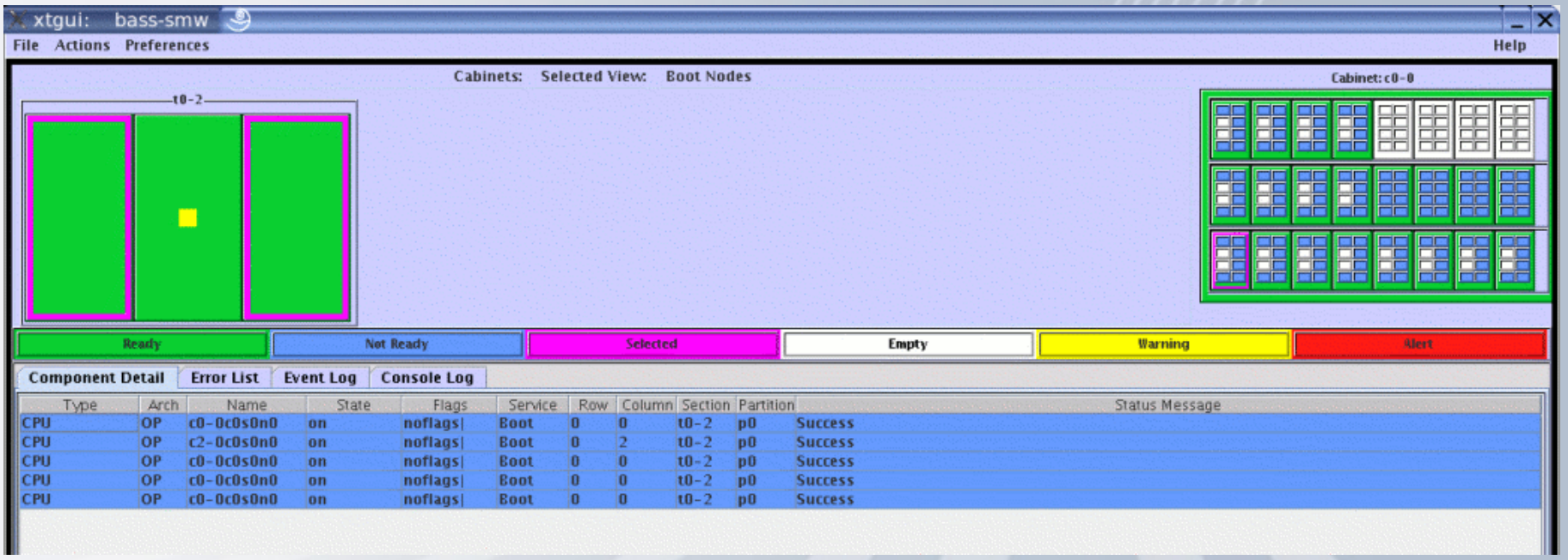
Actions Menu: Show Components

- This option presents a dialog that allows the user to select physical or logical groupings of components:



