



Thomas Cameron, RHCE, RHCX, CNE, MCSE, MCT
Solutions Architect Team Lead
Lon Hohberger
Cluster Software Engineer



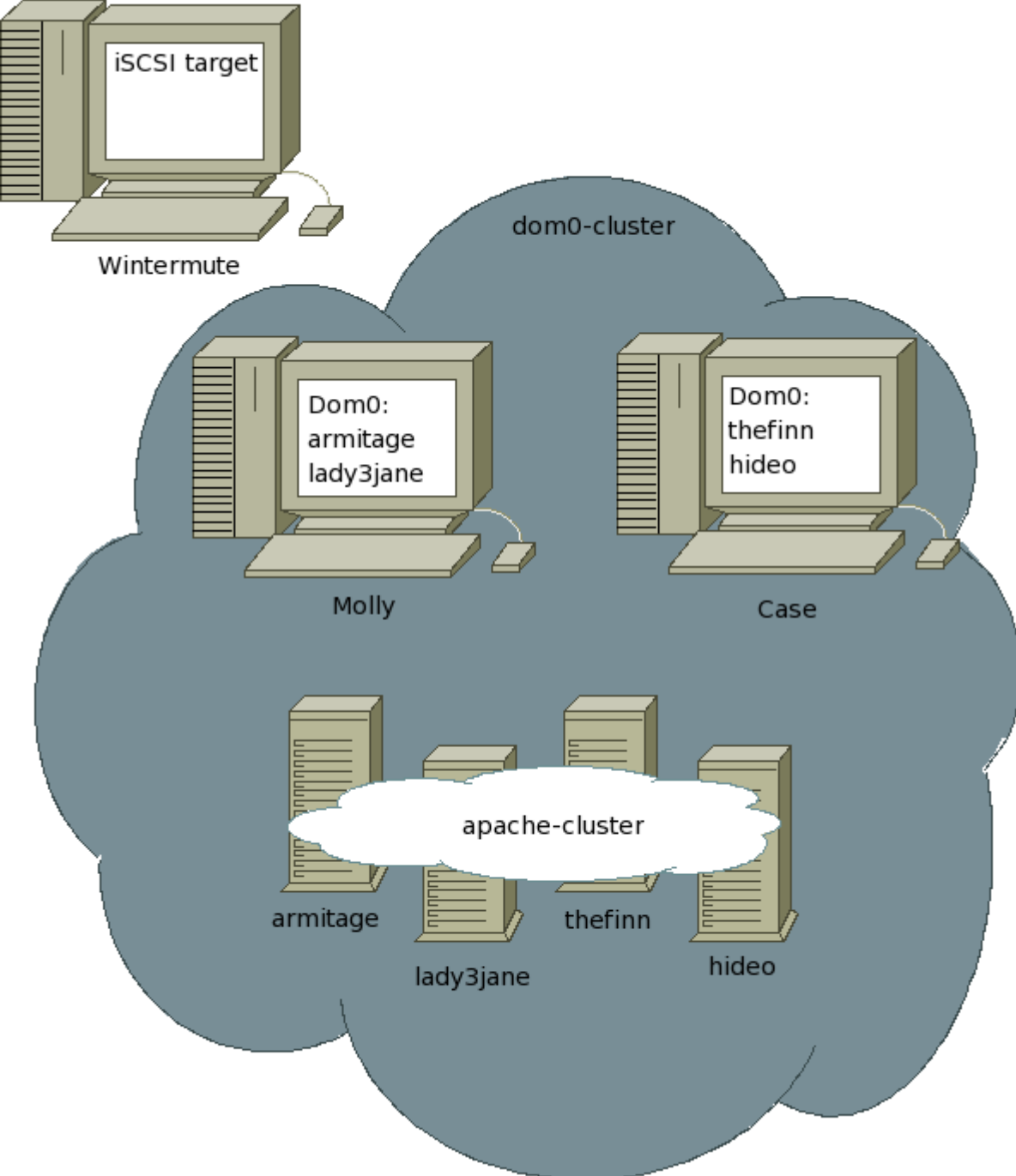
Dedicated to the Memory of
Jim Webb, Global Professional Services Consultant
March 1st, 1963 – May 31st, 2008

Cluster Failover For Any Application

- This session will present the installation and configuration of a high availability cluster of virtual machines using Red Hat Enterprise Linux Advanced Platform (RHEL AP).
- Demonstrations will include:
 - setting up an iSCSI target (server)
 - setting up iSCSI initiators (clients)
 - Red Hat Global Filesystem and Red Hat Cluster Suite functionality included in RHEL AP
 - the Conga web-based cluster management tool

Cluster Failover For Any Application (cont.)

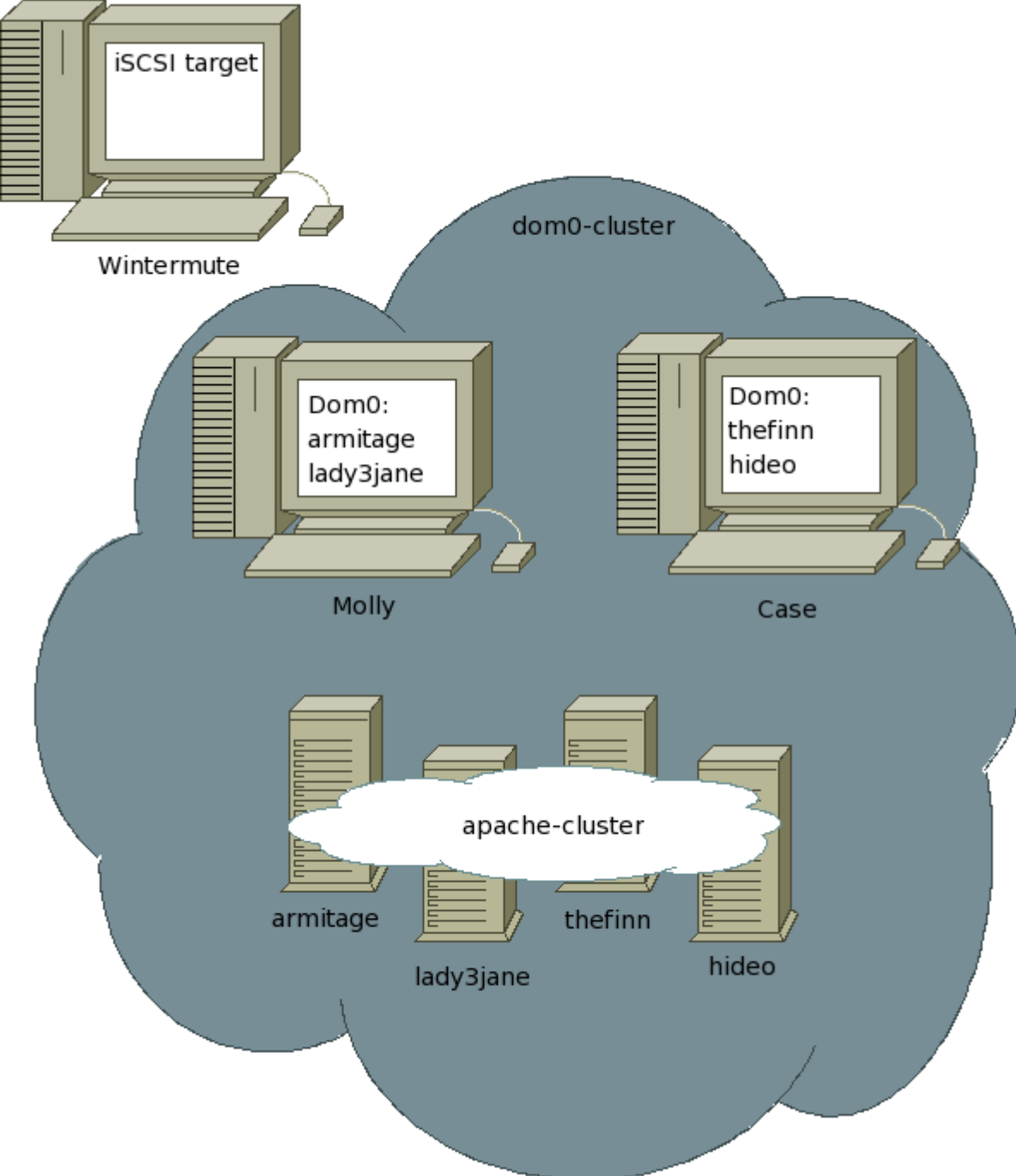
- This presentation will include:
 - setting up clustered logical volumes
 - deploying and clustering virtual machines
 - managing applications on that cluster
- This presentation is available online at <http://people.redhat.com/tcameron>



