Rocks cluster: a cluster oriented Linux distribution
or how to install a computer cluster in a day

Rocks Version 5.5 (Mamba)
http://www.rocksclusters.org

# build
Install New Frontend

# build rescue
Enter "Rescue" mode

(c) 2002-2012 UC Regents
NSF Awards OCI-721623/OCI-1032778
Physical setup
Installing the Frontend
If you have home made rolls or community rolls now is the time to provide them.
Installing the Frontend

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.).

<table>
<thead>
<tr>
<th>Selected</th>
<th>Roll Name</th>
<th>Version</th>
<th>Arch</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CentOS</td>
<td>4.3</td>
<td>i386</td>
</tr>
<tr>
<td></td>
<td>area51</td>
<td>4.2</td>
<td>i386</td>
</tr>
<tr>
<td></td>
<td>base</td>
<td>4.2</td>
<td>i386</td>
</tr>
<tr>
<td></td>
<td>bio</td>
<td>4.2</td>
<td>i386</td>
</tr>
<tr>
<td></td>
<td>condor</td>
<td>4.2</td>
<td>i386</td>
</tr>
<tr>
<td></td>
<td>ganglia</td>
<td>4.2</td>
<td>i386</td>
</tr>
<tr>
<td></td>
<td>grid</td>
<td>4.2</td>
<td>i386</td>
</tr>
<tr>
<td></td>
<td>hpc</td>
<td>4.2</td>
<td>i386</td>
</tr>
<tr>
<td></td>
<td>java</td>
<td>4.2</td>
<td>i386</td>
</tr>
<tr>
<td></td>
<td>kernel</td>
<td>4.2</td>
<td>i386</td>
</tr>
<tr>
<td></td>
<td>sge</td>
<td>4.2</td>
<td>i386</td>
</tr>
<tr>
<td></td>
<td>updates</td>
<td>4.3</td>
<td>i386</td>
</tr>
<tr>
<td></td>
<td>viz</td>
<td>4.2</td>
<td>i386</td>
</tr>
<tr>
<td></td>
<td>vizagra.rocksclusters.org-restore</td>
<td>2006.08.08</td>
<td>i386</td>
</tr>
<tr>
<td></td>
<td>web-server</td>
<td>4.2</td>
<td>i386</td>
</tr>
</tbody>
</table>

Submit
What are “rolls”?

Rolls are packages of packages designed to integrate themselves in the managing system in the same way as the base software, some of them are provided by the distribution developers. On the other hand, extended documentation on how to create new ones has promoted the appearance of others created by the community.

Examples:
- **HPC**: The primary purpose of the HPC Roll is to provide configured software tools that can be used to run parallel applications on your cluster.
- **SGE**
- **BIO**
- **Area51**
What are “rolls”? 

Rolls are packages of packages designed to integrate themselves in the managing system in the same way as the base software, some of them are provided by the distribution developers. On the other hand, extended documentation on how to create new ones has promoted the appearance of others created by the community.

Examples:
- **HPC:** The primary purpose of the HPC Roll is to provide configured software tools that can be used to run parallel applications on your cluster.
  
  The following software packages are included in the HPC Roll:
  - MPI over ethernet environments (OpenMPI, MPICH, MPICH2)
  - PVM
  - Benchmarks (stream, iperf, IOzone)
- **SGE:**
- **BIO:**
- **Area51**
What are “rolls”?

Rolls are packages of packages designed to integrate themselves in the managing system in the same way as the base software, some of them are provided by the distribution developers. On the other hand, extended documentation on how to create new ones has promoted the appearance of others created by the community.

Examples:
- **HPC**: The primary purpose of the HPC Roll is to provide configured software tools that can be used to run parallel applications on your cluster.
- **SGE**: The SGE Roll installs and configures the SUN Grid Engine scheduler.
- **BIO**
- **Area51**
What are “rolls”?

Rolls are packages of packages designed to integrate themselves in the managing system in the same way as the base software, some of them are provided by the distribution developers. On the other hand, extended documentation on how to create new ones has promoted the appearance of others created by the community.

