ZEUS Cloud Administrator's Manual

http://zeus.nitc.ac.in

Supervisor: Dr. Vineeth Paleri
Faculty-In-Charge: Dr. Abdul Nazeer

Authors
Sreeraj S, Shamil C. M.

Department of Computer Science and Engineering
National Institute Of Technology Calicut

Contents

Chapter 1

Introduction

ZEUS Cloud is a private cloud that has been set up within the institute to be used by the students for research or project work. ZEUS Cloud is capable of virtualization of machine images and high performance computing. The cloud is scalable, in terms of adding more resources to the existing system and allocation of resources to users according to their needs. Users will be able to try and test platforms and packages required for their projects, and develop, test and debug their programs. The goal of ZEUS Cloud is to deploy computational resources through Local Area Network within the institute for student projects and research work.

The cloud has been hosted in the cluster at BioComputing Research Laboratory.The Cloud is scalable and can be extended to multiple clusters and different subnets. The cloud is also compatible with Amazon EC2. ZEUS Cloud is powered by the cloud computing software Eucalyptus and cluster management suite, Rocks.

For more information visit

- Rocks Clusters `http://rocksclusters.org`
- Eucalyptus Wiki `http://open.eucalyptus.org/wiki`

Chapter 2

Specifications

2.1 Hardware Specifcation

Master Node (Frontend) - Dell PowerEdge 2950

- *Processor:* Two 2.00GHz Quad Core E 5405 Intel XEON processor with 1333MHz FSB (dual socket)
- *Level 2 Cache:* 12MB
- *RAM:* 8 GB using DDR-ECC
- *Hard Disk:* 1 TB SAS Drives
- *Removable Media:* DVD+RW
- *Network Connections:* 3 Gigabit Ethernet Ports with one dedicated management Port
- *Power Supply:* 800W (at 230 VAC)
- *Serial:* One Serial DB-9 Port (rear)
- *USB:* 4 USB 2.0 Ports
- *RAID:* Hardware RAID5 support

Nodes - Dell PowerEdge 1950

- *Number of systems:* 15
- *Processor:* Two 2.00GHz Quad Core E 5405 Intel XEON processor with 1333MHz FSB (dual socket)
- *Level 2 Cache:* 12MB
- *RAM:* 8 GB using DDR-ECC
- *Graphics Controller:* 32MB shared memory
- *Hard Disk:* 300 GB SAS Drives
- *Network Connections:* 2 Gigabit Ethernet Ports
- *Power Supply:* 980W (Non-Hot Plug)
- *USB:* 4 USB 2.0 Ports

2.2 Software Specification

- *Cluster Management:* Rocks Cluster Distribution Suite 5.4.3 (Viper) x86_64
- *Operating System:* Centos 5.7
- *Cloud Sotware:* Eucalyptus 2.0

Chapter 3

Rocks Cluster Installation

3.1 Introduction

Rocks Cluster Management Suite is software which manages the functioning of the cluster. The version installed is 5.4.3 (Viper) 64-bit. Software is installed in the form of Rolls.

3.2 Installation

3.2.1 Frontend Installation

The installation on the Frontend is done using a disk image either by a DVD or a bootable USB drive. The Jumbo DVD has all the required rolls in one single disk image. The x86_64 version of Rocks 5.4.3 can be downloaded from `http://www.rocksclusters.org/wordpress/?page_id=271`.

- Insert the DVD/USB Drive and restart the main node (Frontend). A boot screen will be displayed with a prompt. Enter the following command to start the installation

  ```
  build
  ```

- This commands starts the installation and shows the first screen 'Configure TCP/IP'. Select the option "Enable IPv4 support" and select "Manual Configuration". Disable IPv6 support. Hit "OK" to go to next screen.
- The next screen shows the "Manual IPv4 Configuration". Enter the values for IPv4 address, gateway and name server(DNS server). The IP address is the public address of the cluster.
- The next screen shows the option to select the installation between CD/DVD based rolls (installation packages) or option to download the rolls from the internet. Select the CD/DVD option.
- The next screen shows the list of all rolls in the DVD. Select the required rolls from the list. The kernel, base, os and web-server rolls are mandatory. Additional rolls can be installed by using DVD based rolls. Hit next to proceed.
- This next screen is for entering Cluster Information. Enter the details for Host name, cluster name, organization, locality, state, country, contact, URL, latitude and longitude. The fully-qualified host name is mandatory and is important for several cluster services.
- Next select the option to set the **eth0** IP address and netmask. This is the IP address of the private network between the Frontend and the nodes. The IP address used is 10.1.1.1 and the netmask is 255.0.0.0
- The next screen has the option to set the **eth1** IP address. This is the public IP of the cluster(connected to internet).
- Configure the gateway and DNS entries. Gateway used is 192.168.41.1 and DNS servers used are 192.168.254.2, 192.168.254.3.