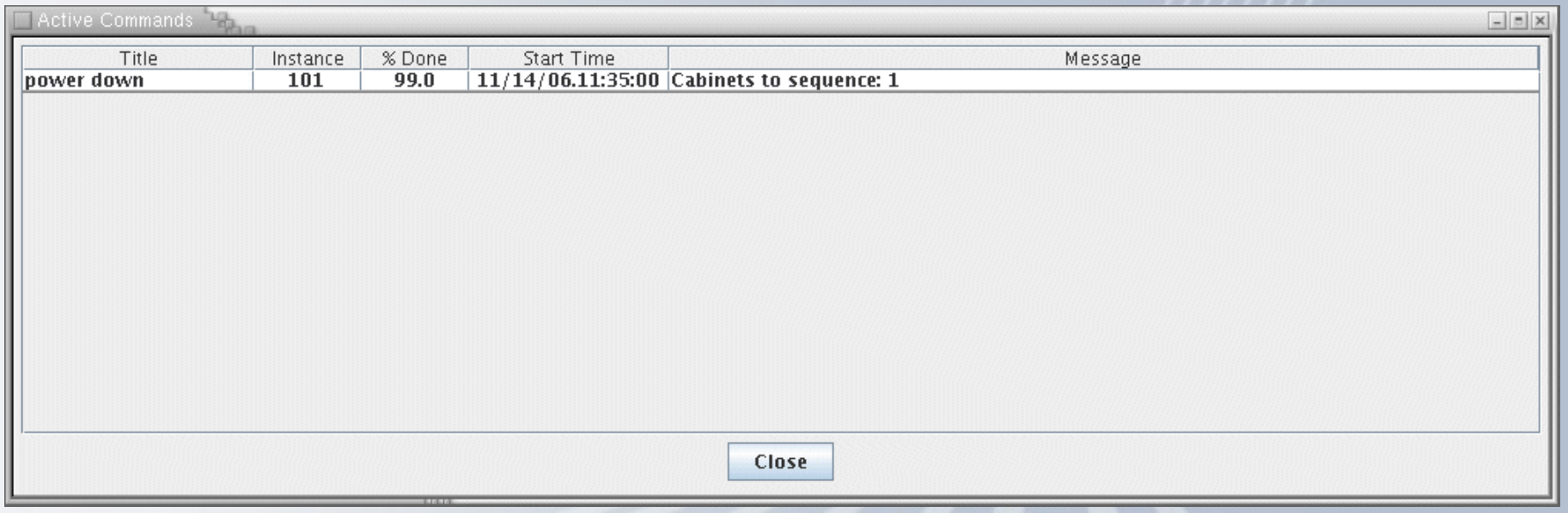
Show Components

- If “Boot Nodes” was chosen
  - the cabinets containing boot nodes would be selected in the system map
  - the cabinet detail window would indicate which blades within a cabinet contained the boot nodes
  - the component detail table would be populated with boot node component records.



