Cluster Computing

Javier Delgado
Grid-Enabledment of Scientific Applications
Professor S. Masoud Sadjadi
Brief History of Clustering
(very brief)

- NOW pioneered the vision for clusters of commodity processors.
  - David Culler (UC Berkeley) started early 90’s
  - SunOS / SPARC
  - First generation of Myrinet, active messages
  - Glunix (Global Unix) execution environment
- Beowulf popularized the notion and made it very affordable.
  - Tomas Sterling, Donald Becker (NASA)
  - Linux
Definition: HPC Cluster Architecture

Frontend Node

Private Ethernet Network

Public Ethernet

Application Network (Optional)

Node

Node

Node

Node

Node

Node

Node

Power Distribution (Net addressable units as option)

© 2006 UC Regents
Essence of a Beowulf

- Hardware
  - One head/master node
  - (Several) compute nodes
  - Interconnection modality (e.g. ethernet)

- Software
  - Parallel Programming Infrastructure
  - Scheduler (optional)
  - Monitoring application (optional)
Scheduling

- Multiple users fighting for resources = bad
  - Don't allow them to do so directly
- Computer users are greedy
  - Let the system allocate resources
- Users like to know job status without having to keep an open session
Cluster Solutions

- Do-it-yourself (DIY)
- OSCAR
- Rocks
- Pelican HPC (formerly Parallel Knoppix)
- Microsoft Windows CCE
- OpenMosix (closed March 2008)
- Clustermatic (no activity since 2005)
DIY Cluster

- Advantages
  - Control
  - Learning Experience
- Disadvantages
  - Control
  - Administration
DIY-Cluster How-To Outline

- Hardware Requirements
- Head Node Deployment
  - Core Software Requirements
  - Cluster-specific Software
  - Configuration
- Adding compute nodes
Hardware Requirements

- Several commodity computers:
  - cpu/motherboard
  - memory
  - ethernet card
  - hard drive (recommended, in most cases)
- Network switch
- Cables, etc.
Software Requirements – Head node

- Core system
  - system logger, core utilities, mail, etc.
  - Linux Kernel
    - Network Filesystem (NFS) server support
- Additional Packages
  - Secure Shell (SSH) server
  - iptables (firewall)
  - nfs-utils
  - portmap
  - Network Time Protocol (NTP)
Software Requirements – Head node

- Additional Packages (cont.)
  - inetd/xinetd – For FTP, globus, etc.
  - Message Passing Interface (MPI) package
  - Scheduler – PBS, SGE, Condor, etc.
  - Ganglia – Simplified Cluster “Health” Logging
    - dependency: Apache Web Server
Initial Configuration

- Share /home directory
- Configure firewall rules
- Configure networking
- Configure SSH
- Create compute node image
Building the Cluster

- Install compute node image on the compute node
  - Manually
  - PXE Boot (pxelinux, etherboot, etc.)
  - RedHat Kickstart
  - etc.
- Configure host name, NFS, etc.
- ... for each node!
Maintainance

- Software updates in head node require update in compute node
- Failed nodes must be temporarily removed from head node configuration files
Building the Cluster

- But what if my boss wants a 200-node cluster?
  - Monster.com
  - OR come up with your own automation scheme
  - OR Use OSCAR or Rocks
Cluster Solutions

- Do-it-yourself (DIY)
- OSCAR
- Rocks
- Pelican HPC (formerly Parallel Knoppix)
- Microsoft Windows CCE
- OpenMosix (closed March 2008)
- Clustermatic (no activity since 2005)
OSCAR

- Open Source Cluster Application Resources
- Fully-integrated software bundle to ease deployment and management of a cluster

Provides
  - Management Wizard
  - Command-line tools
  - System Installation Suite
Overview of Process

- Install OSCAR-approved Linux distribution
- Install OSCAR distribution
- Create node image(s)
- Add nodes
- Start computing
OSCAR Management Wizard

- Download/install/remove OSCAR packages
- Build a cluster image
- Add/remove cluster nodes
- Configure networking
- Reimage or test a node with the Network Boot Manager
OSCAR Command Line tools

- Everything the Wizard offers
- yume
  - Update node packages
- C3 - The Cluster Command Control Tools
  - provide cluster-wide versions of common commands
  - Concurrent execution
  - example 1: copy a file from the head node to all visualization nodes
  - example 2: execute a script on all compute nodes
C3 List of Commands

- cexec: execution of any standard command on all cluster nodes
- ckill: terminates a user specified process
- cget: retrieves files or directories from all cluster nodes
- cpush: distribute files or directories to all cluster nodes
- cpushimage: update the system image on all cluster nodes using an image captured by the SystemImager tool
List of Commands (cont.)

