The Supercomputer Company

Active Manager Tutorial

"Tips and Tricks"

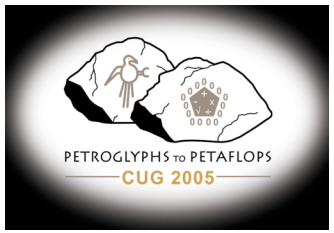
Amar Shan, Ron Westfall, Dave Strenski

CRAY USER GROUP INCORPORATED

CUG 2005

Agenda

- Active Manager Rationale – “why”?
- “Tips and Tricks”
  - Effective Use of Partitions
  - Customizing Alarm Handling
  - Using XML
  - Fabric Error Detection and Correction
  - A few other smaller features …



The Cray XD1

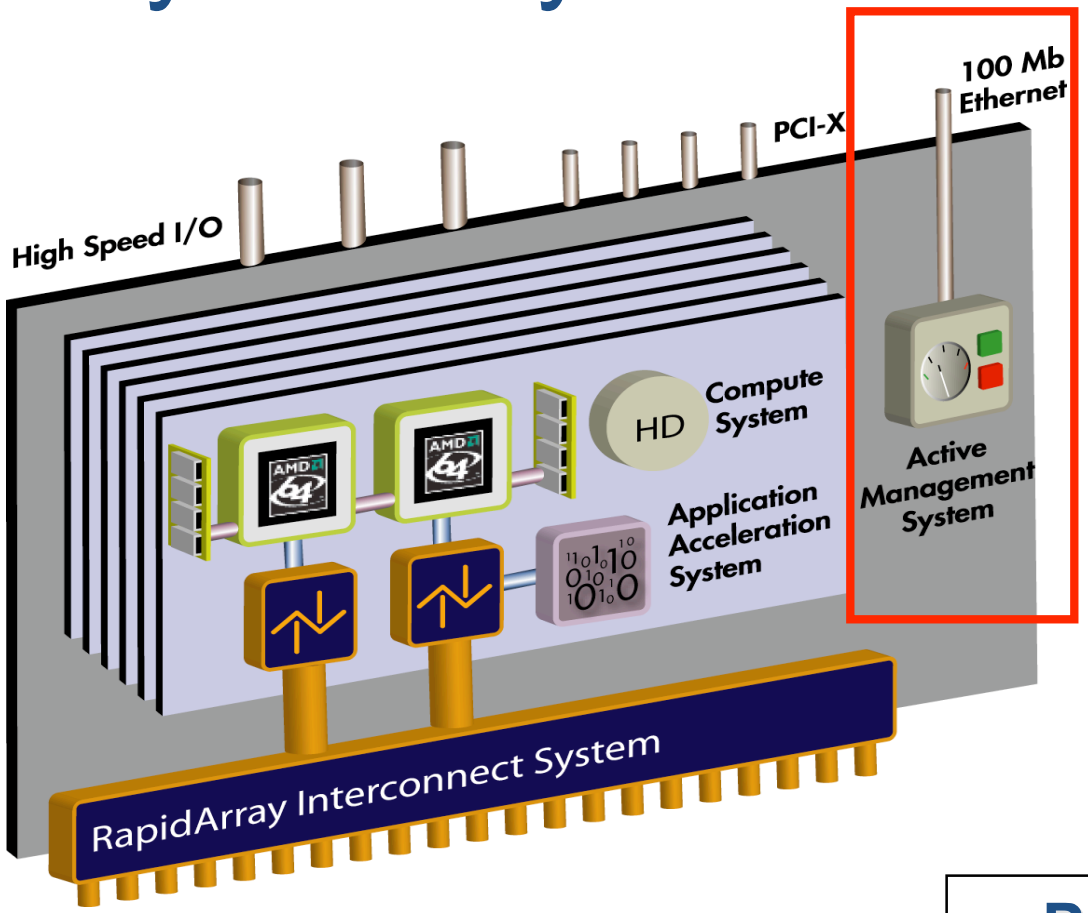
Cray XD1

- Built for price/performance
  - Interconnect bandwidth/latency
  - System-wide process synchronization
  - Application Acceleration FPGAs
- Standards-based
  - 32/64-bit X86, Linux, MPI
- High resiliency
  - Self-configuring, self-monitoring, self-healing
- Single system command & control
  - Intuitive, tightly integrated management software

Purpose-built and optimized for high performance workloads

PETROGLYPHS TO PETAFLOPS
CUG 2005

Cray XD1 System Architecture



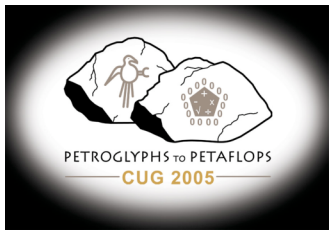
Compute
- **12 AMD Opteron 32/64 bit, x86 processors**
- **High Performance Linux**

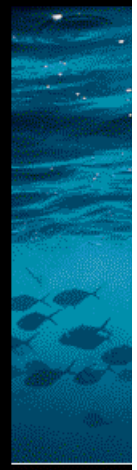
RapidArray Interconnect
- **12 communications processors**
- **1 Tb/s switch fabric**

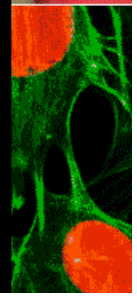
Active Management
- **Dedicated processor**

Application Acceleration
- **6 co-processors**

Processors directly connected via integrated switch fabric







Rationale for Active Manager

CRAY USER GROUP INCORPORATED

What Matters to Users

IDC Cluster User Study – 2004

"Top technical problems facing organizations using clusters"

- Inadequate throughput capacity (31%)
- Difficulties in scaling systems and capacity fast enough to meet user needs (31%)
- Costly maintenance and support (22%)

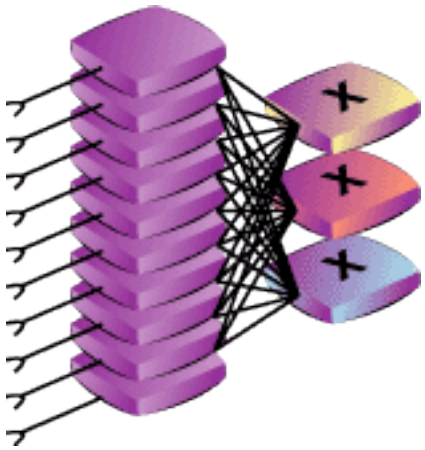
"The biggest challenge to implementing clusters is systems management (43%)"

"Users are primarily concerned about the usability, manageability and long term operation of installed cluster systems"

"Clusters of 15 to 20 machines were the norm….now you typically have clusters of 50 to 100 machines, up to over 1000" (Donald Becker, Beowulf founder, 2004)

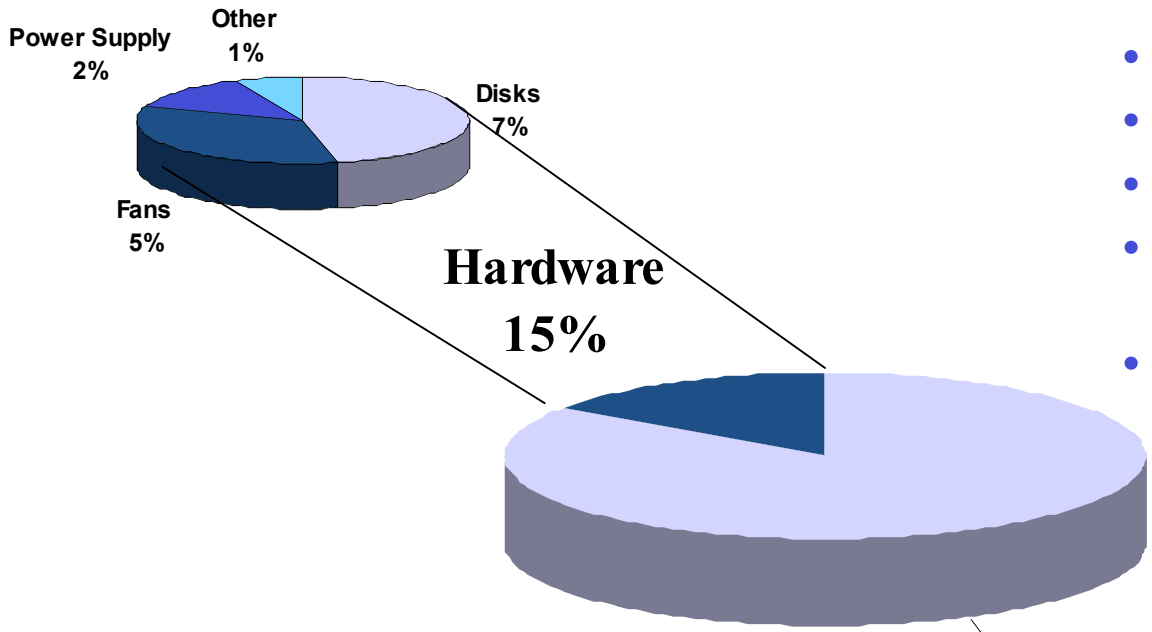
"One cluster user described the management problem this way: All the components work, they just don't always work together"

(IDC 2002)