All hosts are in the fictional “redhat.lan” DNS domain.

wintermute, case and molly are physical servers.

case and molly are iSCSI initiators (clients) and they connect to wintermute, an iSCSI target (server)

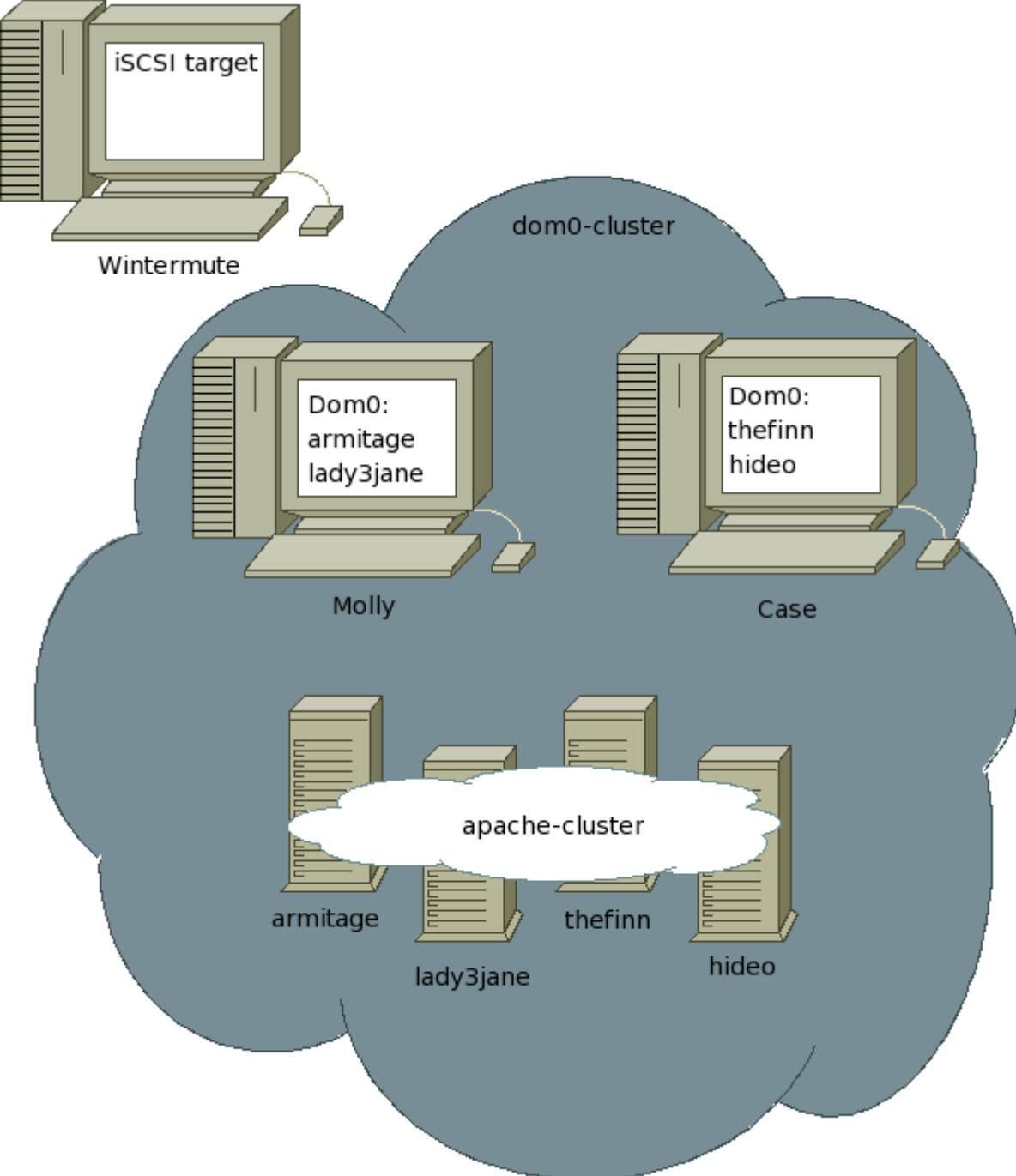


case and molly make up the dom0-cluster.

armitage and lady3jane are domU guests on molly.

thefinn and hideo are domU guests on case.

armitage, lady3jane, thefinn and hideo make up the Apache cluster



The apache-cluster is a cluster of web servers. These servers run on the four domU machines.

Setting up the dom0 cluster

- Ensure that the Clustering and Cluster Storage groups are installed
- `yum grouplist`

```
Installed Groups:  
  Cluster Storage  
  Clustering
```

Setting up the dom0 cluster

- Install luci on a machine which will not be a part of the cluster. For this presentation, wintermute is not actually part of the cluster.
- `yum install luci`
- `luci_admin init` and set administrative password
- `chkconfig luci on`
- `service luci restart`

Setting up the dom0 cluster

```
root@wintermute:~  
File Edit View Terminal Tabs Help  
[root@wintermute ~]# luci_admin init  
Initializing the luci server  
  
Creating the 'admin' user  
  
Enter password:  
Confirm password:  
  
Please wait...  
The admin password has been successfully set.  
Generating SSL certificates...  
The luci server has been successfully initialized  
  
You must restart the luci server for changes to take effect.  
  
Run "service luci restart" to do so  
  
[root@wintermute ~]# service luci restart  
Shutting down luci: [ OK ]  
Starting luci: Generating https SSL certificates... done [ OK ]  
  
Point your web browser to https://wintermute.redhat.lan:8084 to access luci  
  
[root@wintermute ~]#
```

Setting up the dom0 cluster

- Go to the URL mentioned by when restarting the `luci` service
- In this case, <https://wintermute.redhat.lan:8084>
- Note that you'll have to add an exception in Firefox since it is a self-signed certificate

homebase — luci - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/homebase/in

Smart Bookmarks Getting Started Latest Headlines

 **redhat.**  **CLUSTER AND STORAGE SYSTEMS**

homebase cluster storage

help log out

admin

- Add a System
- Add an Existing Cluster
- Add a User

Luci Homebase

Welcome to Luci, admin.

Select an action from the list on the left.

The Conga Cluster and Storage Management System is Copyright © 2000–2008 Red Hat, Inc.
Distributed under the GNU GPL license.

Done

wintermute.redhat.lan:8084

Setting up the dom0 cluster

- Click on the Cluster tab, then on the “Create a new cluster” link
- In this example, the dom0 cluster is called “dom0-cluster.” Not imaginative, but effective.
- The two members are case and molly.
- Choose the “Download packages” radio button – the nodes will download the necessary packages from RHN
- Verify that “Enable Shared Storage Support” is checked
- Verify that “Reboot nodes before joining cluster” is checked to make sure that the systems join cleanly at boot time
- “Check if node passwords are identical” if they are

Luci — cluster — Deploy a cluster - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index

Smart Bookmarks Getting Started Latest Headlines

redhat CLUSTER AND STORAGE SYSTEMS

homebase cluster storage help log out

clusters

- Cluster List
- Create a New Cluster
- Configure

Create a new cluster

Cluster Name dom0-cluster

Node Hostname	Root Password	Key ID
case.redhat.lan		 
molly.redhat.lan		 
		 

Add a cluster node

☒ Download packages
☐ Use locally installed packages.

☒ Enable Shared Storage Support
☒ Reboot nodes before joining cluster
☒ Check if node passwords are identical.

View SSL cert fingerprints

Submit

Done wintermute.redhat.lan:8084

Setting up the dom0 cluster

- Click on Submit to install and reboot

cluster — luci - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index_

Smart Bookmarks Getting Started Latest Headlines

redhat CLUSTER AND STORAGE SYSTEMS

homebase cluster storage help log out

clusters

- Cluster List
- Create a New Cluster
- Configure
- dom0-cluster

dom0-cluster

Please be patient - this cluster's configuration is being modified.



Creating node "case.redhat.lan" for cluster "dom0-cluster"

Currently unable to reach the ricci agent on this node. The install state is not yet complete.

Install Reboot Configure Join



[Stop waiting for this job to complete](#)

Creating node "molly.redhat.lan" for cluster "dom0-cluster"

Currently unable to reach the ricci agent on this node. The install state is not yet complete.