- Set the root password in the next screen.
- The next screen shows the option for the partitioning of the hard disk of the Frontend. Select "Manual Partitioning" since the configuration of "Auto partitioning" provides insufficient space for the /var partition which is used by the Eucalyptus Cloud to upload machine images.
  The partition used for the Frontend is:

Partition Name	Size
/	170 GB
/export	170 GB
/var	473 GB
swap	1 GB

  The frontend will format its file systems, then it will ask for each of the roll CDs you added at the beginning of the frontend installation.

3.2.2 Nodes Installation

Make sure all nodes are switched off. Login to the frontend command line node as root.

- The partition of the nodes must be made manual before starting the installation. To do this open the file

  ```
  /export/rocks/install/site-profiles/5.4.3/nodes/replace-partition.xml
  ```

- Replace all content in the <pre> section with

  ```
  echo "rocks manual" > /tmp/user_partition_info
  ```

- Run the following commands to write the changes to the rocks distro which will be used to install the nodes.

  ```
  cd /export/rocks/install
  rocks create distro
  ```

The following steps must be repeated for every node to be installed:

- Type the command

  ```
  insert-ethers
  ```

  It will launch an application which captures the DHCP request for the managed switch, it will configure it as an ethernet switch and store that information in the MySQL database on the frontend.
- Select **VM Container** from the list.
- Start one of the nodes. The node must be booted through PXE (not through hard disk). Select PXEBoot while booting the nodes. When the frontend machine receives the DHCP request from the compute node, the screen shows a discovered appliance.
- Once a (*) mark appears near the discovered device, you can quit insert ethers by pressing F10.
- The node is named as vm-container-X-Y automatically and installation starts on the node. At this point the installation can be monitored from the frontend using the command

  ```
  rocks-console vm-container-X-Y
  ```