Examples:

- **HPC**: The primary purpose of the HPC Roll is to provide configured software tools that can be used to run parallel applications on your cluster.
- **SGE**: The SGE Roll installs and configures the SUN Grid Engine scheduler.
  Provides:
  - SGE ready to be used (preconfigured queue, hosts groups, pe, etc.)
  - Integrated with HPC roll (no extra configuration is needed to use OpenMPI, MPIC or OpenMPI)
- **BIO**:

---

**SGE**: The SGE Roll installs and configures the SUN Grid Engine scheduler.

Provides:

- SGE ready to be used (preconfigured queue, hosts groups, pe, etc.)
- Integrated with HPC roll (no extra configuration is needed to use OpenMPI, MPIC or OpenMPI)
What are “rolls”?

Rolls are packages of packages designed to integrate themselves in the managing system in the same way as the base software, some of them are provided by the distribution developers. On the other hand, extended documentation on how to create new ones has promoted the appearance of others created by the community.

Examples:
- HPC: The primary purpose of the HPC Roll is to provide configured software tools that can be used to run parallel applications on your cluster.
- SGE: The SGE Roll installs and configures the SUN Grid Engine scheduler.
- BIO: The Bio-informatics Roll is a collection of some of the most common bio-informatics tools that are being used by the community today.
- Area51
What are “rolls”?

**BIO:** The Bio-informatics Roll is a collection of some of the most common bio-informatics tools that are being used by the community today.
- HMMER - From Janelia Farm research institute
- NCBI BLAST - From National Center for Biotechnology Information
- MpiBLAST - From Los Alamos National Laboratory
- biopython
- ClustalW - From the European BioInformatics Institute
- MrBayes - From School of Computational Science at the Florida State University
- T_Coffee - From Information Genomique et Structurale at Centre National de la Recherche Scientifique
- Emboss - From European Molecular Biology Institute
- Phylip - From the Dept. of Biology at the University of Washington
- fasta - From the University of Virginia
- Glimmer - From Center for Bioinformatics and Computational Biology at the University of Maryland
- TIGR Assembler - From the J. Craig Venter Institute

- All the perl utilities mentioned below are from CPAN
  - perl-bioperl
  - perl-bioperl-ext
  - perl-bioperl-run
  - perl-bioperl-db
What are “rolls”?

Rolls are packages of packages designed to integrate themselves in the managing system in the same way as the base software, some of them are provided by the distribution developers. On the other hand, extended documentation on how to create new ones has promoted the appearance of others created by the community.

Examples:
- **HPC**: The primary purpose of the HPC Roll is to provide configured software tools that can be used to run parallel applications on your cluster.
- **SGE**: The SGE Roll installs and configures the SUN Grid Engine scheduler.
- **BIO**: The Bio-informatics Roll is a collection of some of the most common bio-informatics tools that are being used by the community today.
- **Area51**: The Rocks Area51 Roll contains utilities and services used to analyze the integrity of the files and the kernel on your cluster.
What are “rolls”? 

Rolls are packages of packages designed to integrate themselves in the managing system in the same way as the base software, some of them are provided by the distribution developers. On the other hand, extended documentation on how to create new ones has promoted the appearance of others created by the community.

Examples:
- **HPC**: The primary purpose of the HPC Roll is to provide configured software tools that can be used to run parallel applications on your cluster.
- **SGE**: The SGE Roll installs and configures the SUN Grid Engine scheduler.
- **BIO**: The Bioinformatics Roll is a collection of some of the most common bioinformatics tools that are being used by the community today.
- **Area51**: The Rocks Area51 Roll contains utilities and services used to analyze the integrity of the files and the kernel on your cluster.

The following software packages are included in the Area51 Roll:
- Tripwire
- chkrootkit
Installing the Frontend

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.).
Installing the Frontend
Installing the Frontend

Welcome to Rocks

### Help

**Public Network Interface:**
This is the Ethernet network that physically connects your frontend to the outside world.

**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 Public Network

<table>
<thead>
<tr>
<th>Public Network Interface</th>
<th>eth1</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP address</td>
<td>198.202.88.119</td>
</tr>
<tr>
<td>Netmask</td>
<td>255.255.255.0</td>
</tr>
</tbody>
</table>

[Back] [Next]

Done
Installing the Frontend

Welcome to Rocks

Help

Private Network Interface:
This is the Ethernet network that physically connects your frontend to compute nodes.