Install Reboot Configure Join



[Stop waiting for this job to complete](#)

Done wintermute.redhat.lan:8084

Setting up the dom0 cluster

- When the systems come back up the cluster and the nodes should all show up in **green**

Luci — cluster — cluster list - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index

Smart Bookmarks Getting Started Latest Headlines

 **redhat**  **CLUSTER AND STORAGE SYSTEMS**

homebase cluster storage help log out

clusters

- Cluster List
- Create a New Cluster
- Configure

Choose a cluster to administer

Cluster Name: dom0-cluster Restart this cluster

- **Status:** Quorate
- **Total Cluster Votes:** 2
- **Minimum Required Quorum:** 1

Nodes

-  case.redhat.lan
-  molly.redhat.lan

Services

- No Services Defined

The Conga Cluster and Storage Management System is Copyright © 2000–2008 Red Hat, Inc.
Distributed under the GNU GPL license.

Done wintermute.redhat.lan:8084

FileEditViewHistoryBookmarksToolsHelp

←→↺ⓧ🏠🔍

https://wintermute.redhat.lan:8084/luci/cluster/index_html?pagetype=7&clus

Google

Most Visited

Smart Bookmarks

Red Hat, Inc.

Red Hat Network

Support

Shop

Products

Training

redhat

CLUSTER AND
STORAGE SYSTEMS

homebase

cluster

storage

help

log out

clusters

Cluster List

Create a New Cluster

Configure

dom0-cluster

dom0-cluster

Nodes

Services

Resources

Failover Domains

Shared Fence

Devices

dom0-cluster

General

Fence

Multicast

Quorum Partition

Fence Daemon Properties

Post Fail Delay0

Post Join Delay3

Run XVM fence daemon☒

XVM fence daemon key distribution

Enter a node hostname from the host cluster

Enter a node hostname from the hosted (virtual) cluster

Retrieve cluster nodes

Apply

Donewintermute.redhat.lan:8084

Setting up the iSCSI Target

- RHEL 5 AP with
 - @ Clustering
 - @ Cluster Storage
 - scsi-target-utils

Setting up the iSCSI Target

- In this case, wintermute will be the target
- First, set up partition(s) to export
- `/sbin/fdisk` to create partitions
 - For this demonstration, three 9GB partitions were created, `/dev/hda5`, `/dev/hda6` and `/dev/hda7`

Setting up the iSCSI Target

```
root@wintermute:~  
File Edit View Terminal Tabs Help  
[root@wintermute ~]# fdisk -l  
  
Disk /dev/hda: 40.0 GB, 40007761920 bytes  
255 heads, 63 sectors/track, 4864 cylinders  
Units = cylinders of 16065 * 512 = 8225280 bytes  
  
   Device Boot      Start         End      Blocks   Id  System  
/dev/hda1  *           1           13       104391   83  Linux  
/dev/hda2             14          1288     10241437+   83  Linux  
/dev/hda3          1289          1415     1020127+   82  Linux swap / Solaris  
/dev/hda4          1416          4864     27704092+    5  Extended  
/dev/hda5          1416          2510      8795556    83  Linux  
/dev/hda6          2511          3605      8795556    83  Linux  
/dev/hda7          3606          4700      8795556    83  Linux  
[root@wintermute ~]#
```

Setting up the iSCSI Target

- Verify that `scsi-target-utils` is installed
- Enable the `tgtd` service with your favorite service management tool

Setting up the iSCSI Target

```
root@wintermute:~  
File Edit View Terminal Tabs Help  
[root@wintermute ~]# yum list scsi-target-utils  
Loading "security" plugin  
Loading "rhnpplugin" plugin  
rhel-i386-server-cluster- 100% |=====| 1.4 kB 00:00  
rhel-i386-server-producti 100% |=====| 1.2 kB 00:00  
rhel-i386-server-fastrack 100% |=====| 1.2 kB 00:00  
rhel-i386-server-vt-5 100% |=====| 1.4 kB 00:00  
rhel-i386-server-5 100% |=====| 1.4 kB 00:00  
rhel-i386-server-cluster- 100% |=====| 1.4 kB 00:00  
rhel-i386-server-suppleme 100% |=====| 1.2 kB 00:00  
rhn-tools-rhel-i386-serve 100% |=====| 1.2 kB 00:00  
Installed Packages  
scsi-target-utils.i386 0.0-0.20070620snap.el5 installed  
[root@wintermute ~]# chkconfig tgtd on  
[root@wintermute ~]# service tgtd restart  
Stopping SCSI target daemon:  
Starting SCSI target daemon: [ OK ]  
[root@wintermute ~]#
```

Setting up the iSCSI Target

- Use `tgtadm` to define the iSCSI target qualified names
- `tgtadm --lld iscsi --op new --mode target \`
 `--tid 1 -T iqn.2008-06.lan.redhat.disk1`
 - `/usr/sbin/tgtadm` is the SCSI target admin tool
 - `--lld iscsi` defines the low level driver as iSCSI
 - `--op new` defines an operation to create a new definition
 - `--mode target` defines that definition as a target
 - `--tid 1` defines the target identifier as “1”
 - `-T iqn.2008-06.lan.redhat.disk1` defines the iSCSI qualified name with the date (year and month), reversed hostname and the disk label

Setting up the iSCSI Target

- Use `tgtadm` to add the partitions created earlier as targets
- `tgtadm --lld iscsi --op new --mode \`
`logicalunit --tid 1 --lun=1 -b /dev/hda5`
- `--mode localunit` defines the new device as a LUN
- `--tid 1` matches to the target identifier previously defined
- `--lun=1` defines this as the first logical unit (LUN)
- `-b /dev/sdb2` associates that LUN to the block device `/dev/hda5`
- Repeat for each of the partitions created earlier

Setting up the iSCSI Target

- Use tgtadm to allow clients to attach to the targets
- `tgtadm --lld iscsi --op bind --mode target --tid 1 -I 172.31.100.3`
- Or
- `tgtadm --lld iscsi --op bind --mode target --tid 1 -I ALL`
 - `--op bind` defines a bind operation
 - `-I [addr]` or `-I ALL` allows a specific host or everyone to connect

Setting up the iSCSI Target

```
root@wintermute:~  
File Edit View Terminal Tabs Help  
[root@wintermute ~]# tgtadm --lld iscsi --op new --mode target --tid 1 -T iqn.2008-06.lan.redhat.disk1  
[root@wintermute ~]# tgtadm --lld iscsi --op new --mode target --tid 2 -T iqn.2008-06.lan.redhat.disk2  
[root@wintermute ~]# tgtadm --lld iscsi --op new --mode target --tid 3 -T iqn.2008-06.lan.redhat.disk3  
[root@wintermute ~]# tgtadm --lld iscsi --op new --mode logicalunit --tid 1 --lun 1 -b /dev/hda5  
[root@wintermute ~]# tgtadm --lld iscsi --op new --mode logicalunit --tid 2 --lun 2 -b /dev/hda6  
[root@wintermute ~]# tgtadm --lld iscsi --op new --mode logicalunit --tid 3 --lun 3 -b /dev/hda7  
[root@wintermute ~]# tgtadm --lld iscsi --op bind --mode target --tid 1 -I ALL  
[root@wintermute ~]# tgtadm --lld iscsi --op bind --mode target --tid 2 -I ALL  
[root@wintermute ~]# tgtadm --lld iscsi --op bind --mode target --tid 3 -I ALL  
[root@wintermute ~]#
```

Setting up the iSCSI Target

- Use tgtadm to verify that you are exporting all targets correctly:
- `tgtadm --lld iscsi --op show --mode target`

Setting up the iSCSI Target

```
root@wintermute:~  
File Edit View Terminal Tabs Help  
[root@wintermute ~]# tgtadm --lld iscsi --op show --mode target  
Target 1: iqn.2008-06.lan.redhat.disk1  
  System information:  
    Driver: iscsi  
    Status: running  
  I_T nexus information:  
  LUN information:  
    LUN: 0  
      Type: controller  
      SCSI ID: deadbeaf1:0  
      SCSI SN: beaf10  
      Size: 0  
      Backing store: No backing store  
    LUN: 1  
      Type: disk  
      SCSI ID: deadbeaf1:1  
      SCSI SN: beaf11  
      Size: 8G  
      Backing store: /dev/hda5  
  Account information:  
  ACL information:  
    ALL  
Target 2: iqn.2008-06.lan.redhat.disk2  
  System information:
```

