



Cluster Distributions Overview

E. Imamagić, D. D. Žagar

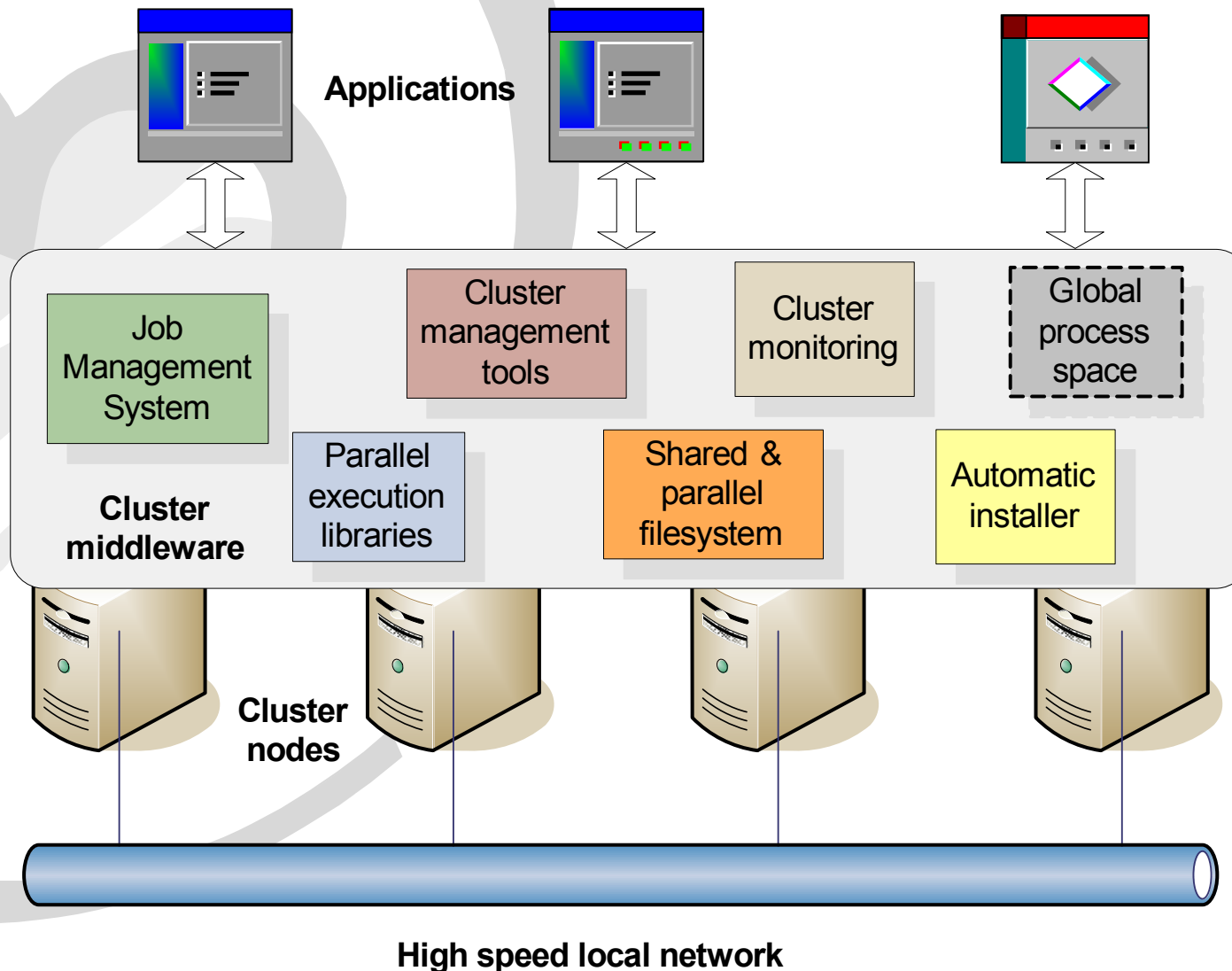
Overview

- Introduction
- Computer cluster
- Cluster middleware
- Cluster distribution
- Cluster distributions overview
- Conclusion

Computer cluster

- Computer cluster
 - Interconnected computers working as single
- High speed local networks
 - Fast Ethernet, Gigabit Ethernet, Myrinet, Infiniband, ...
- Cluster nodes
 - SMPs, workstations, blades, ...
- Cluster middleware
 - Software layer above node OS
 - Enables nodes cooperation

Introduction



Cluster middleware

- Job Management System
 - User interface for submitting and controlling jobs
 - Resource usage policies
 - Cluster nodes utilization
 - Users' jobs information
 - Examples: SGE, PBS, Torque, Condor
- Cluster Monitoring
 - State of resources
 - Resource usage information
 - Examples: Ganglia, Clumon, Supermon, ParmonJob Management System

Cluster middleware (cont.)

- Shared and parallel filesystems
 - Remote file systems
 - Parallel file access (for large files)
 - Examples:
 - Shared: NFS, AFS, Coda
 - Parallel: Lustre, Sestina, PVFS
- Parallel execution libraries
 - Inter-node communication
 - Process dispatching
 - Examples: MPI, PVM

Cluster middleware (cont.)

- Cluster management tools
 - Nodes administration
 - Examples: C3, Rocks tools
- Automatic Installers
 - Software stack & OS deployment on nodes
 - Examples: SystemImager, RedHat Kickstart, FAI, LCFG
- Global process space
 - Global process space on OS kernel level
 - Examples: BProc, KSIX, SCore-D

Cluster distribution

- What is cluster distribution?
 - Integrated cluster middleware
 - OS distribution (in some cases)
- Why do we need cluster distribution?
 - Cluster installation & administration
 - Integration of cluster middleware components
- Critical issues
 - Maintenance & software update
 - Users' feedback & support

Cluster Distributions Overview

	ROCKS	OSCAR	Warewulf	xCAT
JMS	SGE, Torque Condor, SQMS	OpenPBS, Torque	SGE	OpenPBS
Monitoring	Ganglia, SCMS	Ganglia, Clumon	Ganglia	Ganglia
S&P FS	NFS, PVFS	NFS, PVFS	NFS	NFS, PVFS
Parallel Libraries	MPICH (P4, MPD, GM), PVM	MPICH (P4, GM), LAM/MPI, PVM	LAM/MPI, PVM	MPICH (P4, GM), LAM/MPI, PVM
CMT	Rocks tools	C3		IBM Management
AI	Redhat kickstart	System Installation Suite		Redhat kickstart
GPS				

Cluster Distributions Overview (cont.)

	Scyld Beowulf	Clustermatic	OpenSCE	SCore
JMS	PBSPro		SQMS	OpenPBS, SGE
Monitoring		Supermon	SCMS, SCMSWeb	
S&P FS	NFS, PVFS, Lustre, Sestina			
Parallel Libraries	MPI, PVM	ZPL, LA-MPI	MPITH	MPICH, PVM, MPC++, OpenMP
CMT	BProc	BProc	SCMS	SCore-D
AI	BeoBoot	BeoBoot		
GPS	BProc	BProc	KSIX	SCore-D

Conclusion

- Best open source solutions
 - NPACI Rocks
 - Easy installation
 - Huge users community
 - Rich set of cluster middleware components
 - OSCAR
 - Graphical installation tool
 - Modular cluster middleware installation
 - Image based nodes installation
- Warewulf – promising new solution
- Scyld Beowulf – best commercial solution