Platform™

Powering High Performance

Using & Managing HPC Systems

Presenter: William Lu, Ph.D., Marketing Director

Date: October 7, 2008

Platform

Challenges on HPC systems

Using and managing HPC systems, organizations are facing some challenges:

Complex hardware

- Performance relies on CPU, memory, storage, interconnects,…

Reliability

- Commodity hardware has good performance/cost, but…

Users competing resources

- Why my jobs are waiting?

COPYRIGHT ©2008 PLATFORM COMPUTING CORPORATION. ALL RIGHTS RESERVED.

Platform

IDC tracks HPC Management Software

IDC

Analyze the Future

WHITE PAPER

HPC Management Software: Reducing the Complexity of HPC Cluster and Grid Resources

Sponsored by: Platform Computing

Steve Conway
Earl C. Joseph, Ph.D.
Richard Walsh

May 2008

IDC OPINION

The typical architecture of a high-performance computing (HPC) resource is no longer a single custom-engineered, integrated hardware and software system with collections of vendor-ported and tested applications. Today, HPC resources are typically constructed à la carte from standards-based hardware and software component technologies. These resources are likely to include multiple systems at multiple sites. The variety and number of those components, their distinct providers, their potentially different locations, and the costliness of testing user applications against them all have driven up HPC resource management complexity and created an *integration gap*. Most HPC system configurations purchased today are unique assemblages of hardware and software. The growing task of integrating them into a reliable, productive, and secure working environment has fallen heavily on system administration and user support personnel.

The careful design, selection, and integration of the software components between the operating system (OS) and application layers of the software stack are increasingly vital for reducing the complexity of HPC cluster and grid management and for presenting an integrated and responsive HPC resource to users. Often informally referred to as *middleware*, the number of components in midstack HPC management software has grown significantly in the past 10 years. In IDC's opinion, this process has exposed a growing opportunity for vendors to add value to the HPC work cycle by adopting, streamlining, and modularizing the now long list of software functions assumed by HPC management software, including:

- User application compilation, debugging, and profiling
- User and system application input/output translation and scripting
- User application libraries, mathematical, IO, and parallel (e.g., MPI, OpenMP)
- Workload queuing, scheduling, and management (e.g., migration, checkpoint, restart)
- System and application installation, integration, and patching
- System management — servers, software, security, disk storage, and backups
- System and application monitoring, reporting, provisioning, reconfiguration, and failure detection

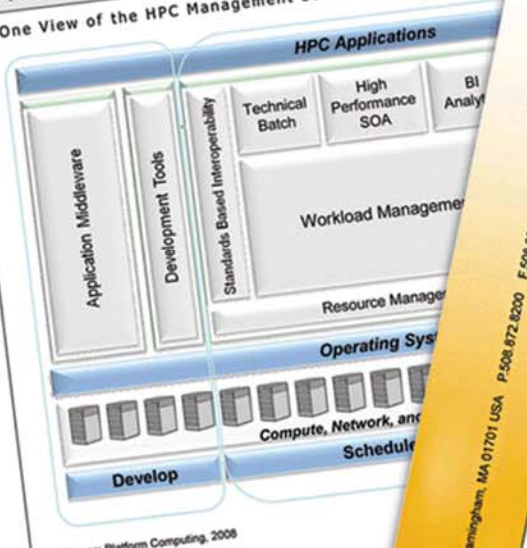
Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com

METHODOLOGY

The sources for this white paper on the origins, recent developments and importance of HPC management

FIGURE 1

One View of the HPC Management Software Stack



Source: Platform Computing, 2008

Is HPC management software middleware? Ye between the application and the OS as the dia HPC management software has been adapte and growing complexity of parallel HPC en anatomy of the HPC management software carefully articulated to be well understoo proper identification of new components evolving infrastructure underlying HPC. further transform HPC and become the to today's HPC management software

SITUAT

From its orig processing (T transactions p applications a diverse role. F of crucial linkin

HPC Man Middlewa

By virtue of this software compo system as middl the many distin working between shortcoming by software presente

IN TI

This wh software and cons elaborate

This whit out the e operating suggests not. It point on a path hardware a with manag managemer leading prov customers a

FIGU

Top R

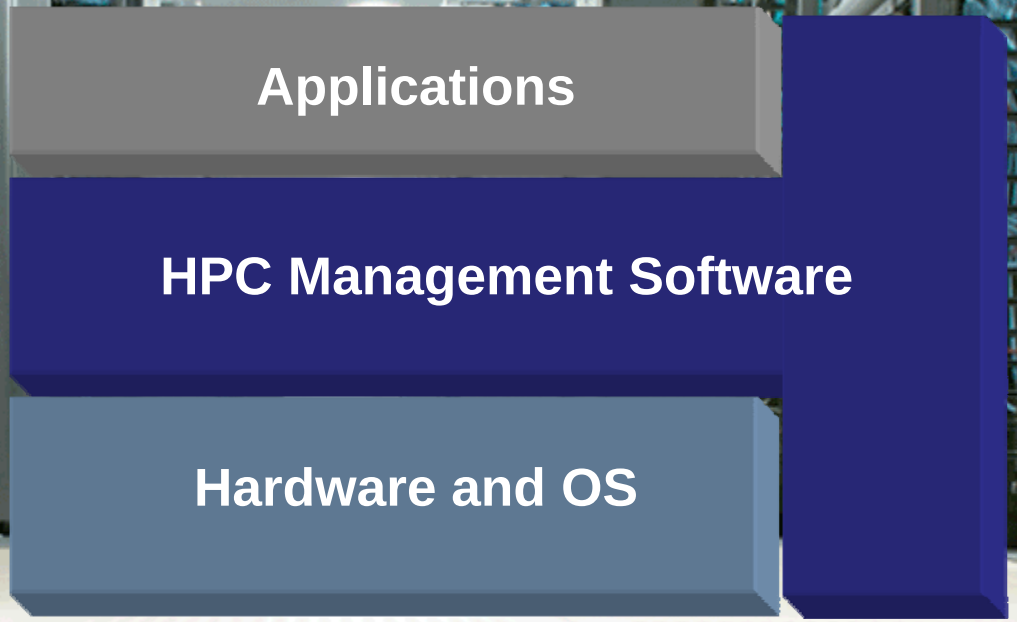
Abil

2

Platform

What is HPC Management Software?