Setting up the iSCSI Target

- To make it persistent, either create a custom startup script or just add the lines to `/etc/rc.d/rc.local`

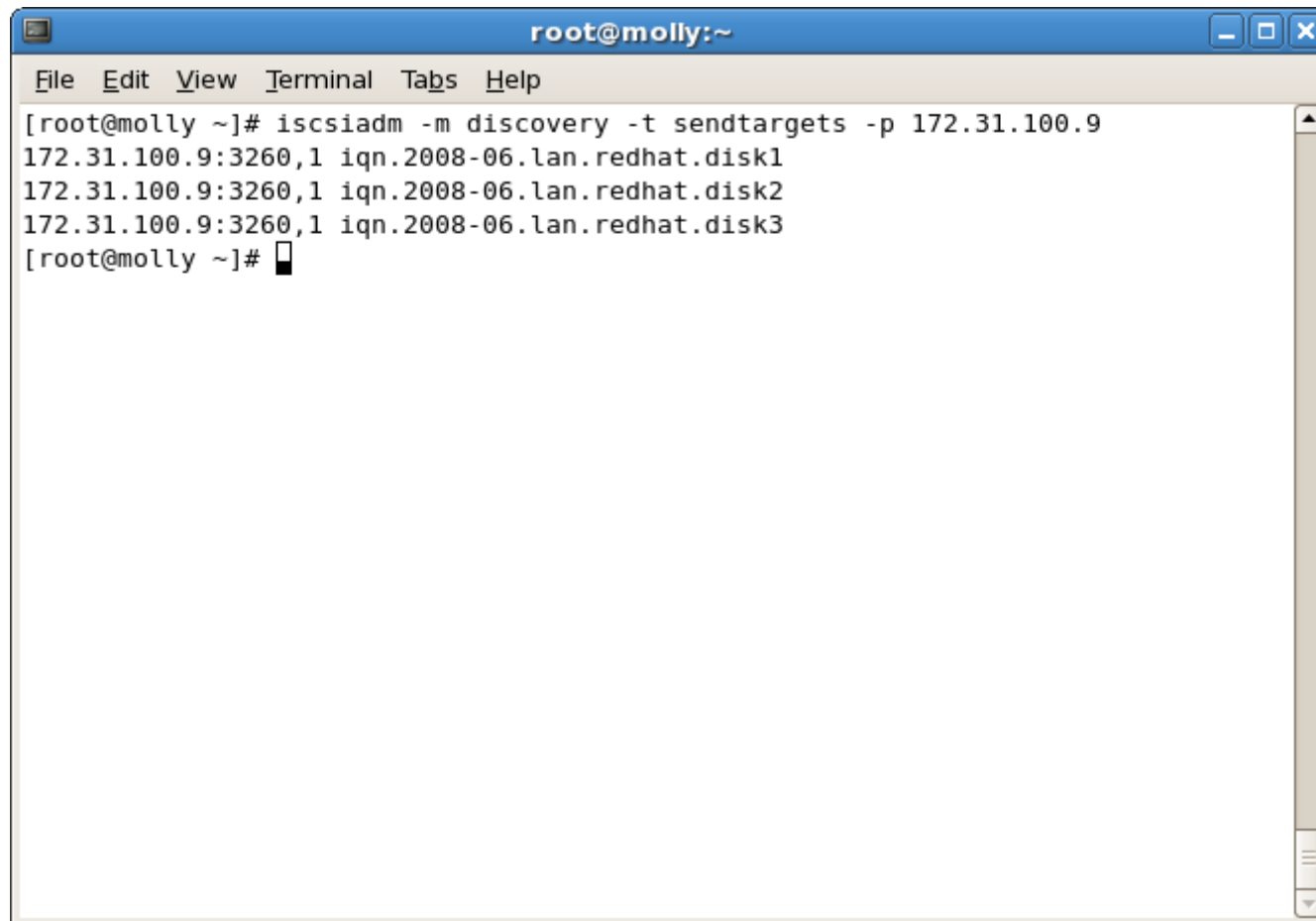
Setting up the iSCSI Target

```
root@wintermute:~  
File Edit View Terminal Tabs Help  
[root@wintermute ~]# cat /etc/rc.d/rc.local  
#!/bin/sh  
#  
# This script will be executed *after* all the other init scripts.  
# You can put your own initialization stuff in here if you don't  
# want to do the full Sys V style init stuff.  
  
tgtadm --lld iscsi --op delete --mode target --tid=1  
tgtadm --lld iscsi --op delete --mode target --tid=2  
tgtadm --lld iscsi --op delete --mode target --tid=3  
tgtadm --lld iscsi --op new --mode target --tid 1 -T iqn.2008-06.lan.redhat.disk1  
tgtadm --lld iscsi --op new --mode target --tid 2 -T iqn.2008-06.lan.redhat.disk2  
tgtadm --lld iscsi --op new --mode target --tid 3 -T iqn.2008-06.lan.redhat.disk3  
tgtadm --lld iscsi --op new --mode logicalunit --tid 1 --lun 1 -b /dev/hda5  
tgtadm --lld iscsi --op new --mode logicalunit --tid 2 --lun 2 -b /dev/hda6  
tgtadm --lld iscsi --op new --mode logicalunit --tid 3 --lun 3 -b /dev/hda7  
tgtadm --lld iscsi --op bind --mode target --tid 1 -I ALL  
tgtadm --lld iscsi --op bind --mode target --tid 2 -I ALL  
tgtadm --lld iscsi --op bind --mode target --tid 3 -I ALL  
tgtadm --lld iscsi --op show --mode target | grep Target  
  
touch /var/lock/subsys/local  
[root@wintermute ~]#
```

Setting up the iSCSI Initiator

- Use the `iscsadm` utility to determine what the target is exporting:
- `iscsiadm -m discovery -t sendtargets -p [host]:3260`
 - `-m discovery` tells `iscsiadm` to discover new information
 - `-t sendtargets` says that information should be target definitions
 - `-p [host]:3260` is the host and optionally the port number to discover

Setting up the iSCSI Initiator

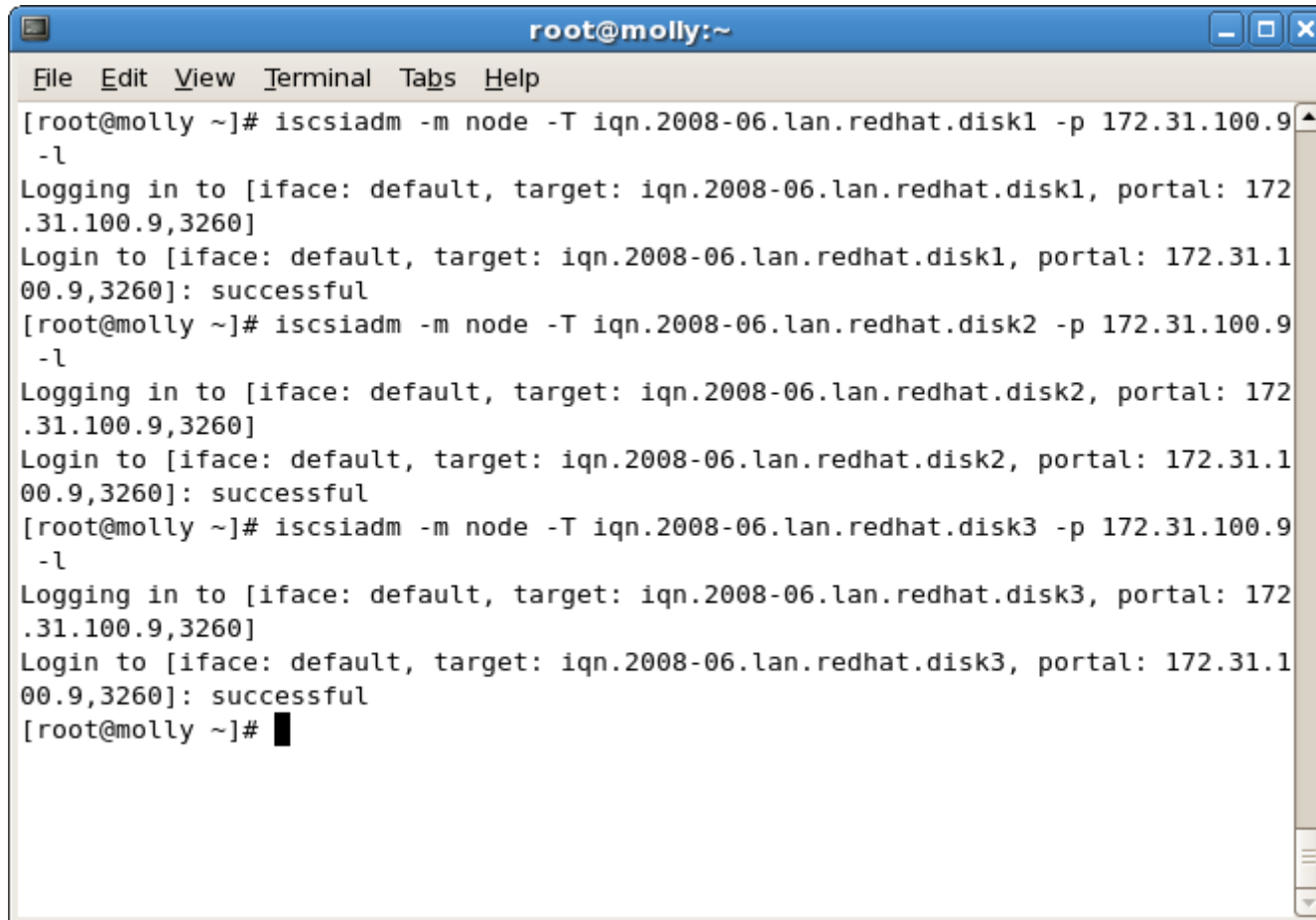


```
root@molly:~  
File Edit View Terminal Tabs Help  
[root@molly ~]# iscsiadm -m discovery -t sendtargets -p 172.31.100.9  
172.31.100.9:3260,1 iqn.2008-06.lan.redhat.disk1  
172.31.100.9:3260,1 iqn.2008-06.lan.redhat.disk2  
172.31.100.9:3260,1 iqn.2008-06.lan.redhat.disk3  
[root@molly ~]#
```