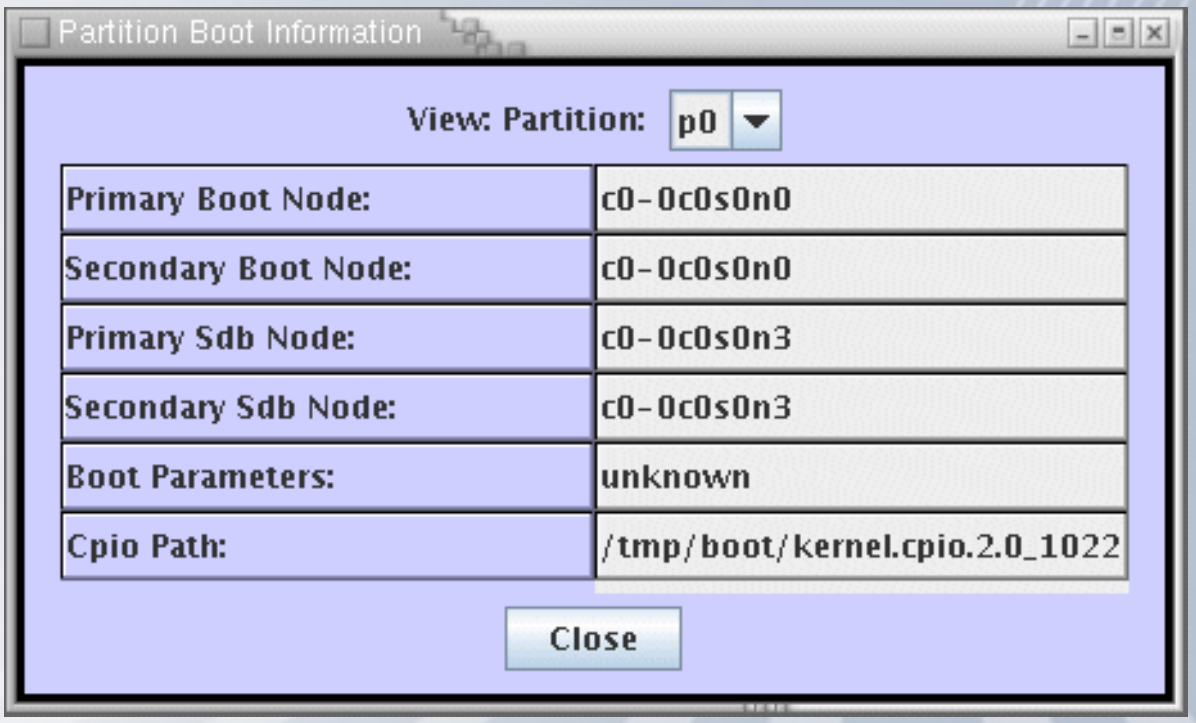
Actions Menu: Show Active Cmds

- This option displays a table of currently active commands that have been started by the xtgui client



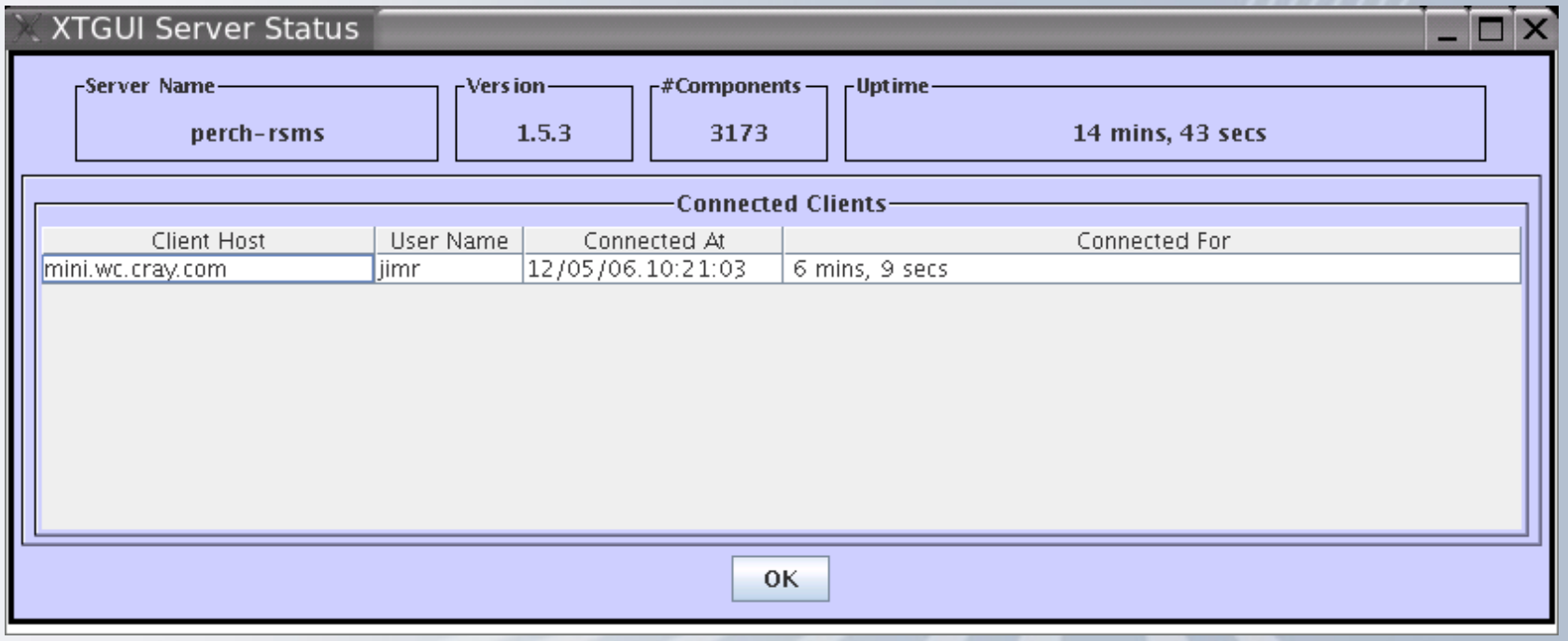
Actions Menu: Show Boot Configuration

- This option displays a dialog that provides way to show the boot configuration of all Cray XT partitions