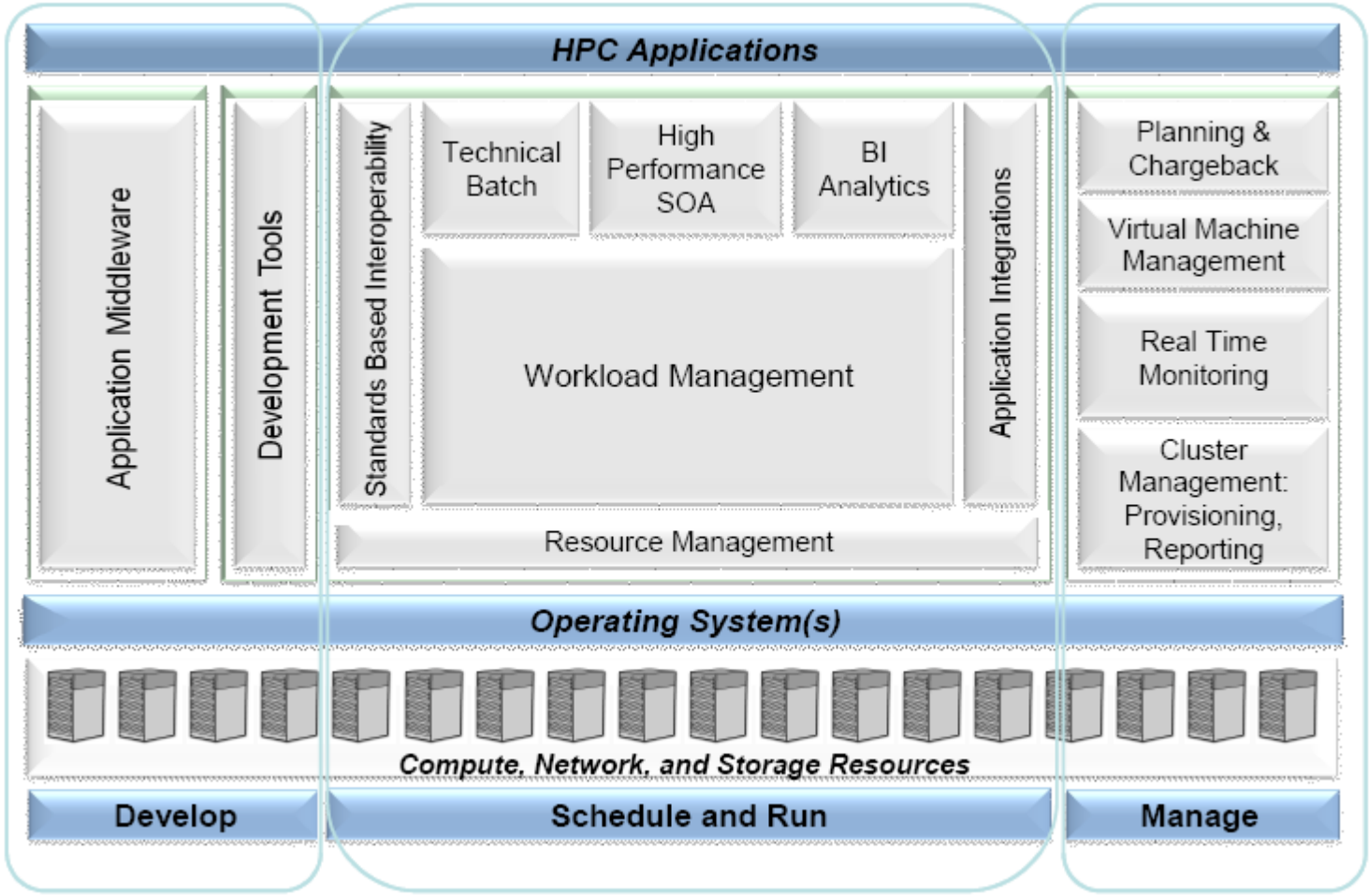
The software between the application and the operating system…