Setting up the iSCSI Initiator

- Use the `iscsadm` utility to log into the target
- `iscsiadm -m node -T \`
`iqn.2008-06.lan.redhat.disk1 \`
`-p 172.31.100.9 -l`
 - `-l` logs the initiator into the target

Setting up the iSCSI Initiator



```
root@molly:~  
File Edit View Terminal Tabs Help  
[root@molly ~]# iscsiadm -m node -T iqn.2008-06.lan.redhat.disk1 -p 172.31.100.9  
-l  
Logging in to [iface: default, target: iqn.2008-06.lan.redhat.disk1, portal: 172  
.31.100.9,3260]  
Login to [iface: default, target: iqn.2008-06.lan.redhat.disk1, portal: 172.31.1  
00.9,3260]: successful  
[root@molly ~]# iscsiadm -m node -T iqn.2008-06.lan.redhat.disk2 -p 172.31.100.9  
-l  
Logging in to [iface: default, target: iqn.2008-06.lan.redhat.disk2, portal: 172  
.31.100.9,3260]  
Login to [iface: default, target: iqn.2008-06.lan.redhat.disk2, portal: 172.31.1  
00.9,3260]: successful  
[root@molly ~]# iscsiadm -m node -T iqn.2008-06.lan.redhat.disk3 -p 172.31.100.9  
-l  
Logging in to [iface: default, target: iqn.2008-06.lan.redhat.disk3, portal: 172  
.31.100.9,3260]  
Login to [iface: default, target: iqn.2008-06.lan.redhat.disk3, portal: 172.31.1  
00.9,3260]: successful  
[root@molly ~]#
```

Setting up the iSCSI Initiator

- Verify that the LUNs are available to the initiator
- `fdisk -l`

Setting up the iSCSI Initiator

```
root@molly:~  
File Edit View Terminal Tabs Help  
[root@molly ~]# fdisk -l  
  
Disk /dev/sda: 100.0 GB, 100030242816 bytes  
255 heads, 63 sectors/track, 12161 cylinders  
Units = cylinders of 16065 * 512 = 8225280 bytes  


| Device    | Boot | Start | End  | Blocks   | Id | System               |
|-----------|------|-------|------|----------|----|----------------------|
| /dev/sda1 | *    | 1     | 13   | 104391   | 83 | Linux                |
| /dev/sda2 |      | 14    | 2563 | 20482875 | 83 | Linux                |
| /dev/sda3 |      | 2564  | 2694 | 1052257+ | 82 | Linux swap / Solaris |

  
Disk /dev/sdb: 9006 MB, 9006649344 bytes  
64 heads, 32 sectors/track, 8589 cylinders  
Units = cylinders of 2048 * 512 = 1048576 bytes  
  
Disk /dev/sdb doesn't contain a valid partition table  
  
Disk /dev/sdc: 9006 MB, 9006649344 bytes  
64 heads, 32 sectors/track, 8589 cylinders  
Units = cylinders of 2048 * 512 = 1048576 bytes  
  
Disk /dev/sdc doesn't contain a valid partition table  
  
Disk /dev/sdd: 9006 MB, 9006649344 bytes  
64 heads, 32 sectors/track, 8589 cylinders  
Units = cylinders of 2048 * 512 = 1048576 bytes  
  
Disk /dev/sdd doesn't contain a valid partition table  
[root@molly ~]#
```

Setting up clustered logical volumes

- Ensure that `clvmd` is running and will survive a reboot
- Create partitions using `fdisk`

Setting up clustered logical volumes

- Use Conga to create a volume group
- Storage tab, select the host, Volume Groups, New Volume Group

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/storage/index_html?pagetype=101& Google

Most Visited Smart Bookmarks Red Hat, Inc. Red Hat Network Support Shop Products Training

 **redhat**  **CLUSTER AND STORAGE SYSTEMS**

homebase cluster **storage** help log out

storage

System List

molly.redhat.lan

Hard Drives
Partition Tables
Software RAIDs
Volume Groups
New Volume Group

Creating New Volume Group

Volume Group Name

Extent Size

Clustered

Select 1 to 3 Physical Volumes

- ☒ /dev/sdb1 (8.38 GB - Partition)
- ☒ /dev/sdc1 (8.38 GB - Partition)
- ☒ /dev/sdd1 (8.38 GB - Partition)

Reset Create

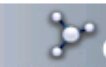
Reprobe Storage

The Conga Cluster and Storage Management System is Copyright © 2000–2008 Red Hat, Inc.
Distributed under the GNU GPL license.

Done wintermute.redhat.lan:8084



redhat

CLUSTER AND
STORAGE SYSTEMS

homebase

cluster

storage

help log out

storage

System List

molly.redhat.lan

Hard Drives

Partition Tables

Volume Groups

molly.redhat.lan

[View recent log activity](#)

Hard Drives

- ▶ [sda](#) 93.16 GB, SCSI ID = SATA_HITACHI_HTS7220071202DP0210DPG6KB1P
- ▶ [sdb](#) 8.38 GB, SCSI ID = 16465616462656166313a3100000000000000000000000000000000
- ▶ [sdc](#) 8.38 GB, SCSI ID = 16465616462656166323a3200000000000000000000000000000000
- ▶ [sdd](#) 8.38 GB, SCSI ID = 16465616462656166333a3300000000000000000000000000000000

Partition Tables

- ▶ [sda](#)
- ▶ [sdb](#)
- ▶ [sdc](#)
- ▶ [sdd](#)

Volume Groups

- ▶ [XenVol](#)

Reprobe Storage

Done

wintermute.redhat.lan:8084



Setting up clustered logical volumes

- Use Conga to create a logical volume
- Storage tab, select the host, Volume Groups, select the VG, then New Logical Volume at the bottom of the screen

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/storage/index_html?storagename=m Google

Most Visited Smart Bookmarks Red Hat, Inc. Red Hat Network Support Shop Products Training

Click cylinders to view properties, unselect all to view Volume Group's properties

Volume Group 'XenVol'

Volume Group Name	XenVol
Extent Size	4.0 MB
Total Extents	6441
Free Extents	6441
Unused	25.16 GB
Size	25.16 GB
Used Extents	0
Maximum Physical Volumes	256
Maximum Logical Volumes	256
Attributes	wz--nc
Clustered	true
UUID	hTrTUQ-Fguf-vZeF-KxiS-eD0H-aqVQ-r71DEU