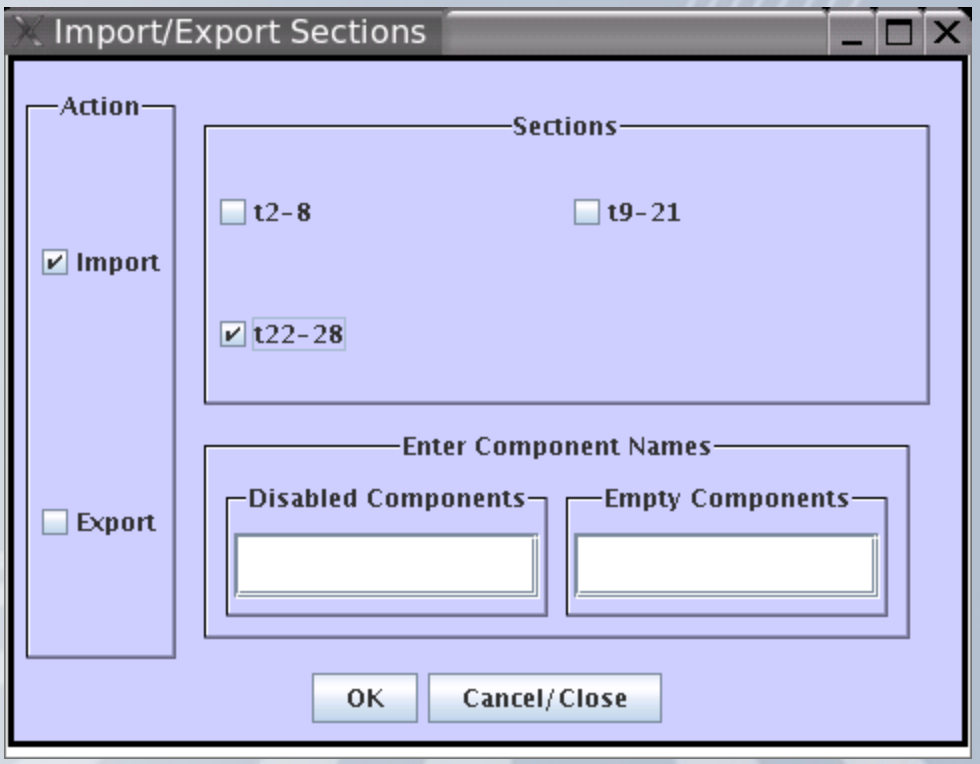
Actions Menu: Show Server Status

- This option displays a window that provides information on the xtgui server process and lists all connected xtgui clients