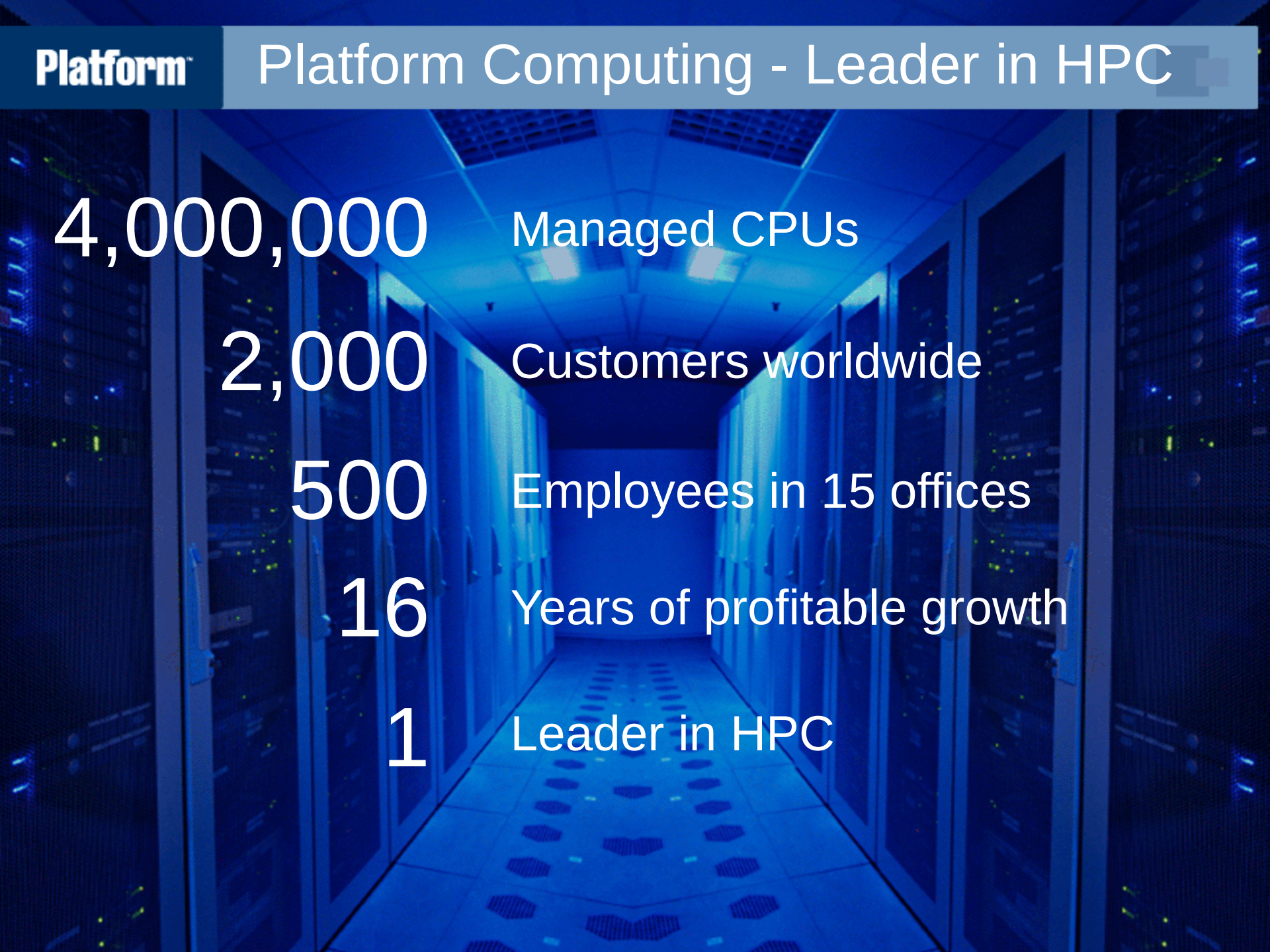
… that enables users to develop, run and manage compute or data intensive applications and manage HPC systems.

HPC Management Components

One View of the HPC Management Software Stack



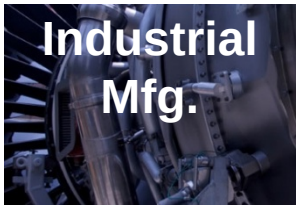
COPYRIGHT ©2008 PLATFORM COMPUTING CORPORATION. ALL RIGHTS RESERVED.


Platform
Platform Computing - Leader in HPC
4,000,000
Managed CPUs
2,000
Customers worldwide
500
Employees in 15 offices
16
Years of profitable growth
1
Leader in HPC

Industries Served by Platform

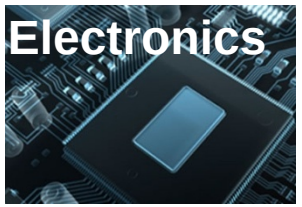
- CERN
- DoD, US
- DoE, US
- ENEA
- Georgia Tech
- Harvard Medical School
- Japan Atomic Energy Inst.
- MaxPlanck Inst.
- MIT
- SSC, China
- Stanford Medical
- TACC
- U. Of Georgia
- U. Tokyo
- Washington U.

- BNP
- Citigroup
- Fortis
- HSBC
- KBC Financial
- JPMC
- Lehman Brothers
- LBBW
- Mass Mutual
- MUFG
- Nomura
- Prudential
- Sal. Oppenheim
- Société Générale



- Airbus
- BAE Systems
- Boeing
- Bombardier
- Deere & Company
- Ericsson
- Honda
- General Electric
- General Motors
- Goodrich
- Lockheed Martin
- Nissan
- Northrop Grumman
- Pratt & Whitney
- Toyota
- Volkswagen

- Agip
- BP
- British Gas
- China Petroleum
- ConocoPhillips
- EMGS
- Gaz de France
- Hess
- Kuwait Oil
- PetroBras
- Petro Canada
- PetroChina
- Shell
- StatoilHydro
- Total
- Woodside



- AMD
- ARM
- Broadcom
- Cadence
- Cisco
- Infineon
- MediaTek
- Motorola
- NVidia
- Qualcomm
- Samsung
- Sony
- ST Micro
- Synopsys
- TI
- Toshiba

- Abott Labs
- AstraZeneca
- Celera
- DuPont
- Eli Lilly
- Johnson & Johnson
- Merck
- National Institutes of Health
- Novartis
- Partners Health Network
- Pharsight
- Pfizer
- Sanger Institute

Other Industries

AT&T Bell Canada Cingular DreamWorks Animation SKG GE
IRI Telecom Italia Telefonica Walt Disney Co.

COPYRIGHT ©2008 PLATFORM COMPUTING CORPORATION. ALL RIGHTS RESERVED.