Done wintermute.redhat.lan:8084

Setting up clustered logical volumes

- Define the content
- Give the cluster a name
- Verify the journal size
- Define the mount point
- Define whether to mount at boot time
- Define whether to modify fstab
- Set the number of journals. Note that with GFS2, the journals are files so adding more later is not painful
- Define whether this is a clustered filesystem

Setting up clustered logical volumes

- PLEASE NOTE THAT GFS2 IS A TECH PREVIEW

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/storage/?pagetype=62&mapper_id=1

Most Visited Smart Bookmarks Red Hat, Inc. Red Hat Network Support Shop Products Training

Click cylinders to view properties, unselect all to view Volume Group's properties

Unused Space - Creating New Logical Volume

Logical Volume Name	XenLV	Content	GFS2 - Global FS v.2
Volume Group Name	XenVol	Cluster Name	dom0-cluster
Size	25.16 (0.00 - 25.16) GB	Unique GFS Name	gfs
Clustered	true	Journal Size	33554432
		Mountpoint	/var/lib/xen/images
		Mount	true
		List in /etc/fstab	true
		Mountable	true
		Locking Protocol	dlm
		Number of Journals	2 (1 - 128)
		Clustered	true

Reset Create

Reprobe Storage

Done

wintermute.redhat.lan:8084

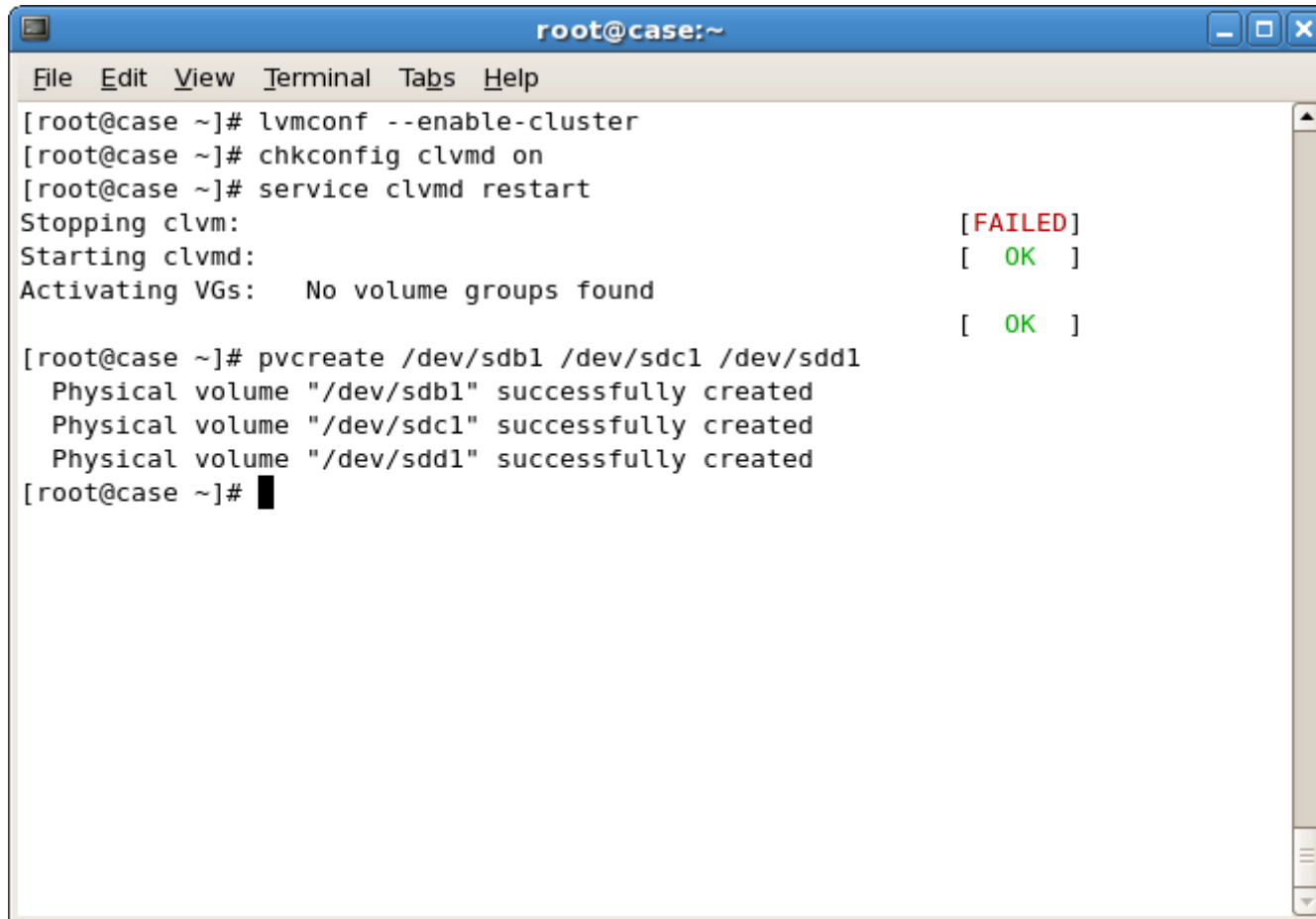
Setting up clustered logical volumes

- Note that you'll need to manually edit the `/etc/fstab` on the other node(s), restart `clvmd` and mount the filesystem there

Setting up clustered logical volumes manually

- Turn on clustered logical volume management with the `/usr/sbin/lvmconf` utility:
- `lvmconf --enable-cluster`
 - Changes `/etc/lvm/lvm.conf` so that the `locking_type` entry is changed to “3” (built-in cluster-wide locking)
- Ensure that `clvmd` is running and will survive a reboot
- Create partitions using `fdisk`
- Create physical volumes using `pvccreate` or `Conga`
 - `pvccreate /dev/sdb1 /dev/sdc1 /dev/sdd1`

Setting up clustered logical volumes manually

A terminal window titled 'root@case:~' with a menu bar (File, Edit, View, Terminal, Tabs, Help). The terminal shows the following commands and output:

```
[root@case ~]# lvmconf --enable-cluster
[root@case ~]# chkconfig clvmd on
[root@case ~]# service clvmd restart
Stopping clvm: [ FAILED ]
Starting clvmd: [ OK ]
Activating VGs: No volume groups found [ OK ]

[root@case ~]# pvcreate /dev/sdb1 /dev/sdc1 /dev/sdd1
Physical volume "/dev/sdb1" successfully created
Physical volume "/dev/sdc1" successfully created
Physical volume "/dev/sdd1" successfully created
[root@case ~]#
```

Setting up clustered logical volumes manually

- Run `pvdiskdisplay /dev/sdb1 /dev/sdc1 /dev/scd1` on the other initiator(s) in the cluster to verify that they see the physical volumes

Setting up clustered logical volumes manually

```
root@molly:~  
File Edit View Terminal Tabs Help  
[root@molly ~]# pvdisplay /dev/sdb1 /dev/sdc1 /dev/sdd1  
"/dev/sdb1" is a new physical volume of "8.39 GB"  
--- NEW Physical volume ---  
PV Name          /dev/sdb1  
VG Name  
PV Size          8.39 GB  
Allocatable      NO  
PE Size (KByte)  0  
Total PE         0  
Free PE          0  
Allocated PE     0  
PV UUID          R7Kq0S-H5B2-tpZC-Ydr3-NnaX-m0yY-wpZUsT  
  
"/dev/sdc1" is a new physical volume of "8.39 GB"  
--- NEW Physical volume ---  
PV Name          /dev/sdc1  
VG Name  
PV Size          8.39 GB  
Allocatable      NO  
PE Size (KByte)  0  
Total PE         0  
Free PE          0  
Allocated PE     0  
PV UUID          2I87Pd-ueT9-bZWU-x9Kl-545d-pti0-DHC0lK
```

Setting up clustered logical volumes manually

- Create volume groups using `/usr/sbin/vgcreate`
 - `vgcreate XenVol \`
`/dev/sdb1 /dev/sdc1 /dev/sdd1`
- You can determine how many free physical extents available using `/usr/sbin/vgdisplay`