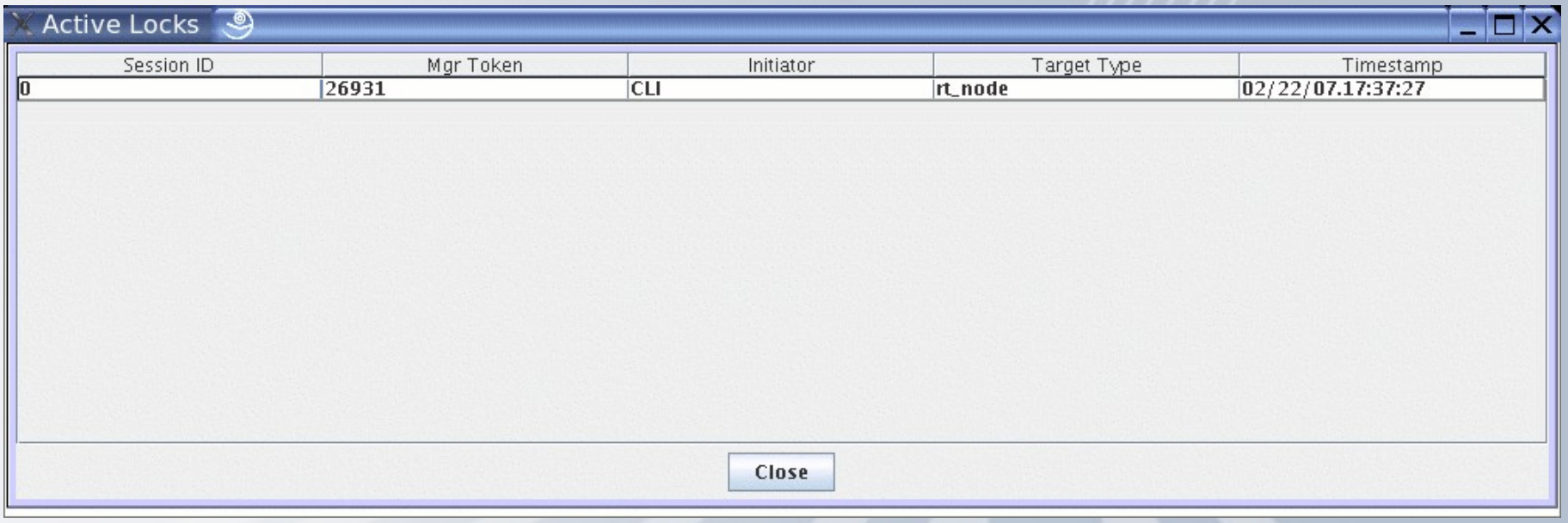
Actions Menu: Import/Export Sections

- This option displays a dialog allowing the user to import or export sections of the Cray XT system.
- The option is enabled only if more than one section has been defined.



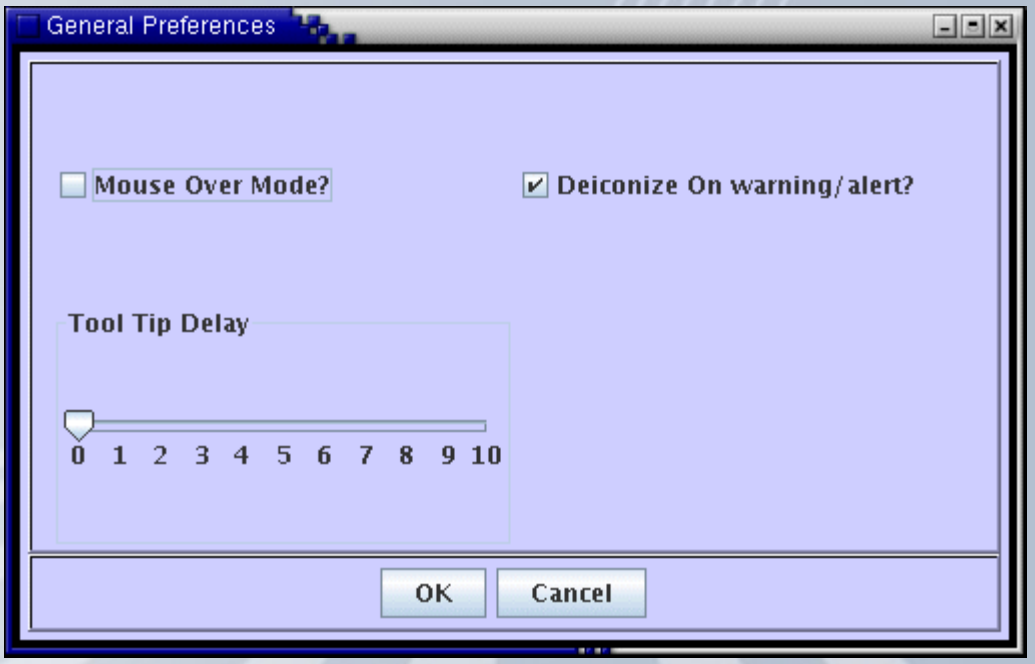
Actions Menu: Show Locks

- This option displays all currently active session locks.
- A left mouse click selects rows in the table and a right mouse click pops up a menu allowing display of the effected components for each session and the option of dismissing the lock.