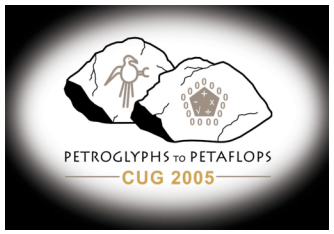
It is crucial to address the management issues

Causes of System Outages



- **Most problems occur during:**
  - Upgrades,
  - Problem diagnosis,
  - Configuration
- **Complex systems**
- **Patchwork software**
- **Inadequate tools**
- **Multivendor h/w, s/w**
- **Driver & software compatibility**
- **Corrupt, stale software & configs**

Source: UC Berkeley



Active Manager System

CLI and Web Access

Active Management Software

Usability

- Single System Command and Control

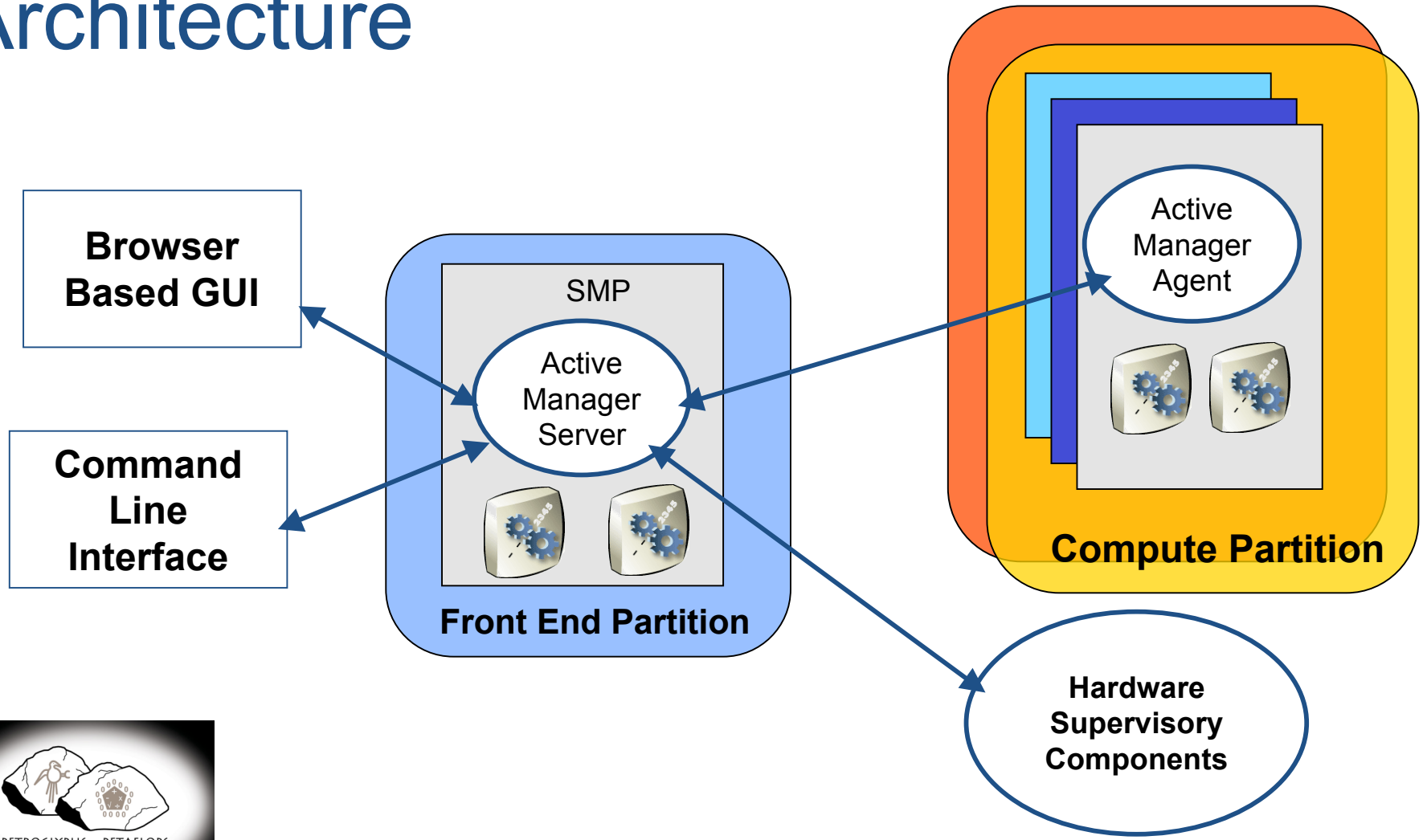
Resiliency

- Dedicated management processors, real-time OS and communications fabric.
- Proactive background diagnostics with self-healing.