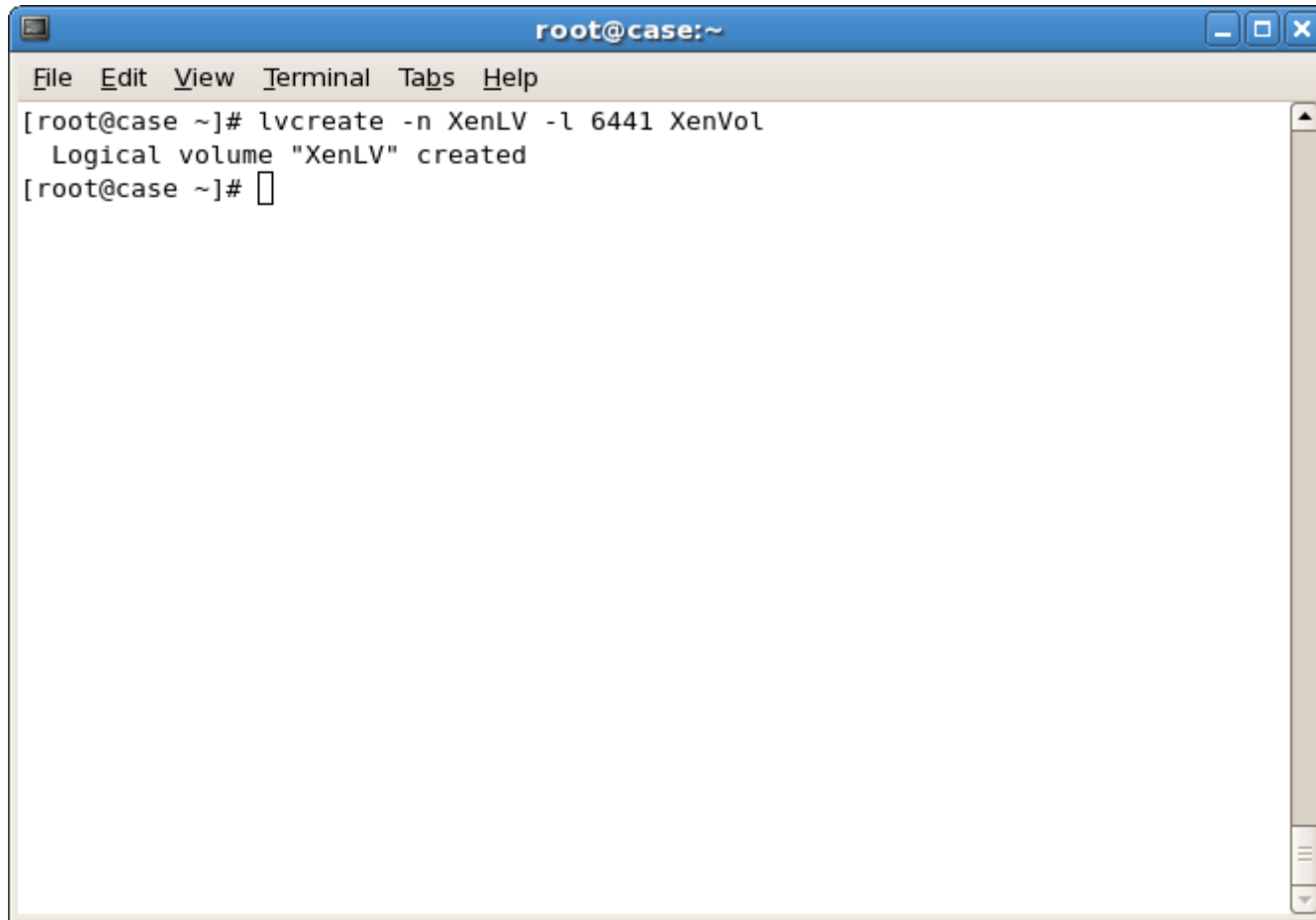
Setting up clustered logical volumes manually

```
root@case:~  
File Edit View Terminal Tabs Help  
[root@case ~]# vgdisplay  
--- Volume group ---  
VG Name                XenVol  
System ID  
Format                 lvm2  
Metadata Areas         3  
Metadata Sequence No   1  
VG Access              read/write  
VG Status              resizable  
Clustered              yes  
Shared                 no  
MAX LV                 0  
Cur LV                0  
Open LV               0  
Max PV                 0  
Cur PV                3  
Act PV                3  
VG Size                25.16 GB  
PE Size                4.00 MB  
Total PE              6441  
Alloc PE / Size        0 / 0  
Free PE / Size         6441 / 25.16 GB  
VG UUID                K7wokx-l0mg-r7MU-quPG-XWLR-dhCp-pa02am
```

Setting up clustered logical volumes manually

- Create logical volume using `/usr/sbin/lvcreate`
 - `lvcreate -n XenLV -l 6441 XenVol`
- Alternatively, you can specify a percentage of free space as in
 - `lvcreate -n XenLV -l 100%FREE XenVol`

Setting up clustered logical volumes manually

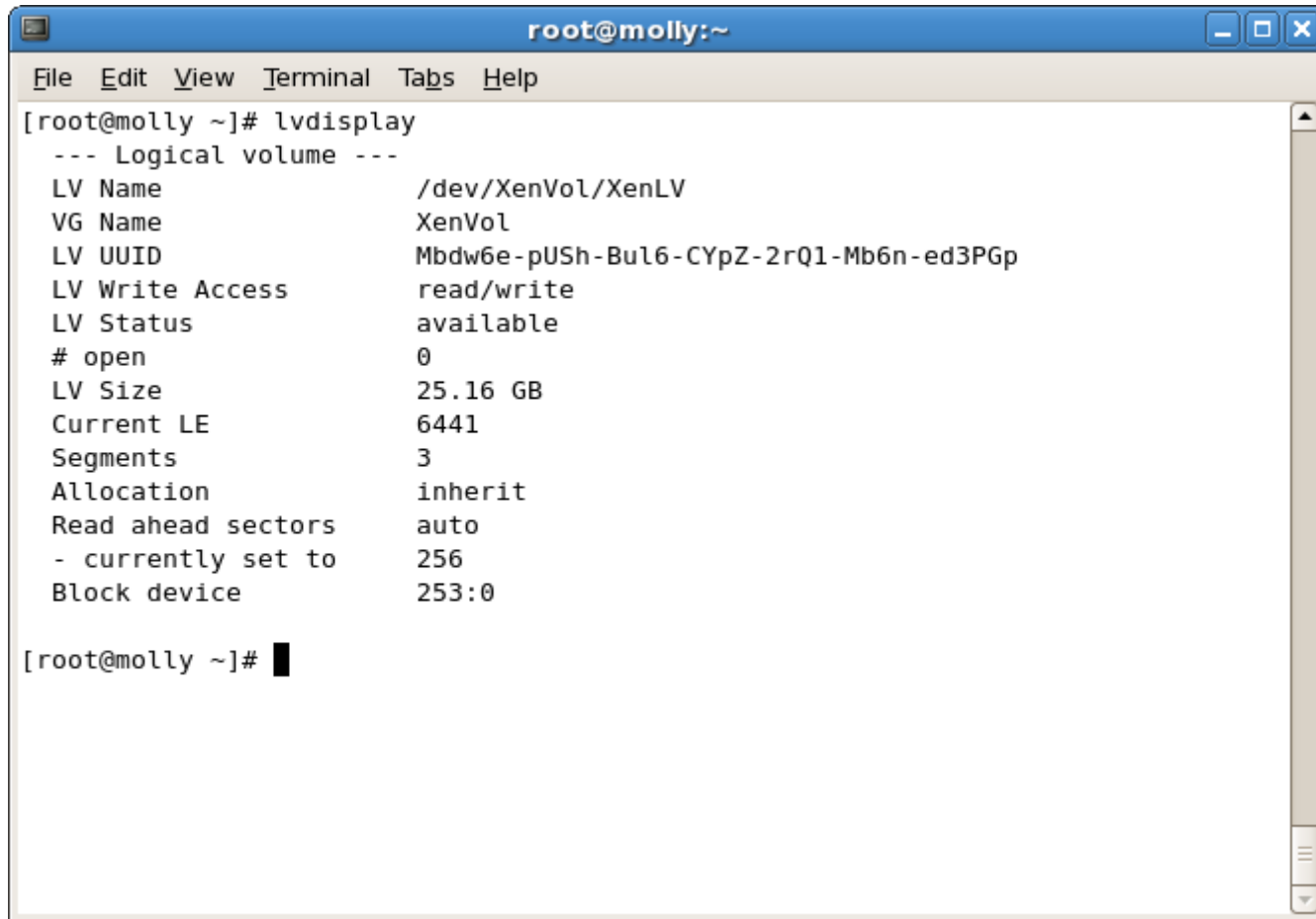
A terminal window titled 'root@case:~' with a menu bar containing 'File', 'Edit', 'View', 'Terminal', 'Tabs', and 'Help'. The terminal shows the command 'lvcreate -n XenLV -l 6441 XenVol' being executed, followed by the output 'Logical volume "XenLV" created' and a new prompt '[root@case ~]#'.

```
root@case:~  
File Edit View Terminal Tabs Help  
[root@case ~]# lvcreate -n XenLV -l 6441 XenVol  
Logical volume "XenLV" created  
[root@case ~]#
```

Setting up clustered logical volumes manually

- Verify that the other node can “see” the logical volume
 - `/usr/sbin/lvdisplay`

Setting up clustered logical volumes manually

A terminal window titled 'root@molly:~' with a menu bar (File, Edit, View, Terminal, Tabs, Help). The terminal shows the command '[root@molly ~]# lvdisplay' and its output. The output is a list of logical volume details for 'XenLV' in volume group 'XenVol'.