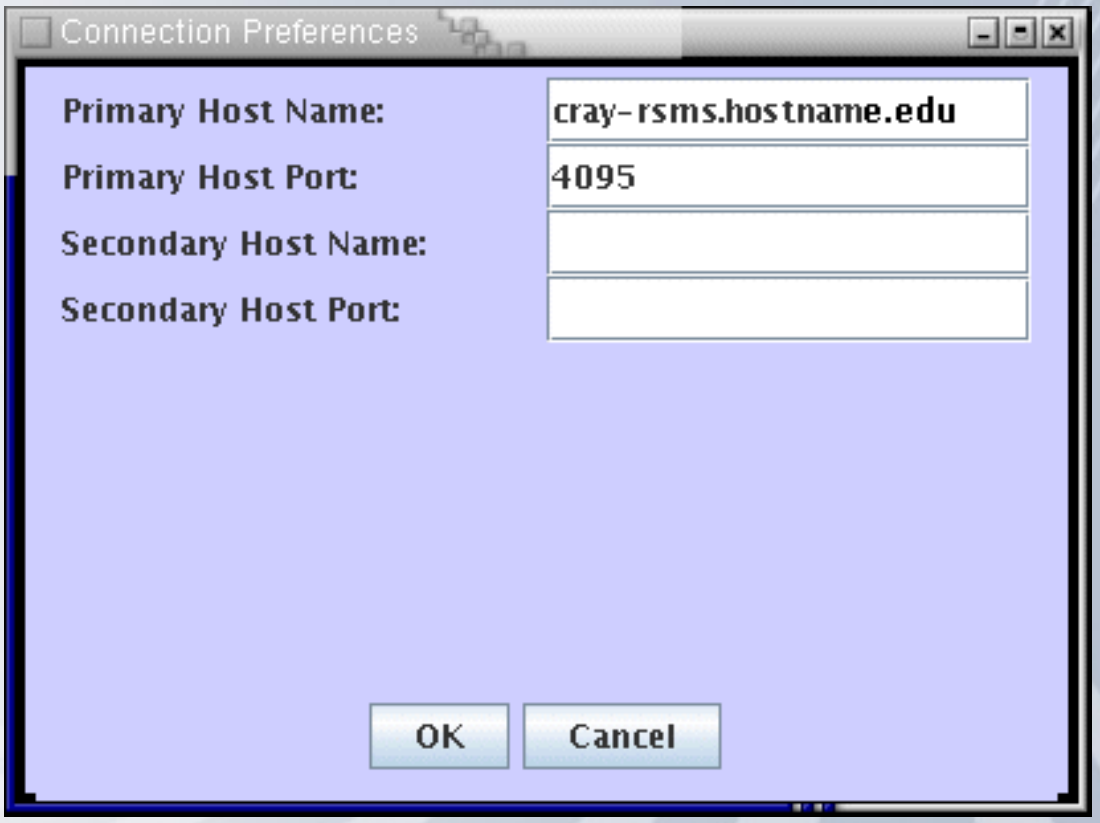
Preferences Menu: General Preferences

- This option displays a dialog that allows the configuration of three options:
  1. Mouse over mode (which means to automatically switch the cabinet detail window to the cabinet that the mouse is currently hovering over).
  2. Deiconize on warning/alert.
  3. Tool tip delay.