Platform Solutions with Partners

Platform OCS 5 and Platform Manager integrated in Dell cluster systems

Platform LSF, Platform Manager form key parts of Unified Cluster Portfolio

Platform enterprise solutions support a wide range of IBM HPC systems

Platform delivers first certified Intel® Cluster Ready solution, Platform OCS 5

Integrates Platform LSF and Platform Symphony in grid solutions

Platform OCS 5 powers the Red Hat® HPC Solution

OEMs Platform's core technology in SAS® applications

COPYRIGHT ©2008 PLATFORM COMPUTING CORPORATION. ALL RIGHTS RESERVED.

Platform HPC Management Solutions

Develop	Schedule & Run	Manage

Platform Accelerate

Platform MPI
Platform LSF
Platform LSF MultiCluster
Platform LSF License Scheduler
Platform LSF Session Scheduler
Platform Process Manager
Platform EGO
EnginFrame

Platform Manage

Platform Manager
Platform Analytics
Platform RTM
Platform VM Orchestrator

Platform Open Cluster Stack 5: Accelerate & manage single Linux cluster

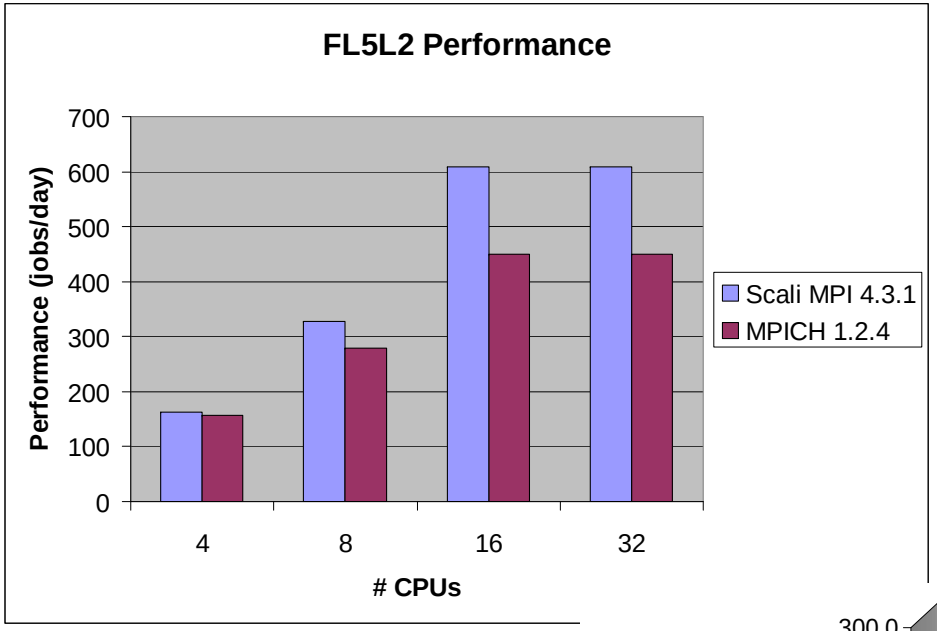
COPYRIGHT ©2008 PLATFORM COMPUTING CORPORATION. ALL RIGHTS RESERVED.

Accelerate Parallel Application – Platform MPI

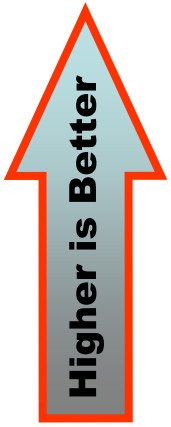
- Platform MPI contains *forward looking* MPI features:
  - Parallel Checkpoint/Restart
  - Dynamic affinity (load monitoring)
  - Windows CCS support (OS agnostic clusters)
  - Dynamic process creation
  - Network Failover
  - Replace run time (.so file) without recompile/relink

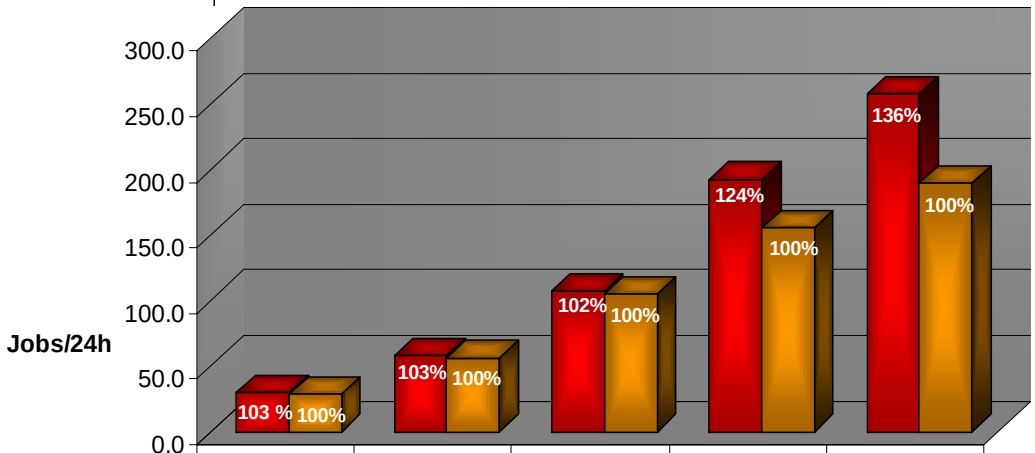
COPYRIGHT ©2008 PLATFORM COMPUTING CORPORATION. ALL RIGHTS RESERVED.