PETROGLYPHS TO PETAFLOPS
CUG 2005

Automated management for exceptional reliability, availability, serviceability

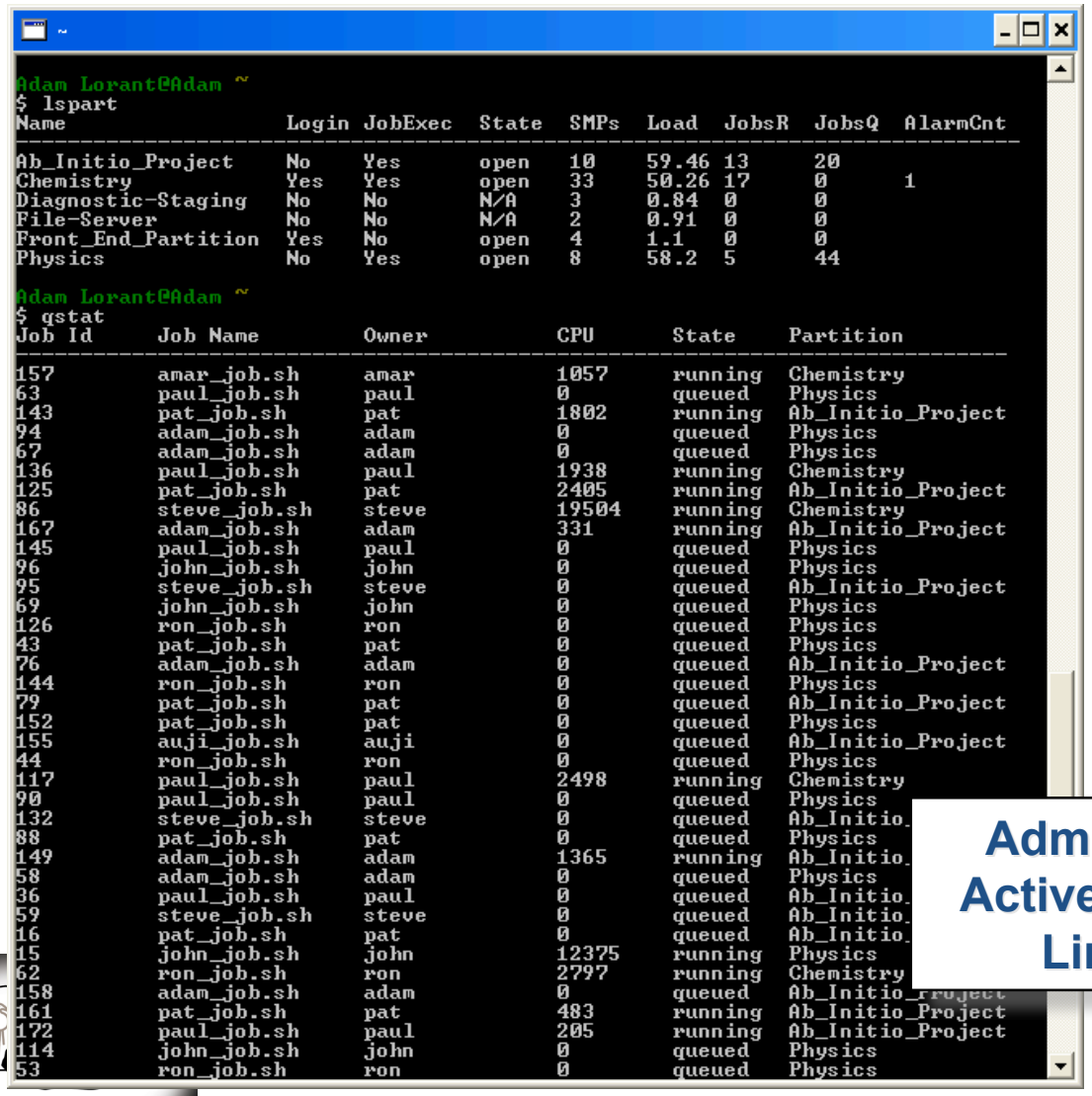
Active Manager Software Architecture



Active Manager GUI: SysAdmin Portal

GUI provides quick access to status info and system functions

Active Manager Command Line Interface



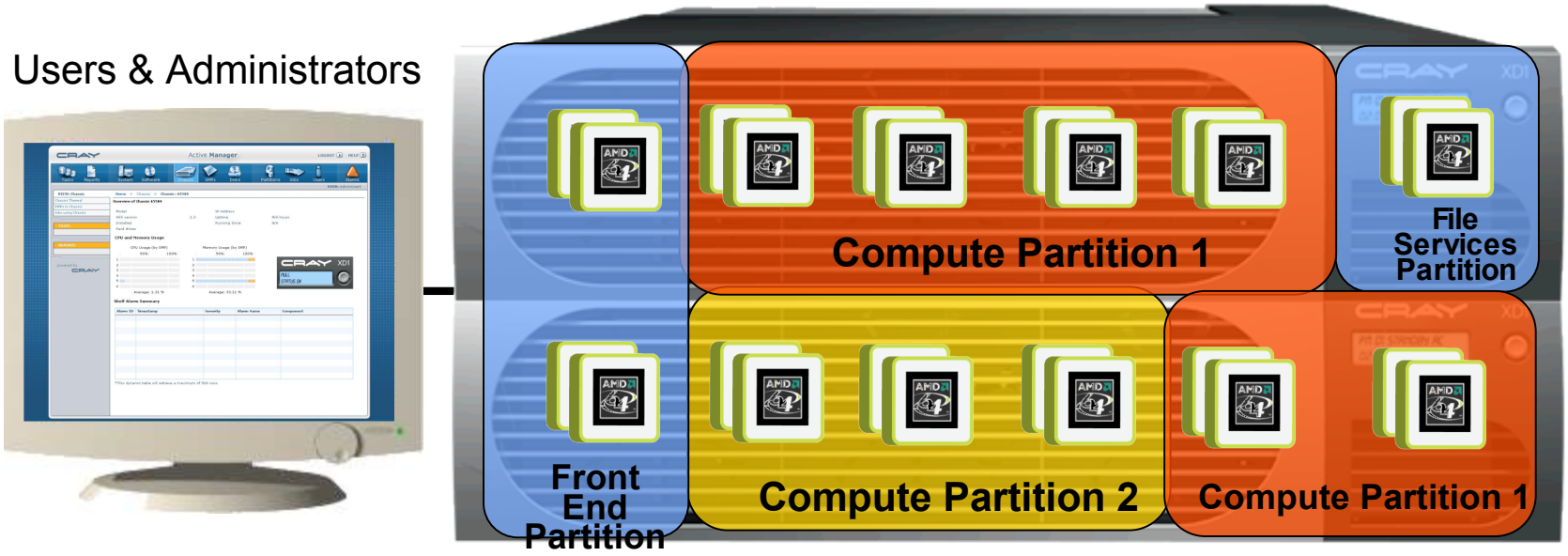
```
Adam Lorant@Adam ~
$ lspart
Name                 Login JobExec  State  SMPs  Load  JobsR  JobsQ  AlarmCnt
-------------------------------------------------------------------------------
Ab_Initio_Project    No    Yes      open   10    59.46 13     20
Chemistry            Yes   Yes      open   33    50.26 17     0      1
Diagnostic-Staging   No    No       N/A    3     0.84  0      0
File-Server          No    No       N/A    2     0.91  0      0
Front_End_Partition  Yes   No       open   4     1.1   0      0
Physics              No    Yes      open   8     58.2  5      44

Adam Lorant@Adam ~
$ qstat
Job Id     Job Name        Owner          CPU      State    Partition
-------------------------------------------------------------------
157        amar_job.sh     amar           1057     running  Chemistry
63         paul_job.sh     paul           0        queued   Physics
143        pat_job.sh      pat            1802     running  Ab_Initio_Project
94         adam_job.sh     adam           0        queued   Physics
67         adam_job.sh     adam           0        queued   Physics
136        paul_job.sh     paul           1938     running  Chemistry
125        pat_job.sh      pat            2405     running  Ab_Initio_Project
86         steve_job.sh    steve          19504    running  Chemistry
167        adam_job.sh     adam           331      running  Ab_Initio_Project
145        paul_job.sh     paul           0        queued   Physics
96         john_job.sh     john           0        queued   Physics
95         steve_job.sh    steve          0        queued   Ab_Initio_Project
69         john_job.sh     john           0        queued   Physics
126        ron_job.sh      ron            0        queued   Physics
43         pat_job.sh      pat            0        queued   Physics
76         adam_job.sh     adam           0        queued   Ab_Initio_Project
144        ron_job.sh      ron            0        queued   Physics
79         pat_job.sh      pat            0        queued   Ab_Initio_Project
152        pat_job.sh      pat            0        queued   Physics
155        auji_job.sh     auji           0        queued   Ab_Initio_Project
44         ron_job.sh      ron            0        queued   Physics
117        paul_job.sh     paul           2498     running  Chemistry
90         paul_job.sh     paul           0        queued   Physics
132        steve_job.sh    steve          0        queued   Ab_Initio_
88         pat_job.sh      pat            0        queued   Physics
149        adam_job.sh     adam           1365     running  Ab_Initio_
58         adam_job.sh     adam           0        queued   Physics
36         paul_job.sh     paul           0        queued   Ab_Initio_
59         steve_job.sh    steve          0        queued   Ab_Initio_
16         pat_job.sh      pat            0        queued   Ab_Initio_
15         john_job.sh     john           12375    running  Physics
62         ron_job.sh      ron            2797     running  Chemistry
158        adam_job.sh     adam           0        queued   Ab_Initio_Project
161        pat_job.sh      pat            483      running  Ab_Initio_Project
172        paul_job.sh     paul           205      running  Ab_Initio_Project
114        john_job.sh     john           0        queued   Physics
53         ron_job.sh      ron            0        queued   Physics
```

Administrators can access Active Manager software and Linux through the CLI

Partitions

System Partitions

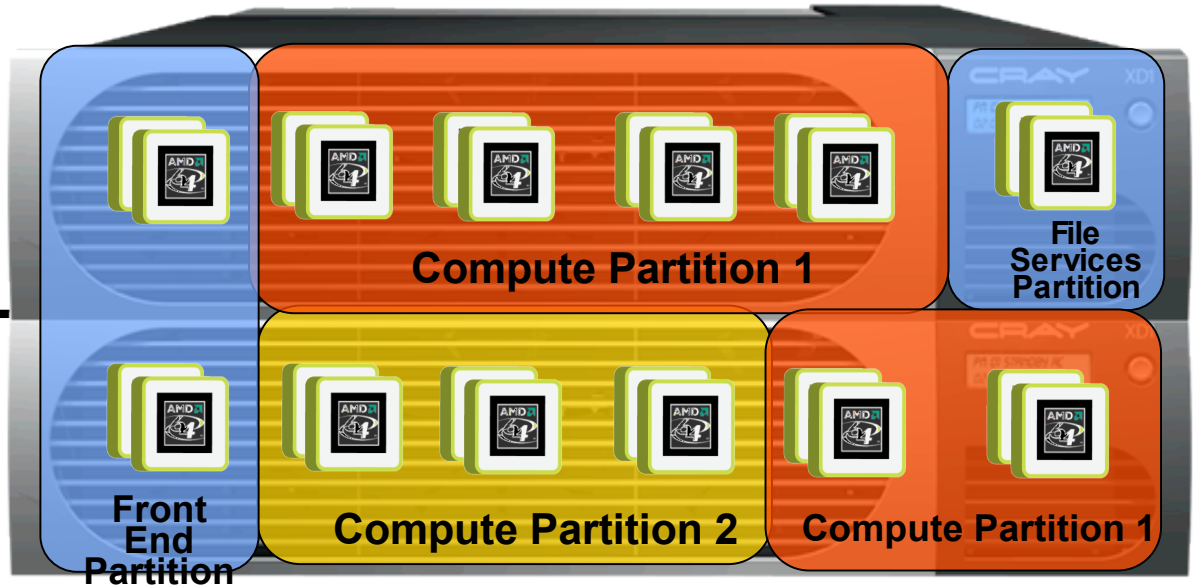


- Front End Partition
- Compute Partition
- Service Partition
  - File Services
  - Database
  - DNS

Manage multiple processors and copies of Linux as single, unified system

Automated Management

Users & Administrators



- Partition management
- Linux configuration
- Hardware monitoring
- Software upgrades
- File system management
- Data backups

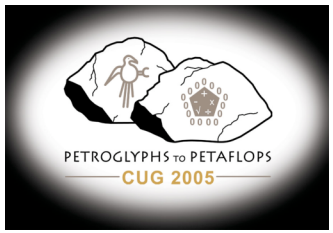
- Network configuration
- Accounting & user management
- Security
- Performance analysis
- Resource & queue management