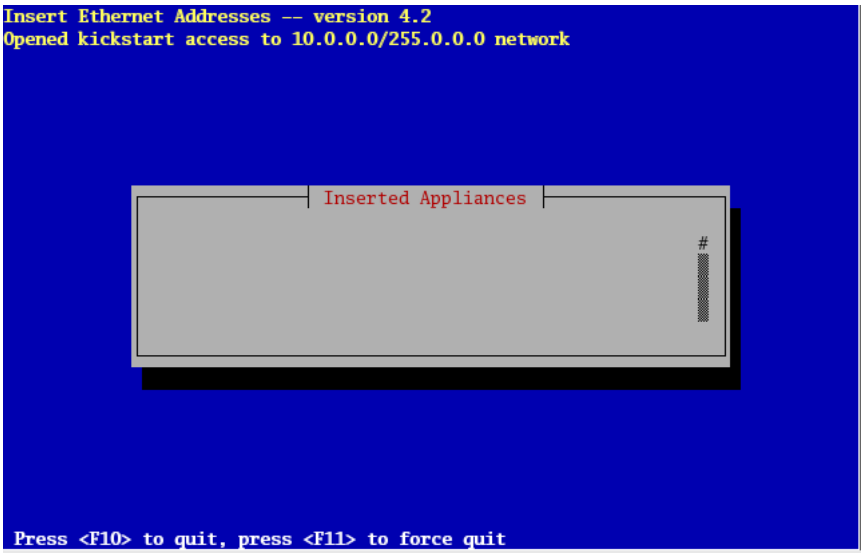

Figure 3.1: Insert-Ethers waiting for node

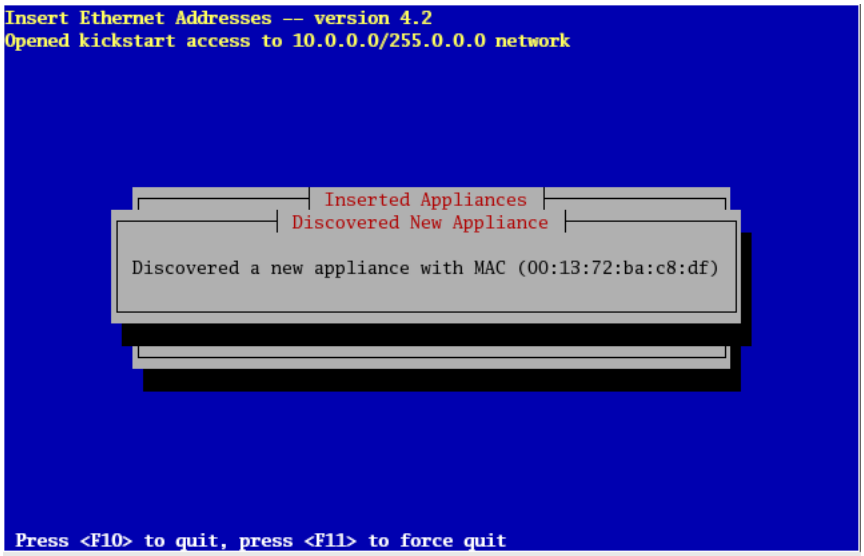

Figure 3.2: Discovering the node

The next screen shows option to partition the nodes. The partition used for nodes is :

Partition Name	**Size**
/	86 GB
/export	100 GB
/var	86 GB
swap	1 GB

Partition the nodes according to the given partition scheme for Eucalyptus to have enough space to upload images, and continue installation.

- Repeat the process for other nodes.

Chapter 4

Eucalyptus Installation

In this chapter, rocks commands to execute certain operations on nodes have been given to run in all the nodes. To run those commands on a single node instead of all the nodes, replace `vm-container` in the commands with `vm-container-0-X` where `X` corresponds to the node number.

4.1 Installing Prerequisites

The yum repositories that are configured now are Rocks specific repositories which doesn't contain few of the required packages. The base repositories of rocks needs to be restored to install these packages using yum.
On the **Frontend**,

```
yum reinstall centos-release

yum install -y java-1.6.0-openjdk ant ant-nodeps dhcp \
            bridge-utils perl-Convert-ASN1.noarch \
            scsi-target-utils httpd
```

4.2 Firewall Configuration

Firewall rules must permit the Eucalyptus components to communicate with one another, and clients to communicate with Eucalyptus.
On the **Frontend**, ports **8443**, **8773**, **8774** and **9001** must be open. Refer to *Rocks User's Guide*

```
rocks add firewall host=frontend network=public protocol=tcp service=8443
chain=INPUT  action=ACCEPT rulename=E10-PORT-8443
rocks add firewall host=frontend network=public protocol=tcp service=8773
chain=INPUT   action=ACCEPT rulename=E10-PORT-8773
rocks add firewall host=frontend network=public protocol=tcp service=8774
chain=INPUT  action=ACCEPT rulename=E10-PORT-8774
rocks add firewall host=frontend network=public protocol=tcp service=9001
chain=INPUT   action=ACCEPT rulename=E10-PORT-9001
```

Edit */etc/sysconfig/iptables* to accept TCP and UDP.

```
#  R900-PRIVILEGED-TCP (global) :
-A INPUT -i eth1 -p tcp --dport 0:1023 -j ACCEPT
#  R900-PRIVILEGED-UDP (global) :
-A INPUT -i eth1 -p udp --dport 0:1023 -j ACCEPT
```

Now synchronize the new rules with the Rocks Database

```
rocks sync host firewall frontend
```

On the **Nodes**, only port **8775** needs to be open. However, enabling firewall causes bridge netfiltering to fail. So, the firewall in nodes needs to be disabled.
Adding firewall rule to open port 8775 in all nodes.

```
rocks add firewall appliance=vm-container protocol=tcp service=8775 network=all \
chain=INPUT action=ACCEPT rulename=E10-PORT-877

rocks sync host firewall vm-container
```

Rename */etc/sysconfig/iptables* to */etc/sysconfig/iptables.backup* . This is to disable firewall and create backup of firewall rules in all nodes.

```
rocks run host vm-container 'mv /etc/sysconfig/iptables /etc/sysconfig/iptables.backup; \
/etc/init.d/iptables restart;'
```

NOTE : To restore iptables in all nodes (if iptables and its backup is corrupted/missing)

```
rocks run host vm-container "cd /etc/sysconfig ; co -f iptables"
```