Platform MPI Benchmarks



Platform MPI out-performs by ~30%





	1	2	4	8	16
Scali MPI Connect 5.5	29.7	57.9	107.3	192.4	257.1
Commercial MPI	28.8	56.3	105.0	155.1	189.1

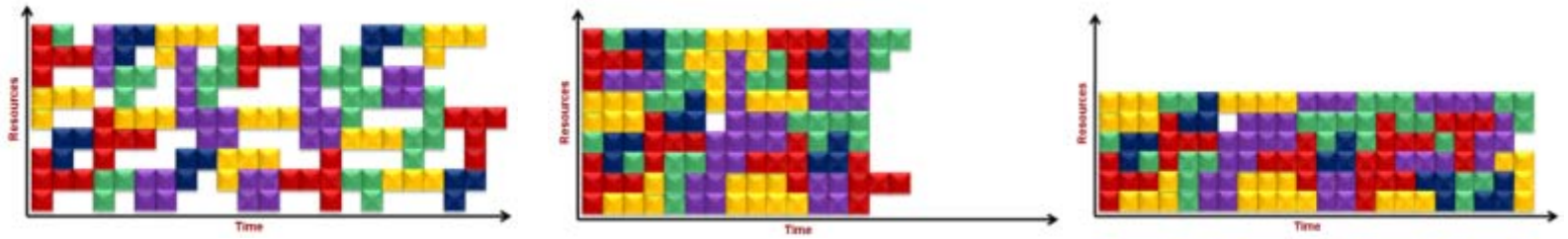
Number of nodes

COPYRIGHT ©2008 PLATFORM COMPUTING CORPORATION. ALL RIGHTS RESERVED.

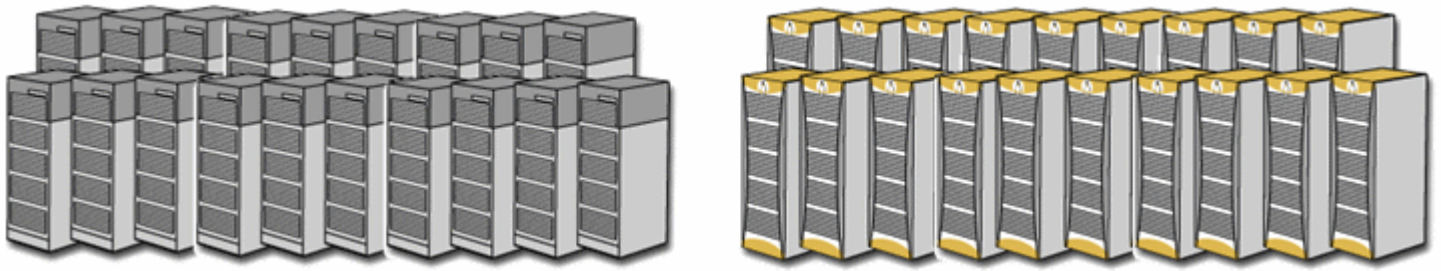
Powerful workload scheduler – Platform LSF

- Robust, fault tolerance, and high performance
- Powering mission critical environment
- Special price for education

By scheduling workloads intelligently according to policy, Platform LSF reduces application run-time and optimizes resource use.



VIRTUALIZED VIEW OF COMPUTE, NETWORK AND STORAGE RESOURCES



COPYRIGHT ©2008 PLATFORM COMPUTING CORPORATION. ALL RIGHTS RESERVED.

Customers using Platform LSF

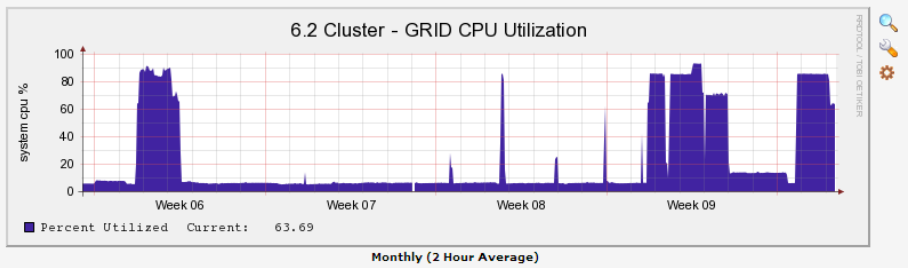
Platform Computing Simplifies

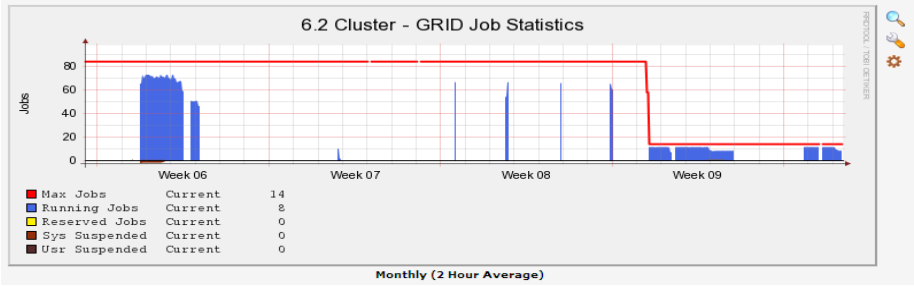
- Why we use Platform products: simplicity and flexibility
  - User base: hundreds of heterogeneous users
  - Mixed modes: from EP to MPI
  - Flexibility: open API
  - Monitoring: reporting and analytics
  - Ease of use: simple and clearly documented