PETROGLYPHS TO PETAFLOPS
CUG 2005

Single System Command and Control

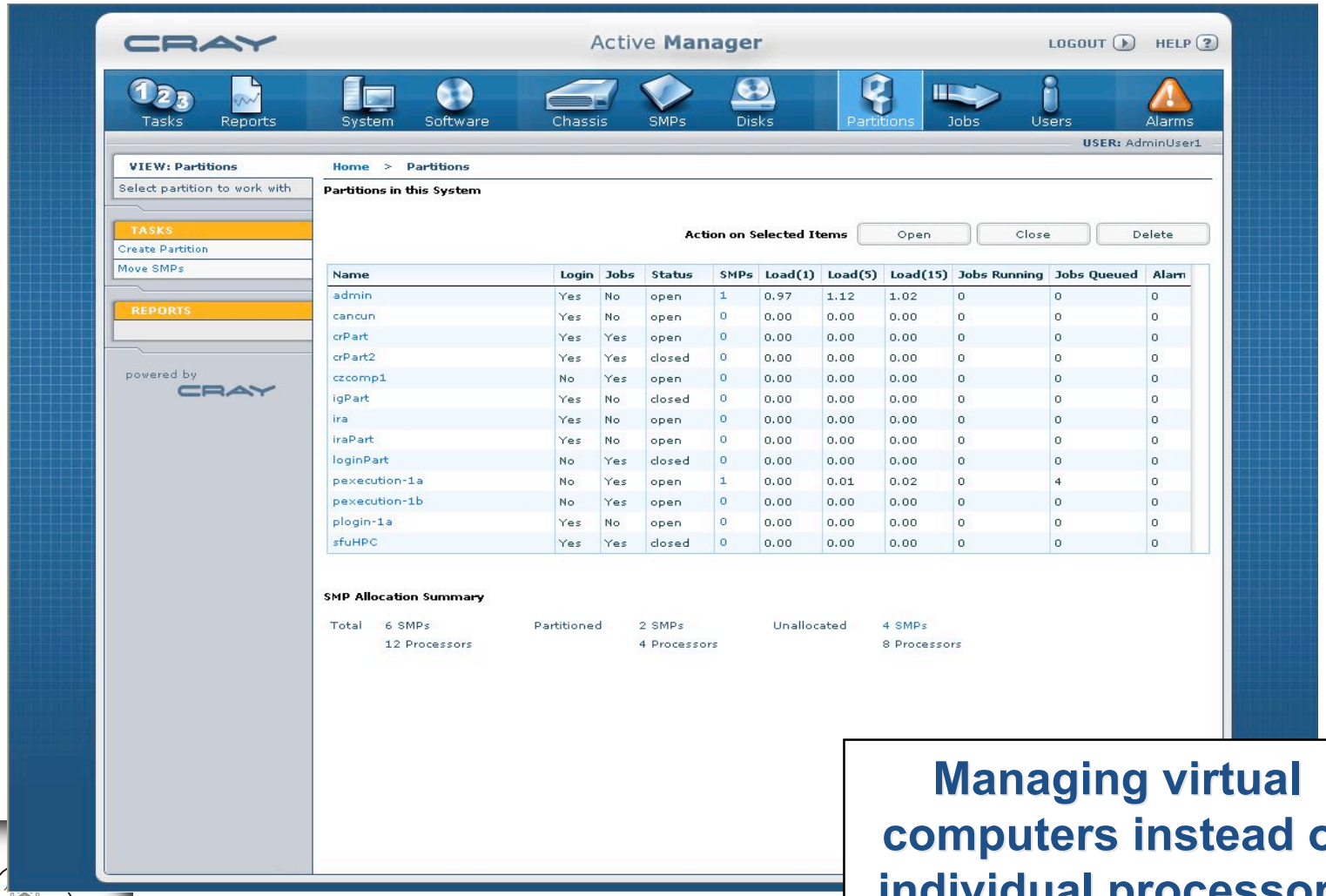
Linux Management using Partitions

- Linux RPM Install/Deployment/Audit
- Third Party Software
- Patches
- Disk Storage Management
- File System Management
- NFS Mount Administration
- Network Configuration (IP Addresses, DNS)
- User Authentication (LDAP, NIS)
- Workload Management Queue configuration