IP address:
Enter the IP address for Private (cluster) Network. This is the interface that connects the frontend to the compute nodes.

Netmask:
Enter the netmask for private network.

Ethernet Configuration for Private Network

Private Network Interface
eth0

IP address
10.1.1.1

Netmask
255.255.0.0
Installing the Frontend

Welcome to Rocks

Help

Private Network Interface:
This is the Ethernet network that physically connects your frontend to compute nodes.

Private Network Interface
- eth0

IP address: 10.1.1.1
Netmask: 255.255.0.0

IP address:
Enter the IP address for Private (cluster) Network. This is the interface that connects the frontend to the compute nodes.

Netmask:
Enter the netmask for private network.
Installing the Frontend

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

Gateway: 172.19.119.1
DNS Servers: 132.239.1.52

Back Next
Installing the Frontend

Welcome to Rocks

Help

Password:
The root password for your cluster.

Root Password

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

Back   Next
Installing the Frontend

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**
Installing the Frontend

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

Back  Next
Installing the Frontend

www.rocksclusters.org

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, delete, or resize partitions as needed.
Installing the Frontend

Compute node default partitioning

<table>
<thead>
<tr>
<th>Partition Name</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>/</td>
<td>16 GB</td>
</tr>
<tr>
<td>Swap</td>
<td>1 GB</td>
</tr>
<tr>
<td>/var</td>
<td>4 GB</td>
</tr>
<tr>
<td>/state/partition1</td>
<td>remainder of root disk</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.

DONE!!!
Installing compute nodes

]# insert-ethers
Installing compute nodes
Adding external NFS servers

```bash
# echo "
data0 -fstype=nfs4,rsize=32768,wsize=32768,nodev,nosuid,_netdev,intr,noatime,nostrict 10.3.1.3:/&
data1 -fstype=nfs4,rsize=32768,wsize=32768,nodev,nosuid,_netdev,intr,noatime,nostrict 10.3.1.3:/&
data2 -fstype=nfs4,rsize=32768,wsize=32768,nodev,nosuid,_netdev,intr,noatime,nostrict 10.3.1.3:/&
apps -fstype=nfs4,rsize=32768,wsize=32768,nodev,nosuid,_netdev,intr,noatime,nostrict 10.3.1.3:/&"
>>/etc/auto.share

# rocks sync config
# cd /var/411
# make
```

Adding extra RPMs

```bash
# cp my_new_rpm.el5.x86_64.rpm /export/rocks/install/contrib/5.4/x86_64/RPMS/
# vi /export/rocks/install/site-profiles/5.4/nodes/extend-login.xml
# cd /export/rocks/install/
# rocks sync config
# rocks create distro
```

Now we should reinstall all nodes :( ...but we can do this:

```bash
# rocks run host rpm -Uvh /share/rocks/install/contrib/5.4/x86_64/RPMS/
```
Monitoring the cluster
Monitoring the cluster

Ganlia is installed and configured automatically
http://your_frontend_adress/ganglia

But, most of the times, I prefer to use:

```bash
#] qstat -f -u */|less
```
I have not talk about:
(…and you are probably going to ask)

-Lustre.
There Is a lustre roll and lots of “How to install lustre on rocks”

- Infiniband.
How to configure a fast network for message passing is explained in the basic manual, apart from that, there are infiniband rolls and “Howtos”and I know a company in Barcelona (ANIMA) that installs computer clusters with rocks cluster and infiniband
- By default it is intended that users access the cluster using the frontend. This can be corrected installing a node with the login appliance.

- /home and /share/apps are in the frontend node and are exported to all nodes, even to itself, through NFS, this can cause huge “iowait” problems. It can be corrected “never” using those directories and using external file servers to store your data and applications.

- Compute nodes OS is dependent to frontend OS
This can not be corrected :(