- **crm**: remove files or directories
- **cshutdown**: shutdown or restart all cluster nodes
- **cnum**: returns a node range number based on node name
- **cname**: returns node names based on node ranges
- **clist**: returns all clusters and their type in a configuration file
Example c3 configuration

```bash
# /etc/c3.conf
##
# describes cluster configuration
##
cluster gcb {
  gcb.fiu.edu #head node
dead placeholder #change command line to 1 indexing
compute-0-[0-8] #first set of nodes
exclude 5 #offline node in the range (killed by J. Figueroa)
}
------
```
OPIUM

- The OSCAR Password Installer and User Management
  - Synchronize user accounts
  - Set up passwordless SSH
  - Periodically check for changes in passwords
SIS

- System Installation Suite
- Installs Linux systems over a network
- Image-based
- Allows different images for different nodes
- Nodes can be booted from network, floppy, or CD.
Cluster Solutions

- Do-it-yourself (DIY)
- OSCAR
- Rocks
- Pelican HPC (formerly Parallel Knoppix)
- Microsoft Windows CCE
- OpenMosix (closed March 2008)
- Clustermatic (no activity since 2005)
Rocks

- Disadvantages
  - Tight-coupling of software
  - Highly-automated
- Advantages
  - Highly-automated...
  - But also flexible
Rocks

- The following 25 slides are property of UC Regents
The Dark Side of Clusters

- Clusters are phenomenal price/performance computational engines …
  - Can be hard to manage without experience
  - High-performance I/O is still unsolved
  - Finding out where something has failed increases at least linearly as cluster size increases
- Not cost-effective if every cluster “burns” a person just for care and feeding
- Programming environment could be vastly improved
- Technology is changing very rapidly. Scaling up is becoming commonplace (128-256 nodes)
The Top 2 Most Critical Problems

◆ The largest problem in clusters is \textit{software skew}
  ✈ When software configuration on some nodes is different than on others
  ✈ Small differences (minor version numbers on libraries) can cripple a parallel program

◆ The second most important problem is adequate job control of the parallel process
  ✈ Signal propagation
  ✈ Cleanup
Rocks (open source clustering distribution)
www.rocksclusters.org

- Technology transfer of commodity clustering to application scientists
  - "make clusters easy"
  - Scientists can build their own supercomputers and migrate up to national centers as needed
- Rocks is a cluster on a CD
  - Red Enterprise Hat Linux (opensource and free)
  - Clustering software (PBS, SGE, Ganglia, NMI)
  - Highly programmatic software configuration management
- Core software technology for several campus projects
  - BIRN
  - Center for Theoretical Biological Physics
  - EOL
  - GEON
  - NBCR
  - OptIPuter
- First Software release Nov, 2000
- Supports x86, Opteron/EM64T, and Itanium
- RedHat/CentOS 4.x
Minimum Components

- Local Hard Drive
- Power
- Ethernet

i386 (Pentium/Athlon)
x86_64 (Opteron/EM64T)
ia64 (Itanium) server
Optional Components

- High-performance network
  - Myrinet
  - Infiniband (Infinicon or Voltaire)

- Network-addressable power distribution unit

- Keyboard/video/mouse network not required
  - Non-commodity
  - How do you manage your management network?
Philosophy

- Caring and feeding for a system is not fun
- System Administrators cost more than clusters
  - 1 TFLOP cluster is less than $200,000 (US)
  - Close to actual cost of a fulltime administrator
- The system administrator is the weakest link in the cluster
  - Bad ones like to tinker
  - Good ones still make mistakes
Philosophy continued

- All nodes are 100% automatically configured
  - Zero "hand" configuration
  - This includes site-specific configuration
- Run on heterogeneous standard high volume components
  - Use components that offer the best price/performance
  - Software installation and configuration must support different hardware
  - Homogeneous clusters do not exist
  - Disk imaging requires homogeneous cluster
Philosophy continued

- **Optimize for installation**
  - Get the system up quickly
  - In a consistent state
  - Build supercomputers in hours not months

- **Manage through re-installation**
  - Can re-install 128 nodes in under 20 minutes
  - No support for on-the-fly system patching

- **Do not spend time trying to issue system consistency**
  - Just re-install
  - Can be batch driven