System Software Automatically Installed and Configured Consistently

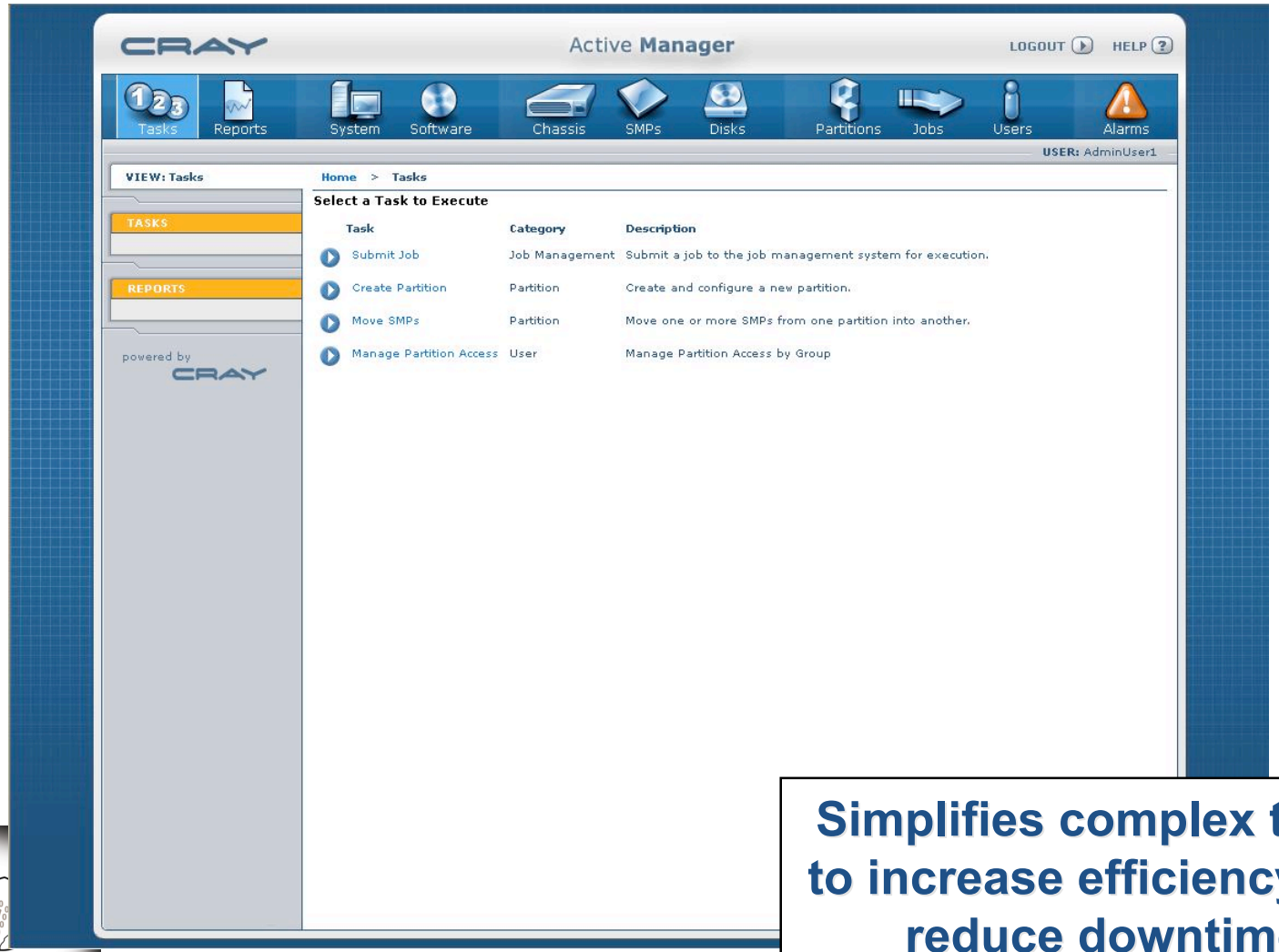
Active Manager System/Partition View



Managing virtual computers instead of individual processors

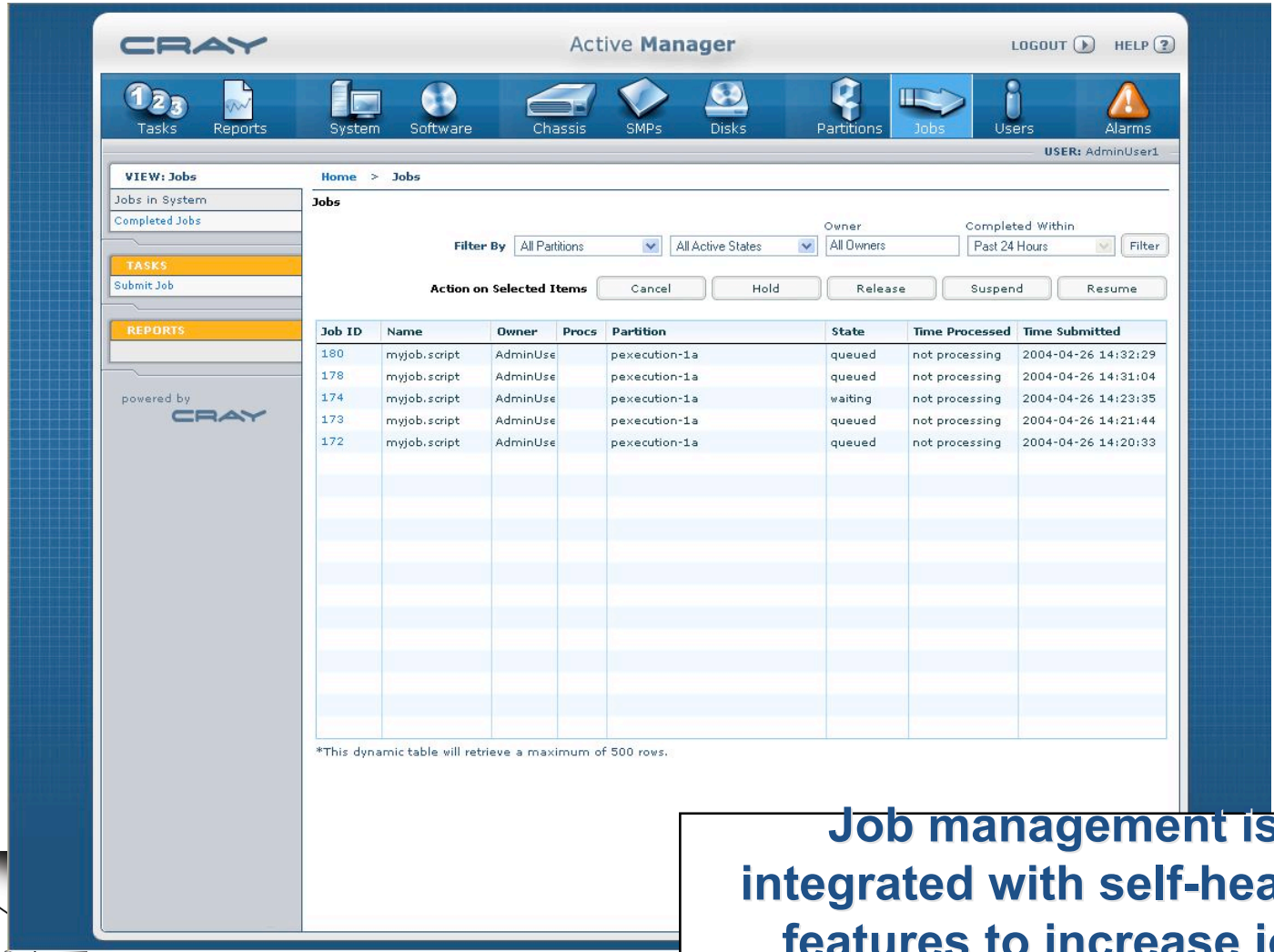
PETROGLYPHS TO PETAFLOPS
CUG 2005

Active Manager Task Wizard



Simplifies complex tasks to increase efficiency and reduce downtime

Active Manager Job Scheduler

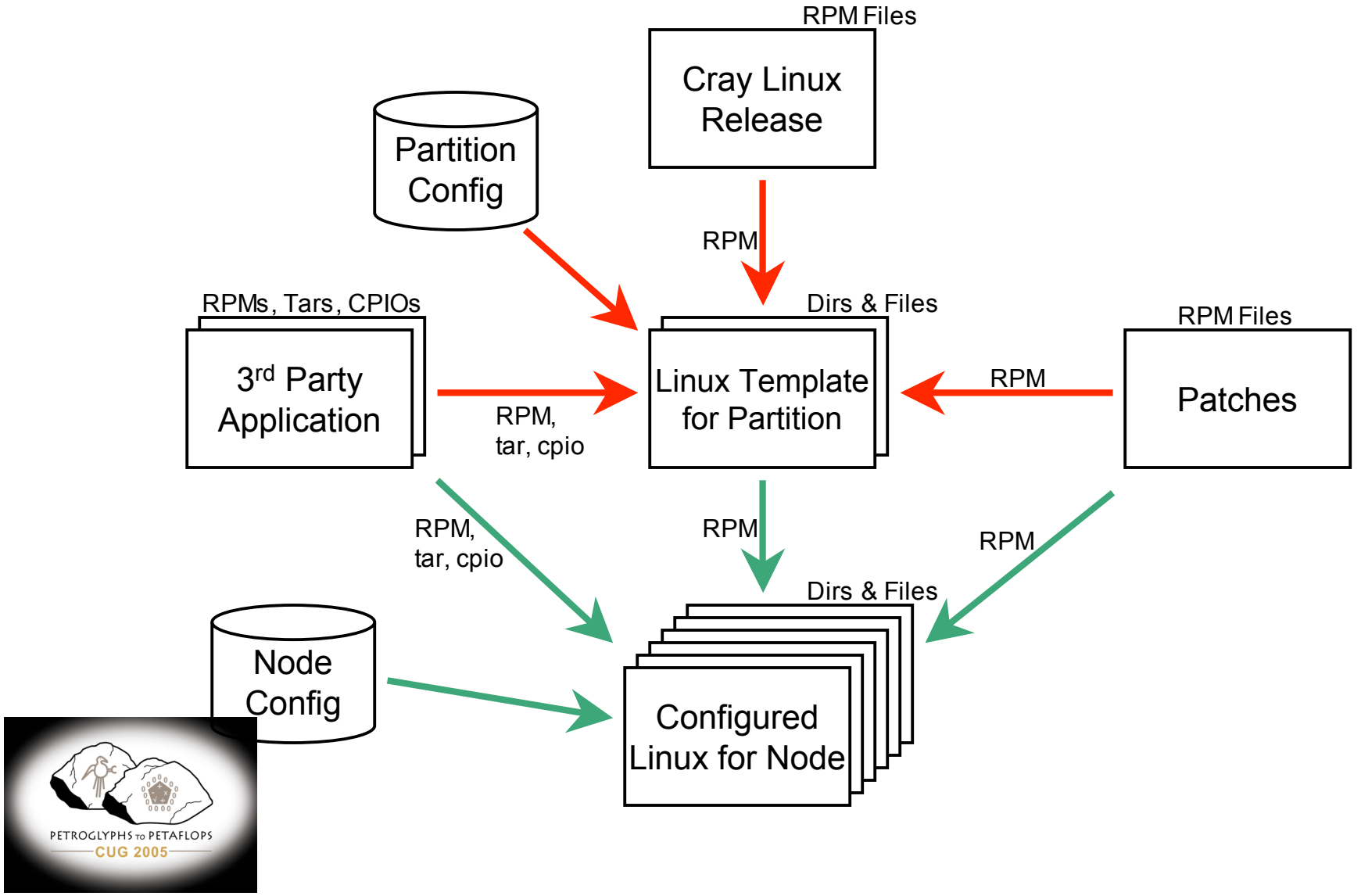


Job management is integrated with self-healing features to increase job completion rates

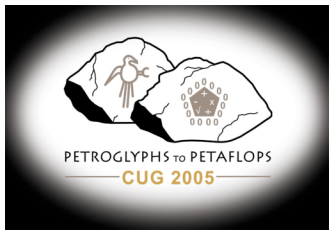
Active Manager User Portal

Intuitive user access for job submission and status lets users focus on applications, not computing

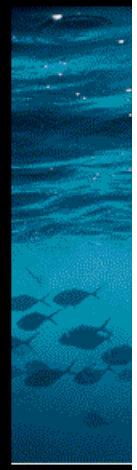
Software Deployment

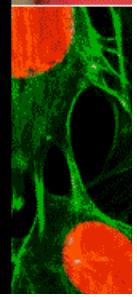


Partition Demo …



The Supercomputer Company



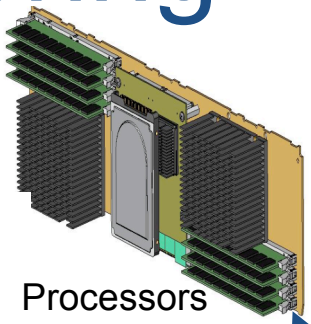


Customizing Event Handling

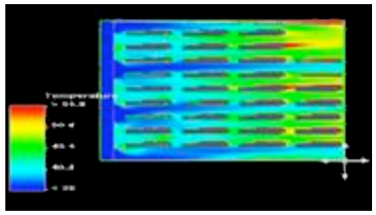
CRAY USER GROUP INCORPORATED

Self-Monitoring

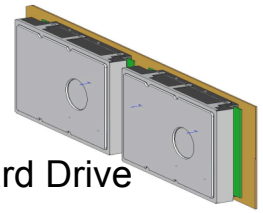
Parity
Heartbeat
Temperature
Fan speed
Diagnostics
Air Velocity
Voltage
Current