4.3 Installation

4.3.1 Frontend Installation

In CentOS Eucalyptus can be installed using the package from the Eucalytpus repository. To access Eucalytpus repository create file `/etc/yum.repos.d/euca.repo` with contents :

```
[euca]
name=Eucalyptus
baseurl=http://www.eucalyptussoftware.com/downloads/repo/eucalyptus \
           /2.0.3/yum/centos/x86\_64
gpgcheck=0
```

To install Eucalyptus on Frontend

```
yum install eucalyptus-cloud eucalyptus-cc eucalyptus-walrus eucalyptus-sc
```

4.3.2 Nodes Installation

Eucalytpus is installed on the Nodes as the same way as the Frontend (from repository). The following command is to be run on all nodes to add Eucalyptus repository

```
ssh vm-container-0-X

echo -e "[euca]\nname=Eucalyptus\nbaseurl=http://www.eucalyptussoftware.com/ \
    downloads/repo/eucalyptus/2.0.3/yum/centos/x86\_64\\ngpgcheck=0" >   \
    /etc/yum.repos.d/euca.repo ; exit
```

To check whether repository is added on all nodes, run this on Frontend

```
rocks run host vm-container "cat /etc/yum.repos.d/euca.repo"
```

To install the Eucalyptus on all Nodes, run this on frontend

```
rocks run host vm-container "yum install eucalyptus-nc"
```

To check if installation was successful, run this on Frontend.

```
rocks run host vm-container "yum list installed | grep -o eucalyptus-nc"
```

4.4 Initial Node Configuration

4.4.1 User and Group Configuration

On Nodes there should be a group for libvirt. This group must be associated with the Eucalytpus user that is automatically created during installation. Run the following commands on Frontend.

```
rocks run host vm-container "groupadd libvirt"
rocks run host vm-container "sed s/'libvirt:x:501:'/'libvirt:x:501:eucalyptus'/ \
            /etc/group > /etc/group1 ; mv -f /etc/group1 /etc/group ;"
```

4.4.2 Xen Configuration on Nodes

The Xen configuration file is located in the location /etc/xen/xend-config.sxp
The following commands must be run on Frontend to edit the Xen configuration files for the Nodes.

```
rocks run host vm-container \
   "echo '(xend-http-server yes)' >> /etc/xen/xend-config.sxp"

rocks run host vm-container \
   "echo '(xend-address localhost)' >> /etc/xen/xend-config.sxp"
```

4.4.3 libvirt Configuration on Nodes

libvirt is A toolkit to interact with the virtualization capabilities of recent versions of Linux (and other OSes).
Uncomment the following lines in */etc/libvirt/libvirtd.conf* on all the nodes.

```
unix_sock_group = "libvirt"
unix_sock_ro_perms = "0777"
unix_sock_rw_perms = "0770"
auth_unix_ro = none
auth_unix_rw = none
```

Script to uncomment these lines in all the nodes.

```
rocks run host vm-container "sed s/'#unix_sock_group'/'unix_sock_group'/ \
     /etc/libvirt/libvirtd.conf > /etc/libvirt/libvirtd1.conf ;  \
     mv -f /etc/libvirt/libvirtd1.conf /etc/libvirt/libvirtd.conf;   \
     sed s/'#unix_sock_ro_perms'/'unix_sock_ro_perms'/ /etc/libvirt/libvirtd.conf >  \
     /etc/libvirt/libvirtd1.conf ; \
     mv -f /etc/libvirt/libvirtd1.conf /etc/libvirt/libvirtd.conf;
     sed s/'#unix_sock_rw_perms'/'unix_sock_rw_perms'/ /etc/libvirt/libvirtd.conf >  \
     /etc/libvirt/libvirtd1.conf ; mv -f /etc/libvirt/libvirtd1.conf  \
     /etc/libvirt/libvirtd.conf;"


rocks run host vm-container "sed s/'#auth_unix_ro'/'auth_unix_ro'/  \
    /etc/libvirt/libvirtd.conf > /etc/libvirt/libvirtd1.conf ; mv -f  \
    /etc/libvirt/libvirtd1.conf /etc/libvirt/libvirtd.conf; sed  \
    s/'#auth_unix_rw'/'auth_unix_rw'/ /etc/libvirt/libvirtd.conf >   \
    /etc/libvirt/libvirtd1.conf ; mv -f /etc/libvirt/libvirtd1.conf  \
    /etc/libvirt/libvirtd.conf;"
```