- **Uptime in HPC is a myth**
  - Supercomputing sites have monthly downtime
  - HPC is not HA
Rocks Basic Approach

- Install a frontend
  1. Insert Rocks Base CD
  2. Insert Roll CDs (optional components)
  3. Answer 7 screens of configuration data
  4. Drink coffee (takes about 30 minutes to install)
- Install compute nodes:
  1. Login to frontend
  2. Execute insert-ethers
  3. Boot compute node with Rocks Base CD (or PXE)
  4. Insert-ethers discovers nodes
  5. Goto step 3
- Add user accounts
- Start computing

Optional Rolls
- Condor
- Grid (based on NMI R4)
- Intel (compilers)
- Java
- SCE (developed in Thailand)
- Sun Grid Engine
- PBS (developed in Norway)
- Area51 (security monitoring tools)
- Many Others...
Red Hat

- **Enterprise Linux 4.0**
  - Recompiled from public SRPMs, including errata updates (source code)
  - No license fee required, redistribution is also fine
  - Recompiled for all CPU types (x86, Opteron, Itanium)
  - *Rocks 5.0 will be based on RHEL 5.0 (Centos, or RHEL)*

- **Standard Red Hat Linux kernel**
  - No Rocks added kernel patches

- **No support for other distributions**
  - Red Hat is the market leader for Linux
    - In the US
    - And becoming so in Europe
  - Trivial to support any Anaconda-based system
  - Others would be harder, and require vendor support (SuSe ~ 12 months work)

- **Excellent support for automated installation**
  - Scriptable installation (Kickstart)
  - Very good hardware detection
Cluster State Management

- **Static Information**
  - Node addresses
  - Node types
  - Site-specific configuration

- **Dynamic Information**
  - CPU utilization
  - Disk utilization
  - Which nodes are online
Common to Any Cluster

- Message Passing / Communication Layer
- Job Scheduling and Launching
- Linux Environment
- Linux Kernel

© 2006 UC Regents
Cluster Software Stack

- Parallel Code / WebFarm / Grid / Computer Lab
- Message Passing / Communication Layer
- Job Scheduling and Launching
- Cluster Software Management
- Cluster State Management / Monitoring
- Linux Environment
- HPC Device Drivers (e.g., Interconnect and Storage)
- Linux Kernel

© 2006 UC Regents
Rocks Cluster Software

Cluster Software Management

Cluster State Management / Monitoring

HPC Device Drivers
(e.g., Interconnect and Storage)

Applications
Middleware
Kernel
Cluster Database
Node Info Stored In A MySQL Database

- If you know SQL, you can execute powerful commands
  - Rocks-supplied command line utilities are tied into the database

- E.g., get the hostname for the bottom 8 nodes of each cabinet:

  # cluster-fork --query="select name from nodes where rank<8" hostname
def get_num_nodes():
    '''return the number of compute nodes (MINUS ONE) in the cluster database. Does MINUS ONE since it is assumed head node is not involved in computations'''
    try:
        from MySQLdb import *
    except:
        print ': (Could not connect to database)'

    link = connect(host='localhost', user='apache',
                   db='cluster', passwd='')

    cursor = link.cursor()

    cursor.execute('select * from nodes')

    return cursor.rowcount - 1
Ganglia (or SCMSWeb / SCE Roll)

- **Scalable cluster monitoring system**
  - Based on ip multi-cast
  - Matt Massie, et al from UCB
  - [http://ganglia.sourceforge.net](http://ganglia.sourceforge.net)

- **Gmon daemon on every node**
  - Multicasts system state
  - Listens to other daemons
  - All data is represented in XML

- **Ganglia command line**
  - Python code to parse XML to English

- **Gmetric**
  - Extends Ganglia
  - Command line to multicast single metrics
Ganglia Screenshot

Host Report for Tue, 18 Mar 2003 01:28:58 +0000

Our Cluster > britannic

britannic Overview

This node is up and running

Time and String Metrics

Name: Value
boottime: Tue. 18 Mar 2003 00:23:20 +0000
gexec: OFF
machine_type: ia64
os_name: Linux
os_release: 2.4.18-0.12smp
sys_clock: Tue. 18 Mar 2003 00:23:34 +0000
uptime: 0 day, 1.5

Constant Metrics

Name: Value
cpu_idle: 97.1 %
cpu_num: 2
cpu_speed: 900 MHz
mem_total: 1011368 KB
mtu: 1500 B
swap_total: 1048544 KB
Software Installation

Collection of all possible software packages (AKA Distribution)