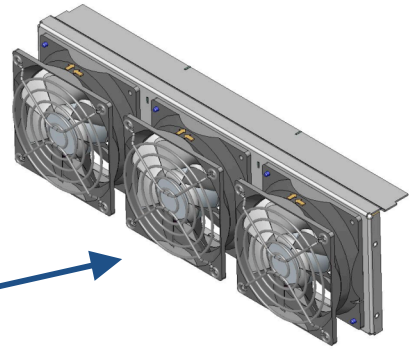
Processors
Memory



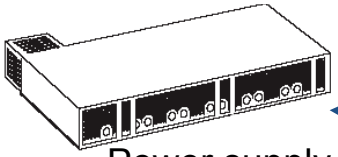
Thermals



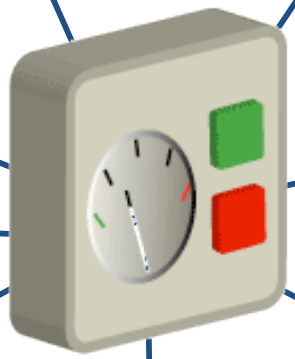
Hard Drive



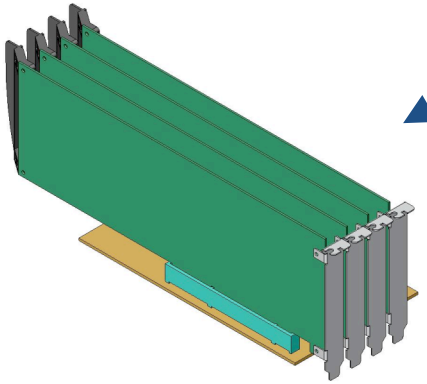
Fans

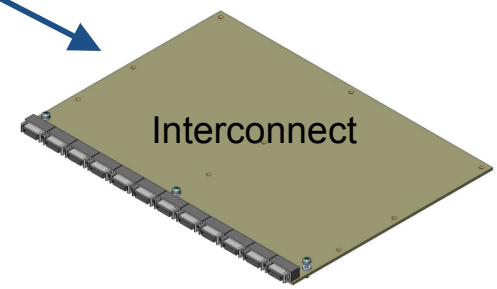


Power supply



Active Manager





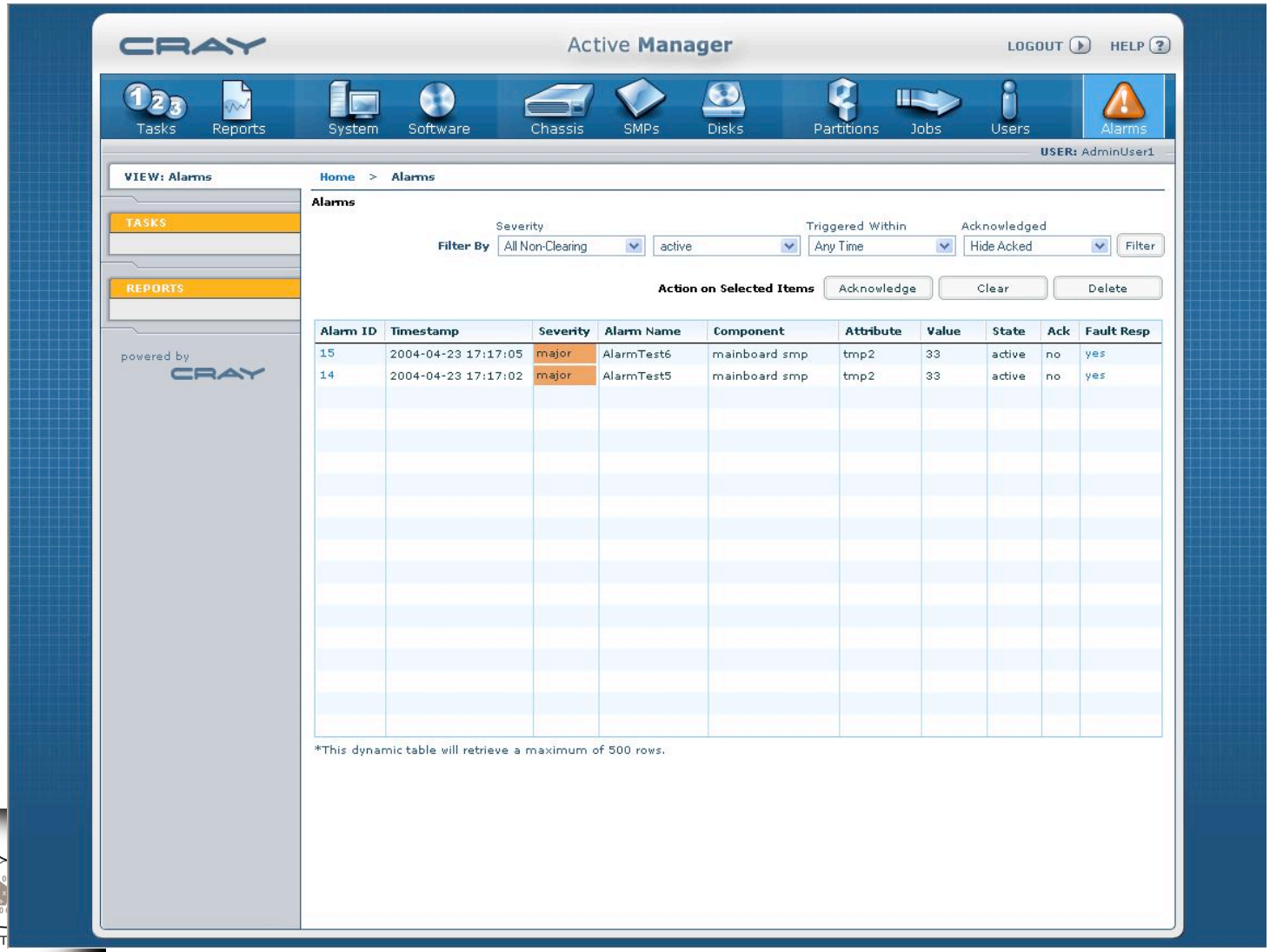
Interconnect

Power Supply

Dedicated Management Processor, OS, Fabric

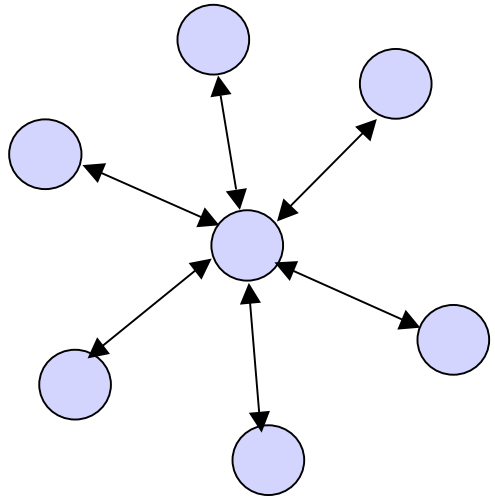
Active Manager Thermal Management

Active Manager Alarm Management



Building Alarm Scripts

- Events are published through a Central Database
- Associate Scripts with Alarms for Custom Handling

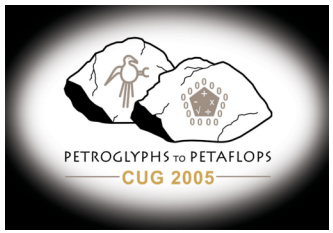


Centralized Event Handling

Through

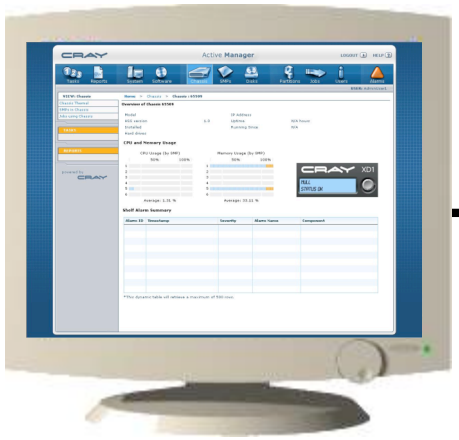
A “Publish/Subscribe”