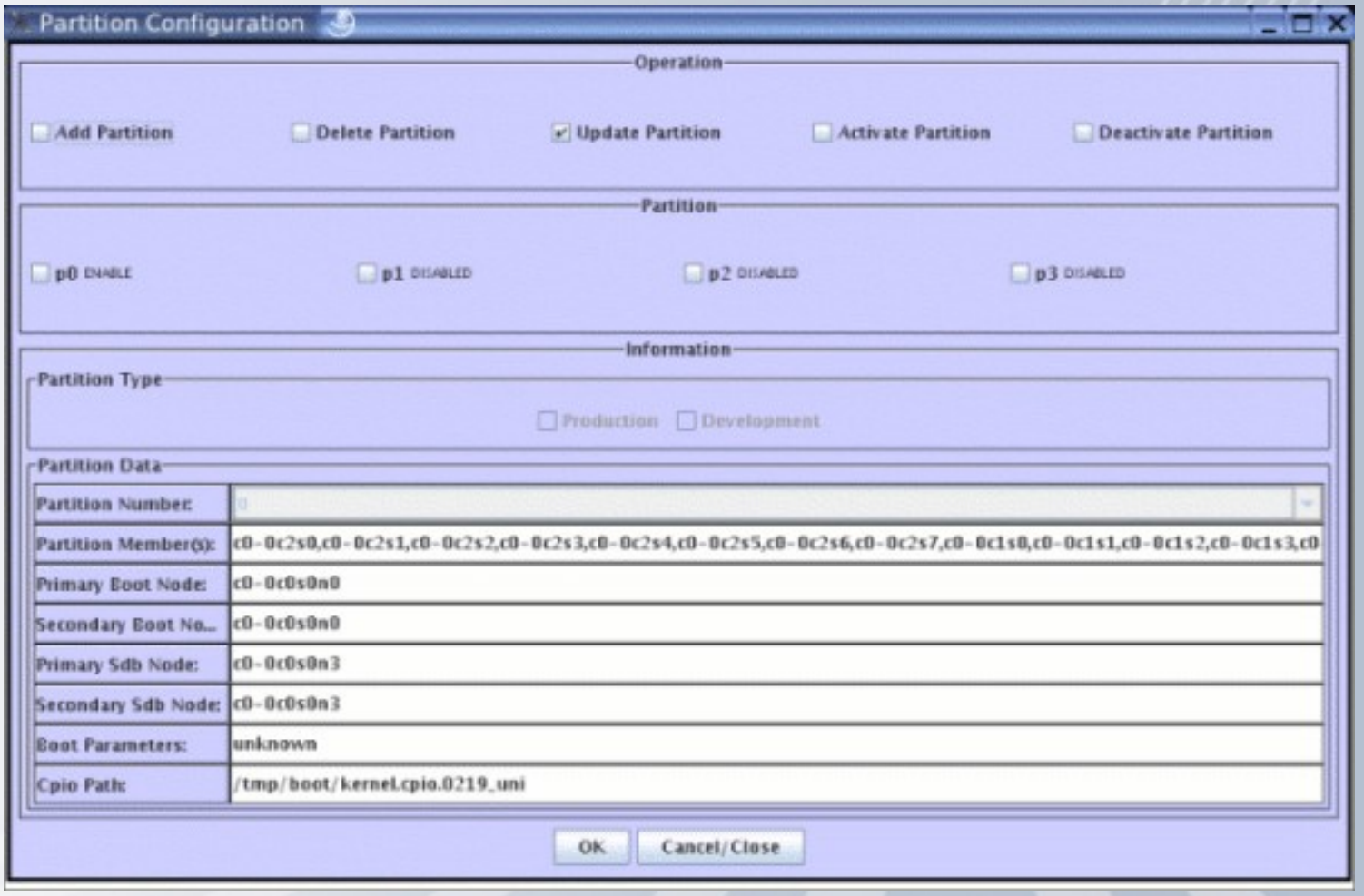
Preferences Menu: Connection

- This option displays a dialog which allows configuration of the host name and port number of the primary and secondary SMW systems to connect to



Preferences Menu: Partition Configuration

- This option presents the partition configuration dialog



Properties File

- The xtgui client uses the file RsmsClientProperties.txt to store configuration options
- This file is copied into the users home directory the first time that the application runs
- Examples of the type of configuration options that are stored in this file:
  - client.primaryPort=4095
  - client.secondaryHost=
  - client.rmiTimeoutMilliseconds=120000
  - client.primaryHost=snowdrift
  - client.toolTipDelay=0
  - client.portNumber=0
  - client.heartBeatDelay=5000
  - client.deiconize=false
  - client.heartBeatFrequency=5000
  - client.secondaryPort=
  - client.mouseOverMode=false

Online Help

- Selecting the Help menu on the toolbar, displays the online help document