Descriptive information to configure a node

RPMs

Kickstart file

Compute Node

IO Server

Web Server

© 2008 UC Regents
Software Repository

Collection of all possible software packages (AKA Distribution)

Descriptive information to configure a node

RPMs

Kickstart file

Compute Node

IO Server

Web Server

Appliances
Installation Instructions

Collection of all possible software packages (AKA Distribution)

RFMs

Kickstart file

Compute Node

IO Server

Web Server

Descriptive information to configure a node
Building a Rocks Distribution

- Start with Red Hat
- Add updates, Rocks (and optional other) software
- Add Kickstart profiles
- Modify Red Hat installation boot image
- Resulting in a Red Hat compatible Rocks distribution
Rocks Installation
Simulation

Slides courtesy of David Villegas and Dany Guevara
Frontend

For a new installation.

frontend rescue
To boot into rescue mode.

Client
do nothing (default)

Boot Roll
v4.2.1 - Cydonia

www.rocksclusters.org

© 2006 university of california regents

NSF

university of california
Welcome to Rocks

Selected Rolls

No rolls have been selected.

If you have CD/DVD-based rolls (that is, ISO images that have been burned onto CDs or a DVD), then click the CD/DVD-based Roll button. The media tray will eject. Then, place your first roll disk in the tray and click Continue. Repeat this process for each roll disk.

If you are performing a network-based installation (also known as a central installation), then input the name of your roll server into the Hostname of Roll Server field and then click the Download button. This will query the roll server and all the rolls that the roll server has available will be displayed. Click the selected checkbox for each roll you will to install from the roll server.

When you have completed your roll selections, click the Next button to proceed to cluster input screens (e.g., IP address selection, root password setup, etc.).

Select Your Rolls

Local Rolls

CD/DVD-based Roll

Network-based Rolls

Hostname of Roll Server: central.rocksclusters.org

Download

Next
Welcome to Rocks

Selected Rolls

No rolls have been selected.

If you have CD/DVD-based rolls (that is, ISO images that have been burned onto CDs or a DVD), then click the CD/DVD-based Roll button. The media tray will eject. Then, place your first roll disk in the tray and click Continue. Repeat this process for each roll disk.

If you are performing a network-based installation (also known as a central installation), then input the name of your roll server into the Hostname of Roll Server field and then click the Download button. This will query the roll server and all the rolls that the roll server has available will be displayed. Click the selected checkbox for each roll you will to install from the roll server.

When you have completed your roll selections, click the Next button to proceed to cluster input screens (e.g., IP address selection, root password setup, etc.).
Welcome to Rocks

Selected Rolls

No rolls have been selected.

If you have CD/DVD-based rolls (that is, ISO images that have been burned onto CDs or a DVD), then click the CD/DVD-based Roll button. The media tray will eject. Then, place your first roll disk in the tray and click Continue. Repeat this process for each roll disk.

If you are performing a network-based installation (also known as a central installation), then input the name of your roll server into the Hostname of Roll Server field and then click the Download button. This will query the roll server and all the rolls that the roll server has available will be displayed. Click the selected checkbox for each roll you will to install from the roll server.

When you have completed your roll selections, click the Next button to proceed to cluster input screens (e.g., IP address selection, root password setup, etc.).
Welcome to Rocks

Selected Rolls

<table>
<thead>
<tr>
<th>Roll Name</th>
<th>Version</th>
<th>Arch</th>
</tr>
</thead>
<tbody>
<tr>
<td>kernel</td>
<td>4.2</td>
<td>x86_64</td>
</tr>
</tbody>
</table>

Select Your Rolls

Local Rolls

CD/DVD-based Roll

Network-based Rolls

Hostname of Roll Server: central.rocksclusters.org

Download

Next
Welcome to Rocks

Help

Fully-Qualified Host Name:
This must be the fully-qualified domain name (required).

Cluster Name:
The name of the cluster (optional).

Certificate Organization:
The name of your organization. Used when building a certificate for this host (optional).

Certificate Locality:
Your city (optional).

Certificate State:
Your state (optional).