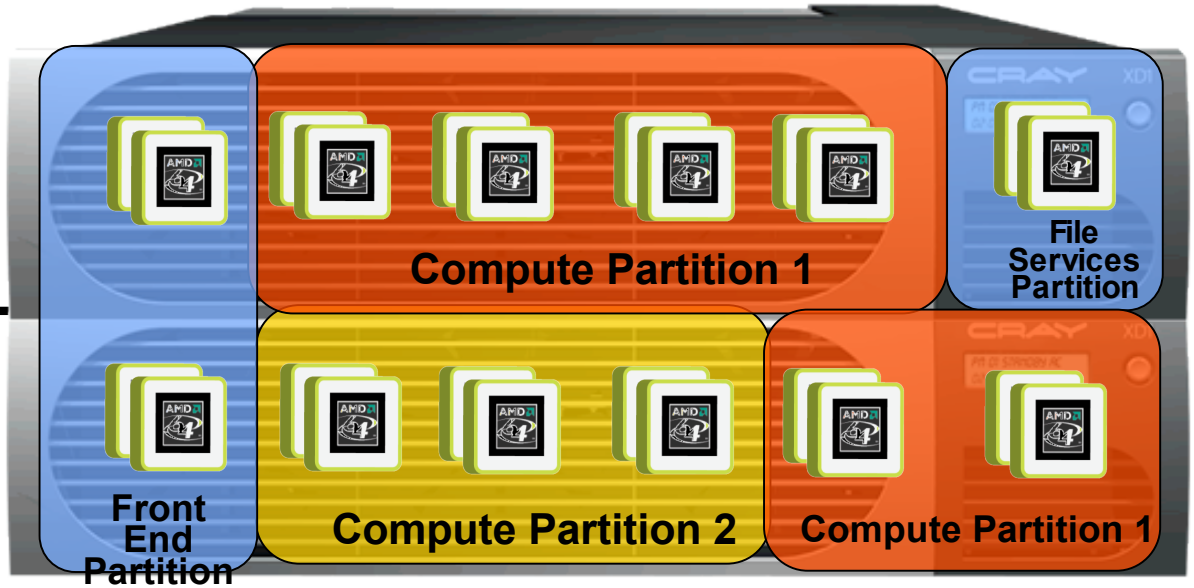
Event Model



Self-Healing

Users & Administrators



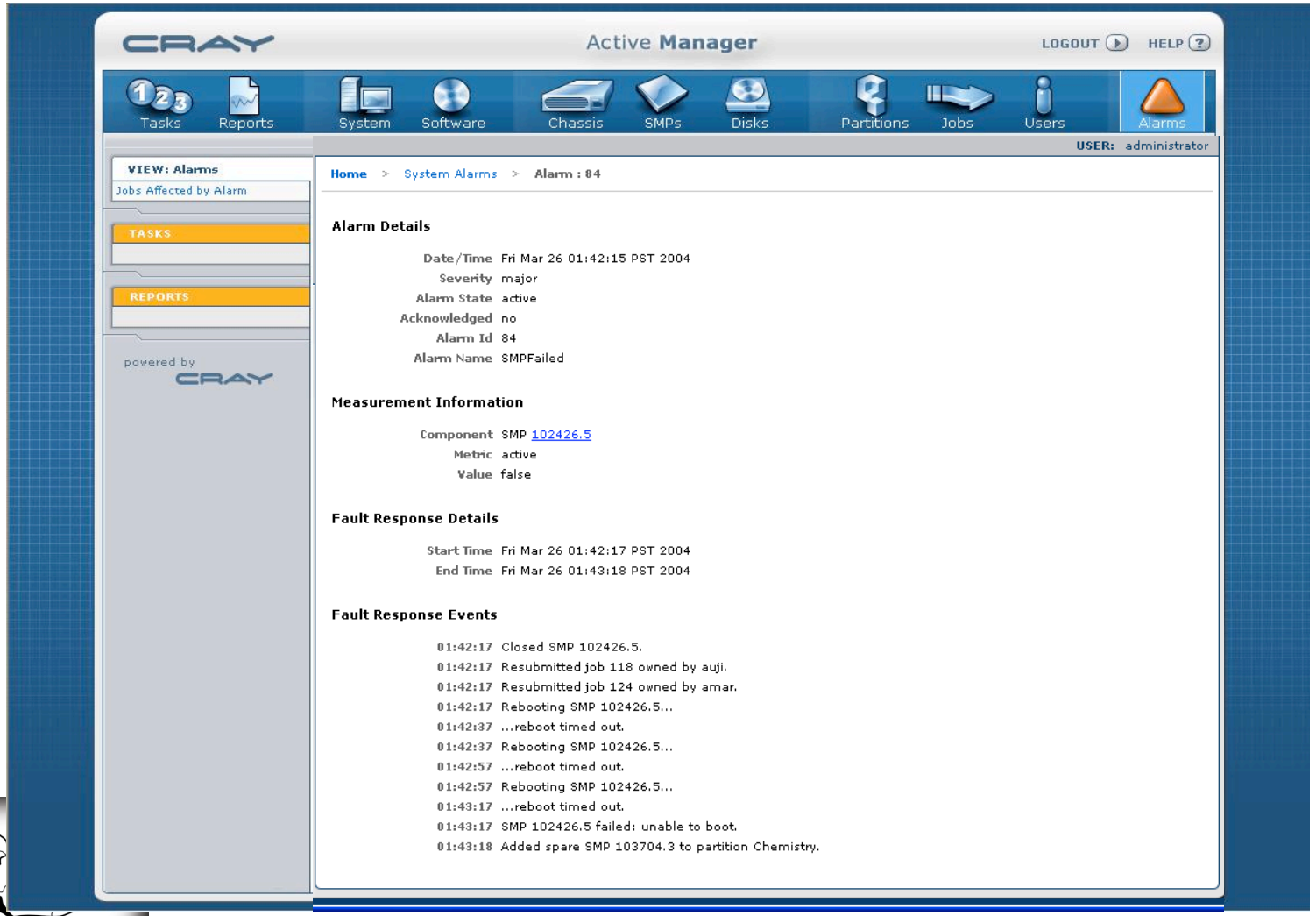


- Continuous Monitoring
- Detect (Future) Failure
- Attempt reset
- Isolate failed component
- Re-allocate resources (N+1 sparing or policy-based)

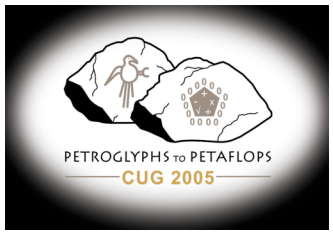
PETROGLYPHS TO PETAFLOPS
CUG 2005

Automated Recovery Reduces MTTR from hours to minutes

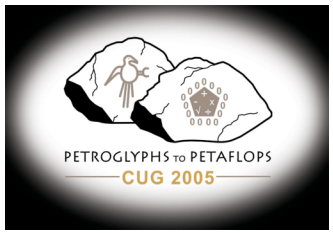
Active Manager Self Healing Policies



Event Handler Demo …



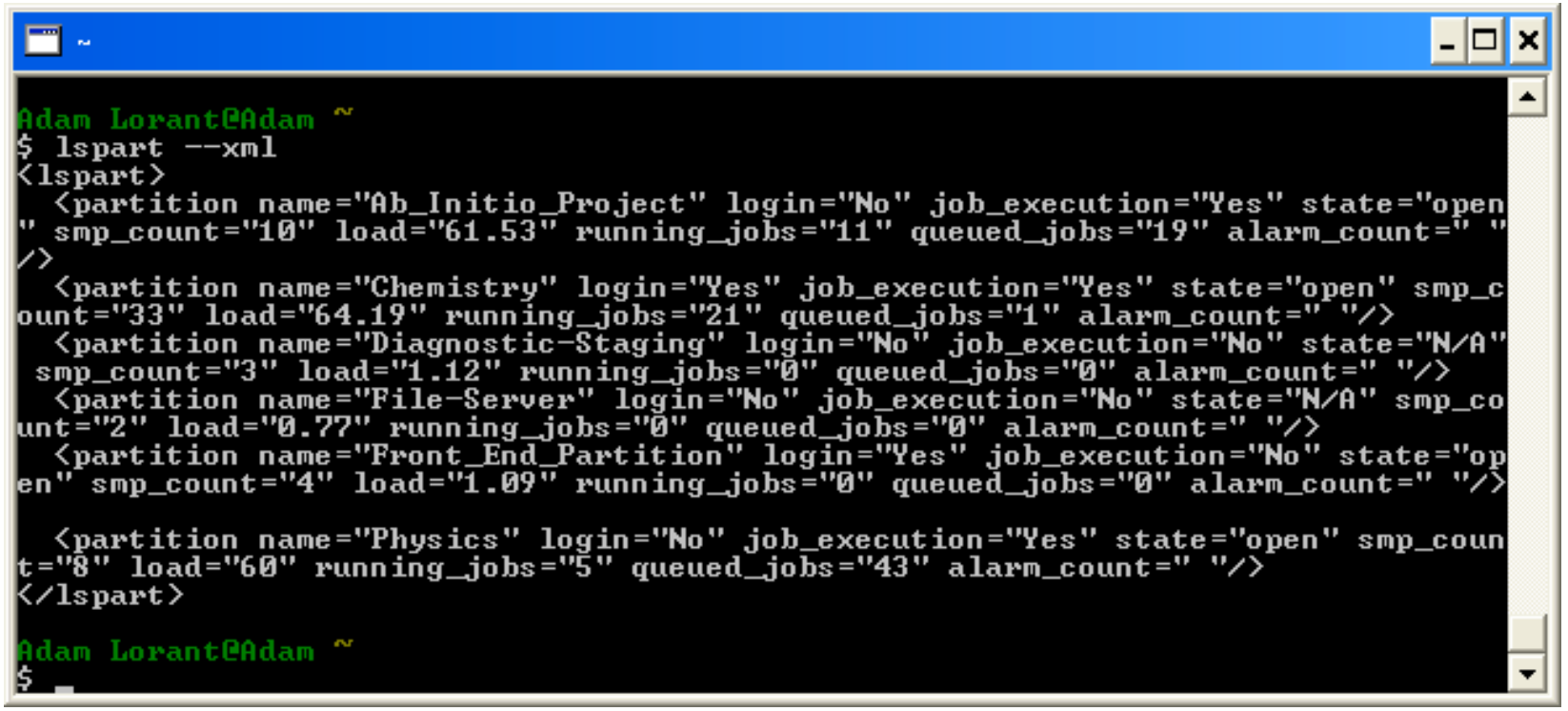
Partition Demo … (reprise)



XML Output from Commands

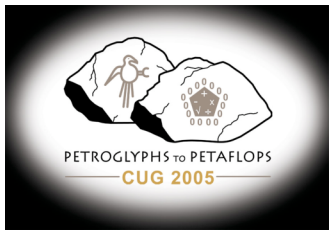
CRAY USER GROUP INCORPORATED

XML Integration Point



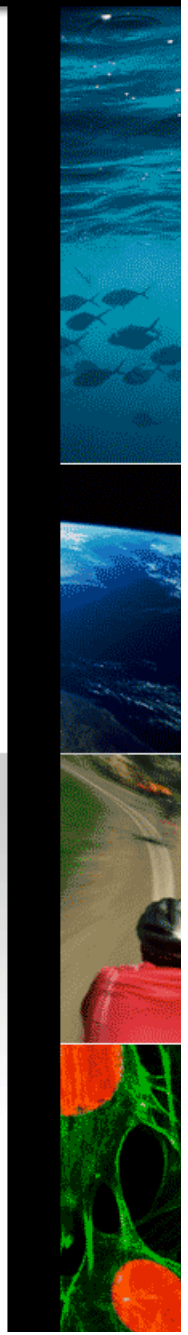
XML – designed to ease integration between different components by simplifying parsing of results