Restart Xen on all Nodes.

```
rocks run host vm-container "/etc/init.d/xend restart"
```

To test Xen functionality on all nodes

```
rocks run host vm-container "su eucalyptus -c 'virsh list'"
```

The output of that command may include error messages (`failed to connect to xend`), but as long as it includes a listing of all domains (at least `Domain-0`), the configuration is in order.
Restart all Eucalyptus services. On the **Frontend**

```
/etc/init.d/eucalyptus-cloud start
/etc/init.d/eucalyptus-cc start
/etc/init.d/eucalyptus-sc start
```

4.5 Registering Cloud

On the **Frontend**,

```
/usr/sbin/euca_conf --register-walrus 192.168.41.203
/usr/sbin/euca_conf --register-cluster ZEUS 192.168.41.203
/usr/sbin/euca_conf --register-sc ZEUS 192.168.41.203

/usr/sbin/euca_conf --register-nodes "10.1.255.254 10.1.255.253 10.1.255.252 10.1.255.251
10.1.255.250 10.1.255.249 10.1.255.248 10.1.255.247 10.1.255.246 10.1.255.245 10.1.255.244
10.1.255.243 10.1.255.242 10.1.255.241 10.1.255.240"
```

Login to the Eucalyptus web interface at `https://192.168.41.203:8443`
Username is *admin* and password is *admin* for first time login.
Make the 'Disk (GB)' 10 for all images, since this is the default value for EMIs.

4.6 Network Configuration

Eucalyptus provides 4 networking configurations, namely STATIC, MANAGED, MANAGED-NOVLAN and SYSTEM. MANAGED-NOVLAN mode has been used for network configuration in ZEUS Cloud. For more information on Networking Modes, refer *Eucalyptus Network Configuration*
If you edit any networking related value in eucalyptus.conf, for the changes to take effect you must perform a "clean restart" of the CC using the following command (make sure to terminate all instances before performing the clean restart):

```
/etc/init.d/eucalyptus-cc cleanrestart
```

4.6.1 Frontend

Edit */etc/eucalyptus/eucalyptus.conf*. Comment This Line.

```
#VNET_MODE="SYSTEM"
```

Uncomment and change the following lines

```
VNET_DHCPDAEMON="/usr/sbin/dhcpd"
VNET_PUBINTERFACE="eth1"
VNET_PRIVINTERFACE="eth0"
VNET_MODE="MANAGED-NOVLAN"
VNET_SUBNET="10.1.0.0"
VNET_NETMASK="255.255.0.0"
VNET_DNS="192.168.254.2"
VNET_ADDRSPERNET="32"
VNET_PUBLICIPS="192.168.41.211 192.168.41.212 192.168.41.213 192.168.41.214
192.168.41.215  192.168.41.216 192.168.41.217 192.168.41.218 192.168.41.219
192.168.41.220  192.168.41.221 192.168.41.222 192.168.41.223 192.168.41.224"
```

4.6.2 Nodes

Edit */etc/eucalyptus/eucalyptus.conf*.

```
#VNET_MODE="SYSTEM"
```

Uncomment and change the following lines

```
VNET_MODE="MANAGED-NOVLAN"

VNET_BRIDGE="xenbr.eth0"

VNET_PUBINTERFACE="xenbr.eth0"
VNET_PRIVINTERFACE="xenbr.eth0"
```

where `xenbr.eth0` is the bridge name which is the output of the follwoing command

```
brctl show
```

For configuring in multiple nodes, download the configuration file from one Node to the Frontend,

```
scp vm-container-0-0:/etc/eucalyptus/eucalyptus.conf eucalyptus.conf
```

Make the necessary changes and then copy it back to all the Nodes.

```
for i in {0..14}
do
scp eucalyptus.conf vm-container-0-$i:/etc/eucalyptus/eucalyptus.conf;
done

/etc/init.d/eucalyptus-nc start ; exit;
```

Chapter 5

Cloud Administration

5.1 Eucalyptus Image Management

5.1.1 Bundling and Uploading Images

A user can bundle and upload Eucalyptus compatible Operating system images which can be used to create instances. The images are uploaded to the walrus storage controller located at the central eucalyptus server.