Certificate Country:

Cluster Information

<table>
<thead>
<tr>
<th>Fully-Qualified Host Name</th>
<th>cluster.hpc.org</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster Name</td>
<td>Our Cluster</td>
</tr>
<tr>
<td>Certificate Organization</td>
<td>SDSC</td>
</tr>
<tr>
<td>Certificate Locality</td>
<td>San Diego</td>
</tr>
<tr>
<td>Certificate State</td>
<td>California</td>
</tr>
<tr>
<td>Certificate Country</td>
<td>US</td>
</tr>
<tr>
<td>Contact</td>
<td><a href="mailto:admin@place.org">admin@place.org</a></td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://www.place.org/">http://www.place.org/</a></td>
</tr>
<tr>
<td>Latitude/Longitude</td>
<td>N32.87 W117.22</td>
</tr>
</tbody>
</table>

Back  Next
Welcome to Rocks

Help

IP address:
Enter the IP address for eth0. This is the interface that connects the frontend to the compute nodes.

Netmask:
Enter the netmask for eth0.

Ethernet Configuration for eth0

IP address: 10.1.1.1
Netmask: 255.0.0.0

Back
Next
Welcome to Rocks

Help

IP address:
Enter the IP address for eth1. This is the interface that connects the frontend to the outside network.

Netmask:
Enter the netmask for eth1.

Ethernet Configuration for eth1

IP address: 172.19.119.230
Netmask: 255.255.255.0

Back
Next
Welcome to Rocks

Help

Gateway:
The IP address of your public gateway.

DNS Servers:
Supply a comma separated list of your DNS servers.

Miscellaneous Network Settings

<table>
<thead>
<tr>
<th>Gateway</th>
<th>172.19.119.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNS Servers</td>
<td>132.239.1.52</td>
</tr>
</tbody>
</table>

Back          Next
Welcome to Rocks

Help

Password:
The root password for your cluster.

Root Password

Password: ***************
Confirm: ***************

Back    Next
Welcome to Rocks

Help

Time Zone:
Select a timezone for your cluster.

NTP Server:
Input a Network Time Protocol (NTP) server that will keep the clock on your frontend in sync.

Time Configuration

Time Zone: America/Los_Angeles
NTP Server: pool.ntp.org

Back Next
Welcome to Rocks

Help

Auto Partitioning:
The first disk on this machine will be partitioned in the default manner. See the documentation at www.rocksclusters.org for details on the default partitioning scheme.

Manual Partitioning:
The user will be required to set all partitioning information for this machine. A subsequent installation screen will allow you to enter your partitioning information.

Disk Partitioning

Auto Partitioning
Manual Partitioning
Disk Setup

Choose where you would like Rocks to be installed.

If you do not know how to partition your system or if you need help with using the manual partitioning tools, refer to the product documentation.

If you used automatic partitioning, you can either accept the current partition settings (click Next), or modify the setup using the manual partitioning tool.

If you are manually partitioning your system, you can see your current hard drive(s) and partitions displayed below. Use the partitioning tool to add, edit, or delete the partitions.

<table>
<thead>
<tr>
<th>Device</th>
<th>Mount Point/RAID/Volume</th>
<th>Type</th>
<th>Format</th>
<th>Size (MB)</th>
<th>Start</th>
<th>End</th>
</tr>
</thead>
<tbody>
<tr>
<td>/dev/hda</td>
<td>/</td>
<td>ext3</td>
<td></td>
<td>8001</td>
<td>1</td>
<td>1020</td>
</tr>
<tr>
<td>/dev/hda2</td>
<td>/var</td>
<td>ext3</td>
<td></td>
<td>4001</td>
<td>1021</td>
<td>1530</td>
</tr>
<tr>
<td>/dev/hda3</td>
<td></td>
<td>swap</td>
<td></td>
<td>996</td>
<td>1531</td>
<td>1657</td>
</tr>
<tr>
<td>/dev/hda4</td>
<td>Extended</td>
<td></td>
<td></td>
<td>63319</td>
<td>1658</td>
<td>9729</td>
</tr>
<tr>
<td>/dev/hda5</td>
<td>/export</td>
<td>ext3</td>
<td></td>
<td>63319</td>
<td>1658</td>
<td>9729</td>
</tr>
</tbody>
</table>
Installing Packages

We have gathered all the information needed to install Rocks on the system. It may take a while to install everything, depending on how many packages need to be installed.