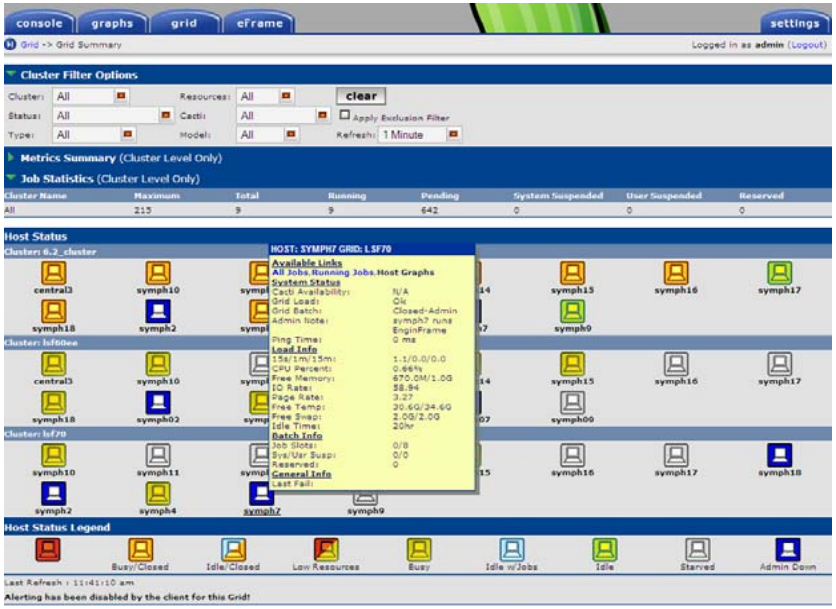
COPYRIGHT ©2008 PLATFORM COMPUTING CORPORATION. ALL RIGHTS RESERVED.

Centralized monitoring/reporting – Platform RTM

- Tool for LSF Administrators and Operators
- Near Real-Time Monitoring for LSF and extendable to other physical devices and application software
- Drill-down to Individual job statistics
- It allows you to graphically see your clusters job history without the use of b-commands
- Based on Open Source Cacti
- *Platform RTM enables system administrators to make timely decisions for proactively managing compute resources against business demand …*







COPYRIGHT ©2008 PLATFORM COMPUTING CORPORATION. ALL RIGHTS RESERVED.

RTM Shows Detailed Job Statistics

General Information

Status:	DONE	**Jobid:**	646527
Ex Host:	symph11	**Jobname:**	airbag.deploy.dyna
User:	ddunlap	**Clustername:**	6.2_cluster
Start Time:	2006-11-29 12:27:00	**End Time:**	2006-11-29 13:12:00
Exit Status:	0		

Submit Detail

Project:	Dyna	**Submit Host:**	symph4
Queue:	normal	**Submit Time:**	2006-11-29 12:24:00
Resource Requirements:	select[mem>=299 && swp>=299]rusage[mem=299:duration=10,swp=299:duration=10]		
Dependencies:			
Submit Command:	#!/bin/sh;#BSUB -P "Dyna";#BSUB -J "airbag.deploy.dyna";#BSUB -q normal;#BSUB -n 1;###BSUB -R "select[mem>=299 && swp>=299 && max_scratch>=3...		

Pending/Suspended Reasons

Reasons:

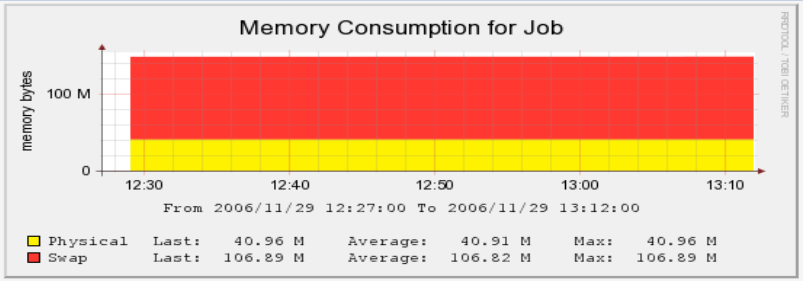
Execution Environment

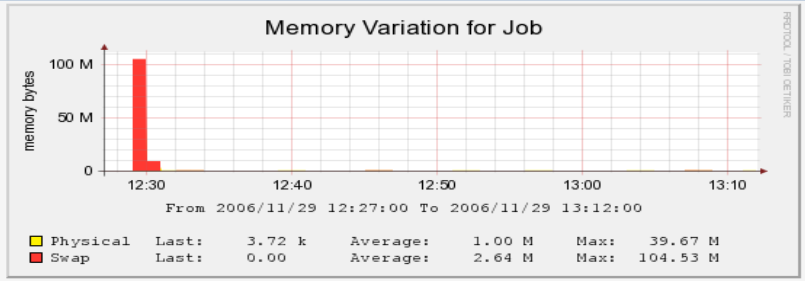
User Home:		**Working Dir:**	
UID String:	0	**Username:**	

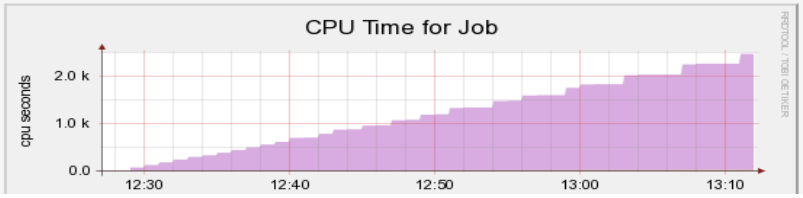
Current/Last Status

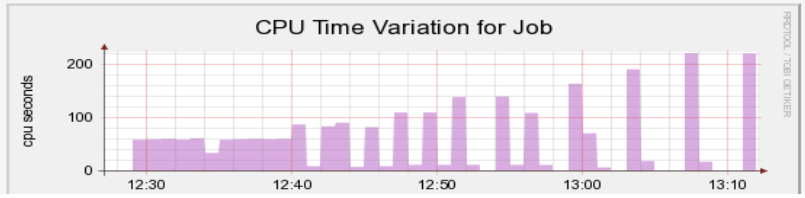
Cumulative CPU:	43.27 Minutes	**System Time:**	11 Seconds	**User Time:**	43.08 Minutes
Mem Used:	40.957M	**Swap Used:**	106.891M		
PGIDS:	16464	**PIDS:**	16464, 16465, 16468, 16470	**Threads:**	4