XML Demo …



CRAY The Supercomputer Company

Fabric Error Detection and Correction

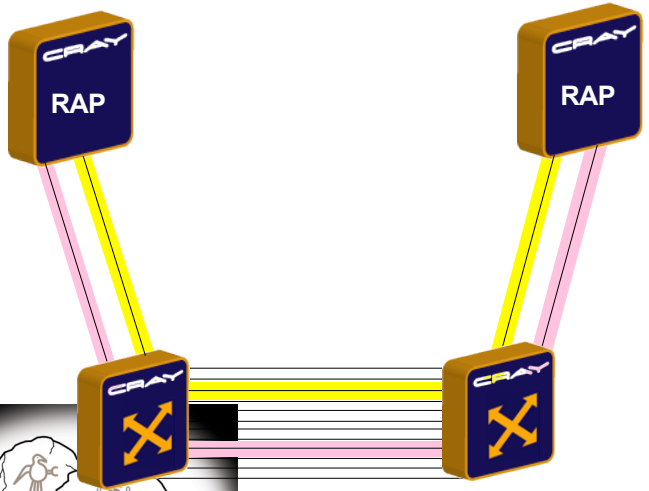


CRAY USER GROUP INCORPORATED

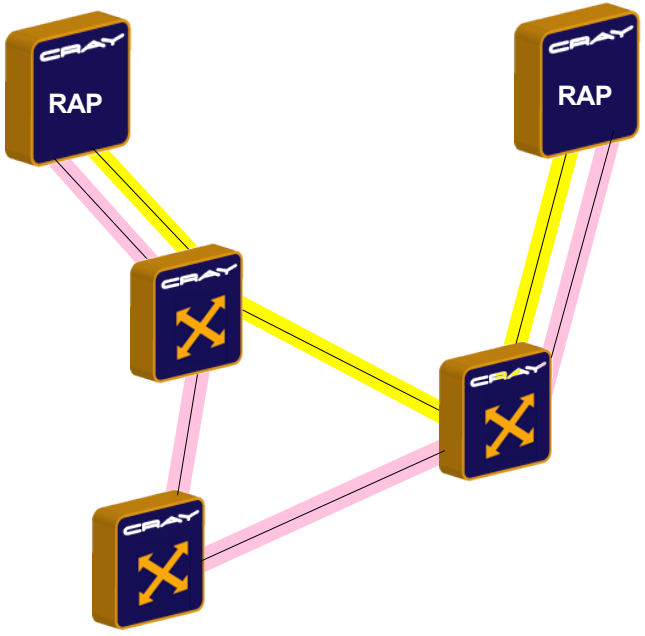
Load Sharing

- Every pair of RAPs have TWO paths between them for load sharing

eg: 2 chassis topology

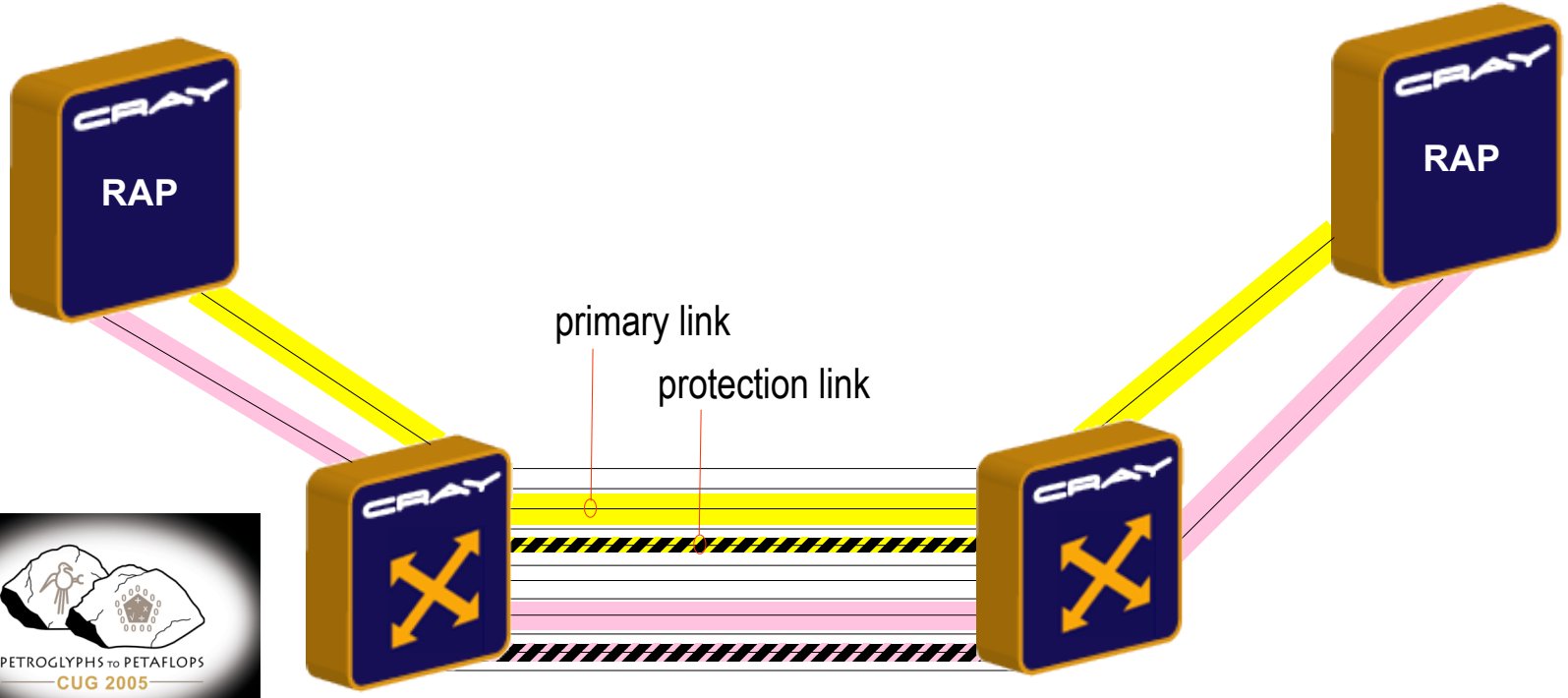


eg: 13 chassis topology

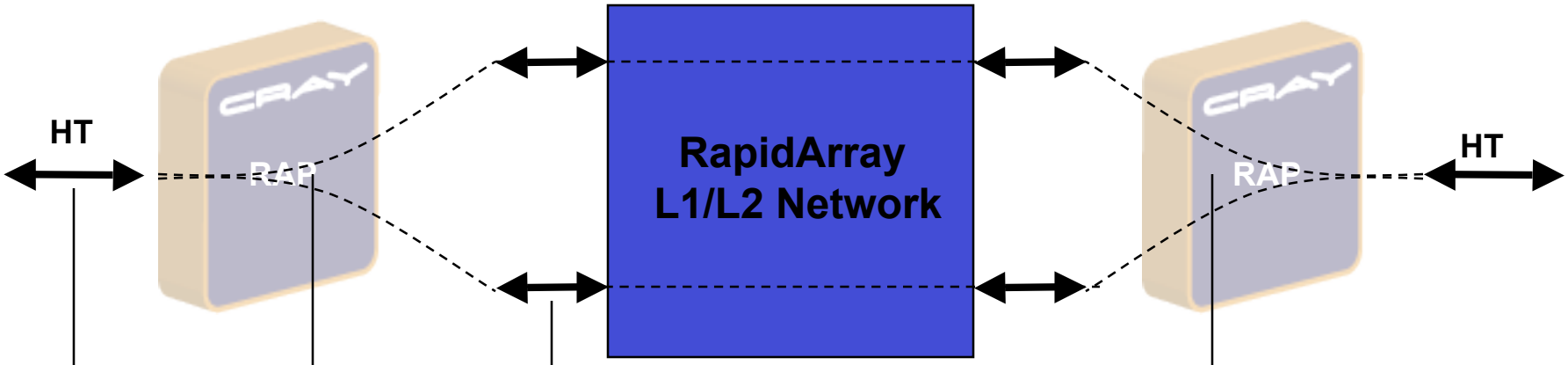


Path Redundancy

- Every RapidArray path is protected with a backup path
- Failover is automatic (alarm will be generated)
- Reliable Transport ensures no data loss



Network Data Integrity and Reliability



CRC Error Detection across all HT Links

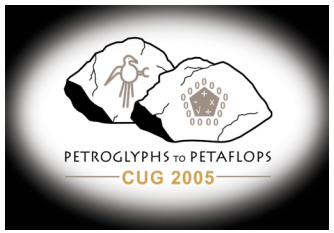
CRC Error Detection across all RA Links

End-to-End Reliable Transport
Retransmission on Lost Packets

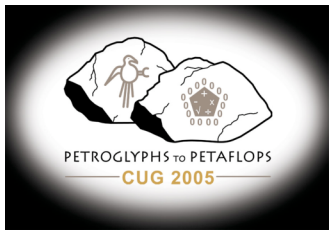
Fallback to single path routing if a path failure is detected

Verifying Fabric with Active Manager

- Present topology
  - Simply connect cables from diagrams
  - Active Manager will auto discover all links
- Expected topology
  - Defined in an XML file
  - Commands to test present vs. expected



Fabric Error Demo …



The Cray XD1

Cray XD1

- Built for price/performance
  - Interconnect bandwidth/latency
  - System-wide process synchronization
  - Application Acceleration FPGAs
- Standards-based
  - 32/64-bit X86, Linux, MPI
- High resiliency
  - Self-configuring, self-monitoring, self-healing
- Single system command & control
  - Intuitive, tightly integrated management software

Purpose-built and optimized for high performance workloads

PETROGLYPHS TO PETAFLOPS
CUG 2005