Install Roll

Put Roll disk 'kernel - Disk 1' in the drive

Status:
Installation of Compute Nodes

Log into Frontend node as root
At the command line run:
> insert-ethers
Insert Ethernet Addresses -- version 4.2
Opened kickstart access to 10.0.0.0/255.0.0.0 network

Choose Appliance Type

Select An Appliance Type:

Compute
- Ethernet Switches
- NAS Appliance
- Power Units
- Remote Management
- Tile

OK
Insert Ethernet Addresses -- version 4.2
Opened kickstart access to 10.0.0.0/255.0.0.0 network

Press <F10> to quit, press <F11> to force quit
Installation of Compute Nodes

- Turn on the compute node
- Select to PXE boot or insert Rocks CD and boot off of it
Insert Ethernet Addresses -- version 4.2
Opened kickstart access to 10.0.0.0/255.0.0.0 network

Inserted Appliances
Discovered New Appliance

Discovered a new appliance with MAC (00:13:72:ba:c8:df)

Press <F10> to quit, press <F11> to force quit
Insert Ethernet Addresses -- version 4.2
Opened kickstart access to 10.0.0.0/255.0.0.0 network

Inserted Appliances

00:13:72:ba:c8:df  compute-0-0  ()  #

Press <F10> to quit, press <F11> to force quit
Insert Ethernet Addresses -- version 4.2
Opened kickstart access to 10.0.0.0/255.0.0.0 network

Insert Appliances

00:13:72:ba:c8:df compute-0-0 (*) #

Press <F10> to quit, press <F11> to force quit
Cluster Administration

- Command-line tools
- Image generation
- Cluster Troubleshooting
- User Management
Command Line Tools

- Cluster-fork – execute command on nodes (serially)
- Cluster-kill – kill a process on all nodes
- Cluster-probe – get information about cluster status
- Cluster-ps – query nodes for a running process by name
Image Generation

- Basis: Redhat Kickstart file
  - plus XML flexibility
  - and dynamic stuff (i.e. support for “macros”)
- Image Location: /export/home/install
- Customization: rolls and extendcompute.xml
- Command: rocks-dist
Image Generation

Cluster Computing - GCB

Example

- **Goal:** Make a regular node a visualization node
- **Procedure**
  - Figure out what packages to install
  - Determine what configuration files to modify
  - Modify *extend-compute.xml* accordingly
  - *(Re-)*deploy nodes
Figure out Packages

- X-Windows Related
  - X, fonts, display manager
- Display wall
  - XDMX, Chromium, SAGE
Modify Config Files

- X configuration
  - xorg.conf
  - Xinitrc
- Display Manager Configuration

```bash
<file name="/etc/X11/xinit/Xclients">
  xhost +
  xset -dpms
  exec mwm
</file>

<file name="/etc/X11/gdm/gdm.conf">
  awk \
  '/id:J:installdefault:/ { print "id:J:installdefault:"; next; } \n  { print; }' /etc/inittab &gt; /tmp/inittab
  mv /tmp/inittab /etc/inittab
  /sbin/chkconfig --level 5 xfs on

  awk \
  '/BEGIN | FS = "-"; \n  \
  $1 ~ /TimedLoginEnabled/ { printf "%s=true\n", $1; next; } \
  $1 ~ /TimedLogin$/ { printf "%s=\n", $1; next; } \
  $1 ~ /#DisallowTCP$/ { printf "DisallowTCP=false\n"; next; } \
  } /etc/X11/gdm/gdm.conf &gt; /tmp/gdm.conf
  mv /tmp/gdm.conf /etc/X11/gdm/gdm.conf
</file>
```
User Management

- **Rocks Directory: /var/411**
  - Common configuration files:
    - Autofs-related
    - `/etc/group`, `/etc/passwd`, `/etc/shadow`
    - `/etc/services`, `/etc/rpc`
  - All encrypted

- **Helper Command**
  - `rocks-user-sync`
Start Computing

- Rocks is now installed
- Choose an MPI runtime
  - MPICH
  - OpenMPI
  - LAM-MPI
- Start compiling and executing
Pelican HPC

- LiveCD for instant cluster creation
- Advantages
  - Easy to use
  - A lot of built-in software
- Disadvantages
  - Not persistent
  - Difficult to add software
Microsoft Solutions

- Windows Server 2003 Compute Cluster Edition (CCE)
- Microsoft Compute Cluster pack (CCP)
- Microsoft MPI (based on MPICH2)
- Microsoft Scheduler
Microsoft CCE

- Advantages
  - Using Remote Installation Services (RIS), compute nodes can be added by simply turning it on
  - May be better for those familiar with Microsoft Environment

- Disadvantages
  - Expensive
  - Only for 64-bit architectures
  - Proprietary
  - Limited Application base
References

- http://pareto.uab.es/mcreel/PelicanHPC/
- http://pareto.uab.es/mcreel/ParallelKnoppix/
- http://www.clustermatic.org
- portmap man page
- http://www.rocksclusters.org/rocksapalooza