Graphical History







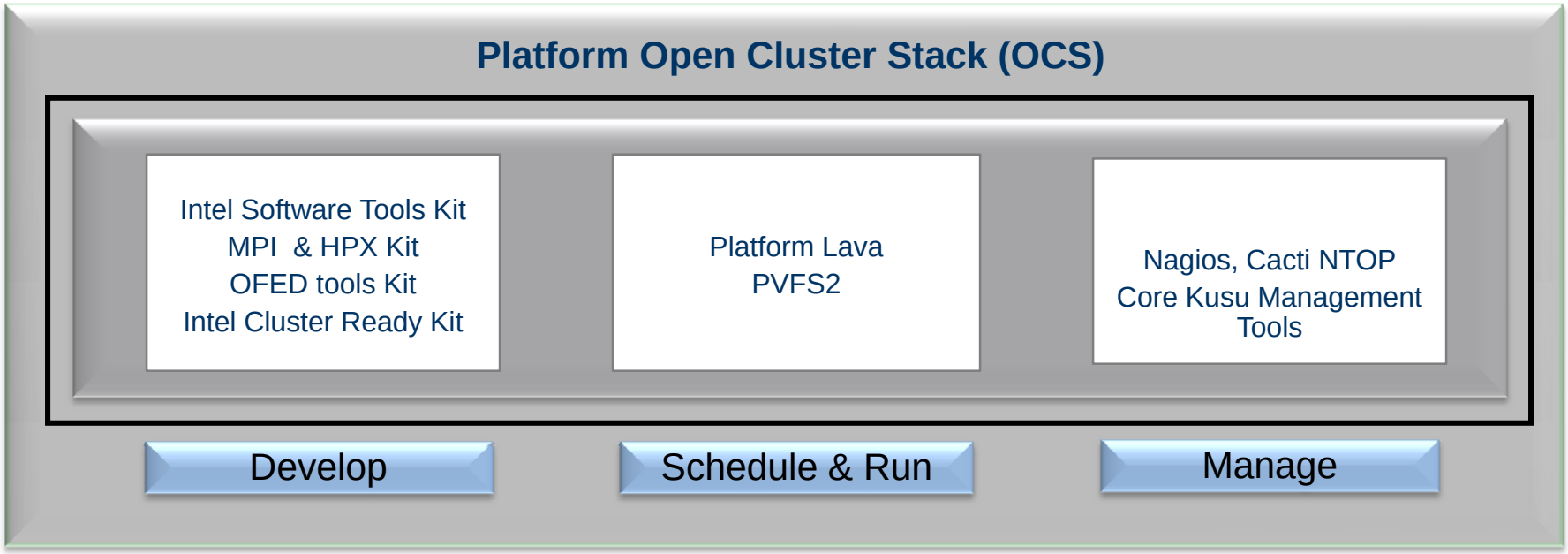


COPYRIGHT ©2008 PLATFORM COMPUTING CORPORATION. ALL RIGHTS RESERVED.

Single Cluster All-in-One Solution

- Simplify HPC cluster

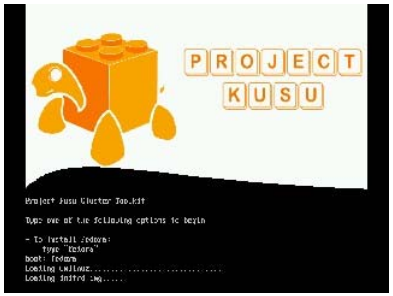
HPC Management Software –
for Single Linux Cluster



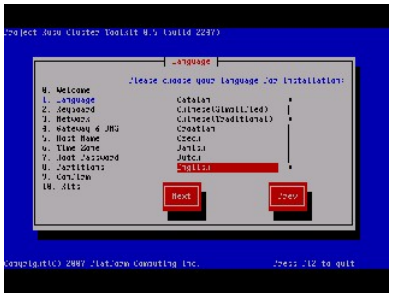
Red Hat HPC Solution = RHEL + Platform OCS 5

COPYRIGHT © 2008 PLATFORM COMPUTING CORPORATION. ALL RIGHTS RESERVED.