- Kernel file (vmlinuz* file) :

  ```
  euca-bundle-image -i <kernel-file> --kernel true
  euca-upload-bundle -b <kernel-bucket> -m /tmp/<kernel-file>.manifest.xml
  euca-register <kernel-bucket>/<kernel-file>.manifest.xml
  ```

  Here <kernel-file> is the vmlinuz* file and <kernel-bucket> is the user specified name to denote this file. euca-register will return an ID (<eki-XXXXXXXX>).

- Ramdisk file (initrd* file) :

  ```
  euca-bundle-image -i <ramdisk-file> --ramdisk true
  euca-upload-bundle -b <ramdisk-bucket> -m /tmp/<ramdisk-file>.manifest.xml
  euca-register <ramdisk-bucket>/<ramdisk-file>.manifest.xml
  ```

  Here <ramdisk-file> is the initrd* file and <ramdisk-bucket> is the user specified name. euca-register will return an ID (<eri-XXXXXXXX>).

- Image file (*.img file) :

  ```
  euca-bundle-image -i <image-file> --kernel <eki-XXXXXXXX> \
  --ramdisk <eri-XXXXXXXX>
  euca-upload-bundle -b <image-bucket> -m /tmp/<image-file>.manifest.xml
  euca-register <image-bucket>/<image-file>.manifest.xml
  ```

  Here <image-file> is the *.img file and <image-bucket> is the user specified name. euca-register will return an ID (<emi-XXXXXXXX>).

5.1.2 Deleting Images

To delete an Image, first it has to be deregistered

```
euca-deregister <emi-XXXXXXXX>
```

To remove the image and bucket, use the following code:

```
euca-delete-bundle -b <bucket-name>
euca-delete-bundle -b <bucket-name> --clear
```

5.2 Web Interface

The cloud is accessed using the web interface of eucalyptus. New users sign up for an account which sends a request to the admistrator account.
Login using adminstrator username and password. The administrator can see the list of users and the pending requests. It is up to the administrator to perform request authentication. Once accepted, users can login using their username and passwords.

Adding users

Users can be added by the administrator explicitly by clicking the 'Users' tab, clicking on the 'Add User' button, and filling out the same user form that a user would fill out if they applied themselves. The user will be automatically 'approved' using this method.

Managing users

The administrator can disable or delete a user by clicking the 'Users' tab, and clicking either the 'disable' or 'delete' link respectively.

5.3 Command-Line based Management

5.3.1 Managing Nodes

Once you have a running Eucalyptus system you can add and remove nodes (hosts running Node Controllers) by executing on the Frontend:

```
$EUCALYPTUS/usr/sbin/euca_conf --register-nodes "<nodename1> ... <nodenameN>"
```

Similarly, to remove a node, execute:

```
$EUCALYPTUS/usr/sbin/euca_conf --deregister-nodes "<nodename1> ... <nodenameN>"
```

Chapter 6

Cluster Maintainence

6.1 Some Cluster Commands

For detailed command reference, see
`http://www.rocksclusters.org/roll-documentation/base/5.4.3/c2083.html`. Here are a few essential Rocks commands.

- `rocks list host [hosts]`
  Displays information for all systems in the cluster. `[hosts]` is optional list of host names for whose information is to be displayed. If it is not specified information about all hosts are displayed.
- `rocks list network [network-names]`
  Displays information for all networks. `[network-names]` is optional list of network names for whose information is to be displayed. If it is not specified information about all networks are displayed.
- `rocks list roll [roll-names]`
  Displays statuses of all rolls installed in the cluster. `[roll-names]` is optional list of roll names whose status is to be displayed. If it is not specified information about all rolls are displayed.
- `rocks run host [hosts] command`
  Used to run a shell `command` on multiple systems within the cluster. `[hosts]` specify an optional list of hosts on which the `command` is to be run. If no hosts are specified, the command is run on all the nodes in the cluster.
- `rocks sync config`
  Rebuild all configuration files and restart relevant services.
- `rocks update`
  Download and install updated packages and Rolls from Rocks. This does not include any OS packages. This does not rebuild the distribution or update the backend nodes.

6.2 Shutting Down the Cluster

All the nodes can be shut down from the frontend. Run the following command:

```
rocks run host vm-container "init 0"
```

After all nodes stop working, shut down the frontend.
If required, you may switch off the power switches to the frontend, nodes and the network switch.