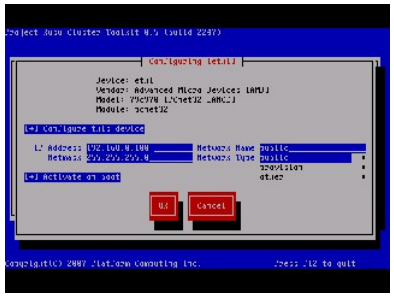
Creating a OCS 5 Installer Node



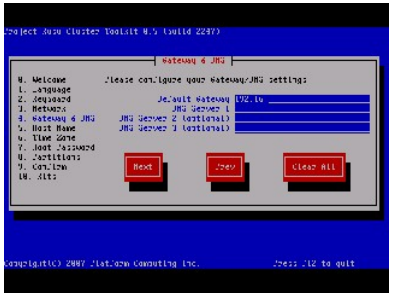
Boot from CD



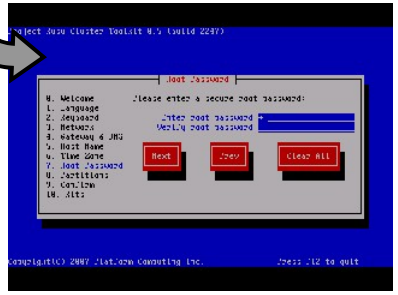
Choose Language



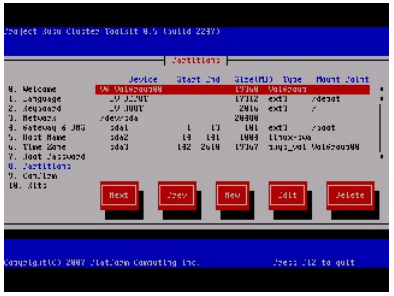
Configure Networks



DNS, & Gateway

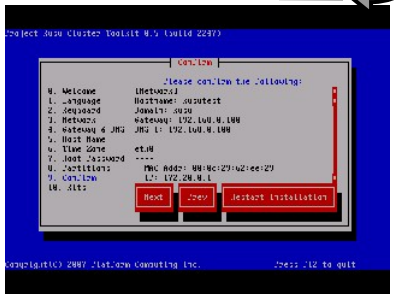


Root Password

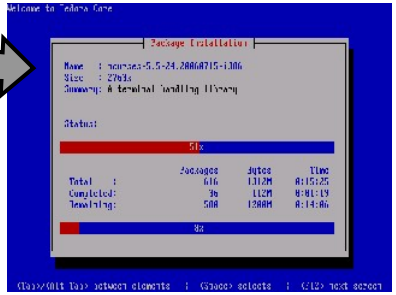


Partitioning & LVM

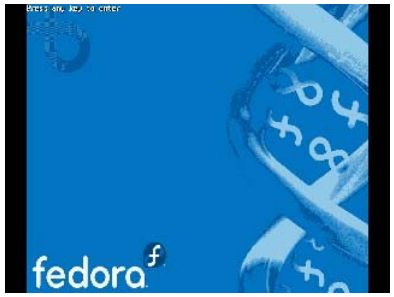
Adding Kits



Install Summary



Installing Packages



Installer Node Boot

Creating an Installer Node

Installation complete.
Installer node is ready to use.

COPYRIGHT ©2008 PLATFORM COMPUTING CORPORATION. ALL RIGHTS RESERVED.

Platform OCS 5 Features

Operating Systems:
- RHEL 4, 5*
- Fedora 6
- Fedora 7
- CentOS 5

Message Passing Libraries:
- MPICH 1,2
- MVAPICH 1
- OpenMPI
- Intel MPI*

Workload Management:
- Platform Lava
- Platform LSF*

Resource Management:
- Platform EGO*

Application Portals:
- NICE EngineFrame*

Certification Tools:
- Intel Cluster Checker+

Applications & Benchmarks:
- Atlas
- blacs
- bonnie++
- fftw
- hdf5
- iozone
- iperf
- linpack
- modules
- netcdf
- scalapack
- stream

Networking Hardware:
- Gigabit Ethernet
- Cisco Infiniband
- Qlogic Infiniband
- Mellanox Infiniband

Cluster Provisioning Methods:
- Package
- Diskless
- Imaged

Repository Management:
- Multiple repositories
- Repository Snapshots
- Full Dependency Checking
- Package management

Cluster Update:
- Update from RHN
- Update from Yum repository
- Update individual rpms

Configuration Management:
- Node Templates/Classes
- Cluster File Synronization
- Driver Management
- Cluster Network Configuration

Cluster Administration
- *Cacti Configured to monitor*
  - Memory on a node
  - CPU Load on a node
  - Logged In users
  - Processes on a node

logwatch configured to monitor activity on the cluster

Pdsh for administering nodes on the cluster

- *Nagios configured to manage & Notify*
  - Nodes in the cluster
  - Web server
  - Load
  - logged in users
  - DNS Server status
  - MySQL Server status
  - NTP daemon status
  - Node 'PING' status
  - Root Partition disk
  - SSH Daemon
  - SWAP Space
  - Total Processes on node
  - Email Notifications to Administrator

Plone CMS containing:
- Manual pages for OCS
- Documentation for
- Platform OCS 5
- List of installed Kits on the cluster

*Optional Commercial products that must be purchased separately

+Available only from Intel Cluster Ready Partners

COPYRIGHT ©2008 PLATFORM COMPUTING CORPORATION. ALL RIGHTS RESERVED.

Platform Technology Roadmap

- Workload based dynamic power management
  - Turn on/off machine
  - Avoid hot spot
  - Slow down CPU speed when MPI task is waiting.
- Workload driven dynamic provisioning
  - Switch machines between Windows and Linux automatically based on workload
- “Cloud” Toolkit
  - User web portal
  - Dynamic provisioning
  - Workload scheduling
  - Accounting and billing

COPYRIGHT ©2008 PLATFORM COMPUTING CORPORATION. ALL RIGHTS RESERVED.

World-class Support & Services

24x7 Support across the globe

"Platform has been proactive, involved and very, very friendly in providing support."

Henry Neeman
Director, Oklahoma University Supercomputing Centre

"Platform's standard of support has been excellent."

Tim Cutts
Platform LSF Administrator
Sanger Institute

COPYRIGHT ©2008 PLATFORM COMPUTING CORPORATION. ALL RIGHTS RESERVED.

Platform

Summary

- Platform's Value proposition
  - Total solution for HPC management
  - 16 years of HPC experiences and broader customer base
  - Top quality technical support and services

COPYRIGHT ©2008 PLATFORM COMPUTING CORPORATION. ALL RIGHTS RESERVED.

Platform™

Powering High Performance

www.platform.com

info@platform.com

1-877-528-3676 (1-87-PLATFORM)