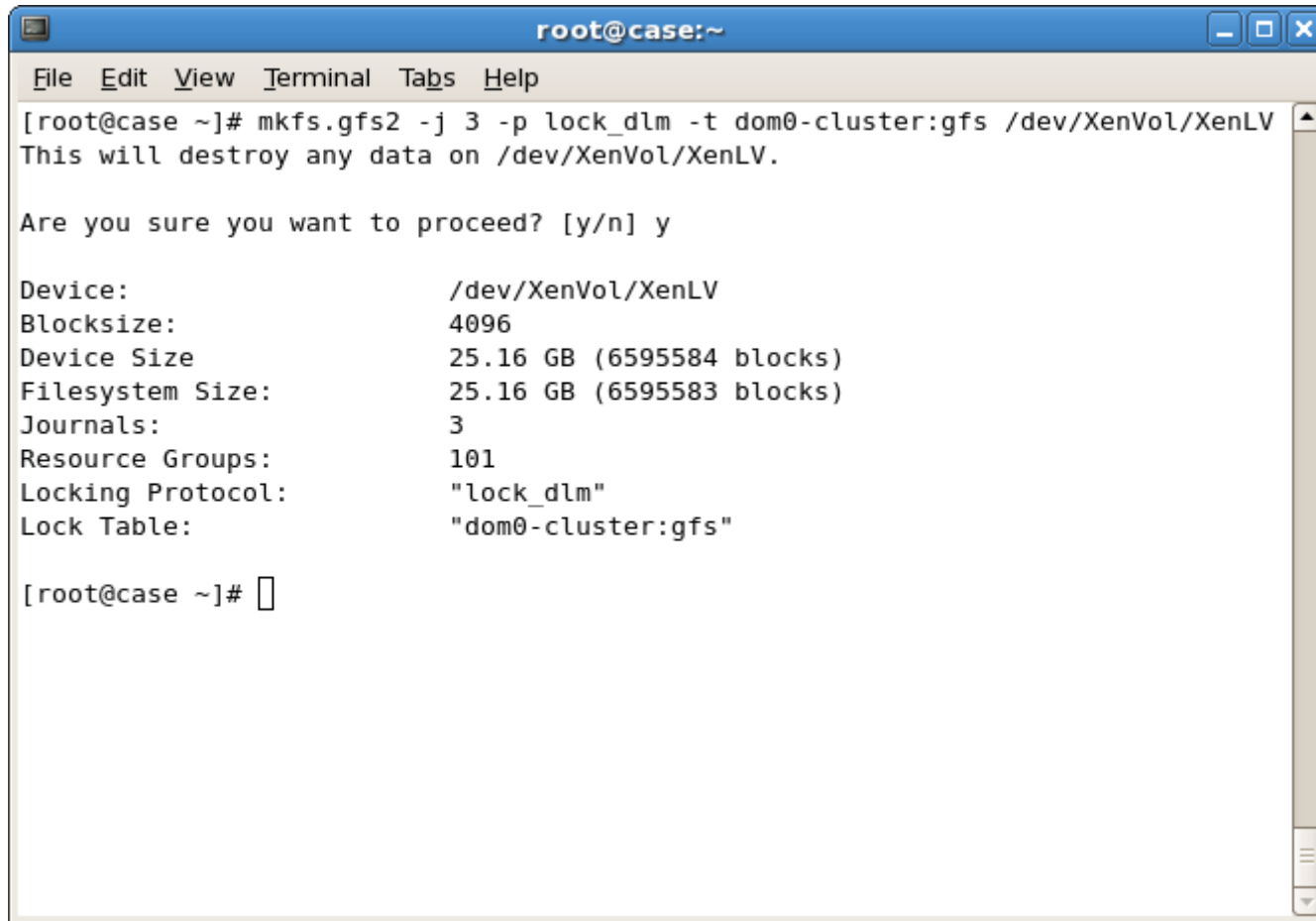
```
[root@molly ~]# lvdisplay
--- Logical volume ---
LV Name                /dev/XenVol/XenLV
VG Name                XenVol
LV UUID                Mbdw6e-pUSh-Bul6-CYpZ-2rQ1-Mb6n-ed3PGp
LV Write Access        read/write
LV Status              available
# open                 0
LV Size                25.16 GB
Current LE             6441
Segments               3
Allocation             inherit
Read ahead sectors     auto
 - currently set to    256
Block device           253:0

[root@molly ~]#
```

Setting up clustered logical volumes manually

- Set up the shared GFS version 2 filesystem using `/sbin/mkfs.gfs2` (TECH PREVIEW)
 - `mkfs.gfs2 -j 3 -p lock_dlm -t \dom0-cluster:gfs /dev/XenVol/XenLV`
 - `mkfs.gfs2` creates the gfs2 filesystem
 - `-j 3` defines the number of journals – in this case, one for each node plus one extra in case the cluster is expanded
 - `-p lock_dlm` defines the locking protocol. In this case, distributed lock manager
 - `-t dom0-cluster:gfs` specifies the lock table name, which is the cluster name and the shared filesystem name
 - `/dev/XenVol/XenLV` is the device to be formatted

Setting up clustered logical volumes manually



```
root@case:~  
File Edit View Terminal Tabs Help  
[root@case ~]# mkfs.gfs2 -j 3 -p lock_dlm -t dom0-cluster:gfs /dev/XenVol/XenLV  
This will destroy any data on /dev/XenVol/XenLV.  
  
Are you sure you want to proceed? [y/n] y  
  
Device:                /dev/XenVol/XenLV  
Blocksize:             4096  
Device Size            25.16 GB (6595584 blocks)  
Filesystem Size:       25.16 GB (6595583 blocks)  
Journals:              3  
Resource Groups:       101  
Locking Protocol:      "lock_dlm"  
Lock Table:            "dom0-cluster:gfs"  
  
[root@case ~]#
```

Setting up clustered logical volumes manually

- Mount the newly created GFS filesystem. In this case on `/var/lib/xen/images` for both dom0 machines
 - `mount /dev/XenVol1/XenLV /var/lib/xen/images`
- You can check `/var/log/messages` to see that the shared storage was mounted correctly

Setting up clustered logical volumes manually

```
root@case:~  
File Edit View Terminal Tabs Help  
[root@case ~]# mount /dev/XenVol/XenLV /var/lib/xen/images/  
[root@case ~]# tail /var/log/messages  
Jun  2 18:29:35 case clvmd: Cluster LVM daemo  
Jun  2 20:01:50 case kernel: dlm: Using TCP  
Jun  2 20:01:50 case kernel: dlm: connecting  
Jun  2 20:01:51 case clvmd: Cluster LVM daemo  
Jun  2 20:18:45 case kernel: GFS 0.1.23-5.el5  
lled  
Jun  2 20:18:45 case kernel: Trying to join c  
s"  
Jun  2 20:18:45 case kernel: Joined cluster.  
Jun  2 20:18:46 case kernel: GFS: fsid=dom0-c  
ire journal lock...  
Jun  2 20:18:46 case kernel: GFS: fsid=dom0-c  
rnal...  
Jun  2 20:18:46 case kernel: GFS: fsid=dom0-c  
[root@case ~]#  
  
root@molly:~  
File Edit View Terminal Tabs Help  
[root@molly ~]# mount /dev/XenVol/XenLV /var/lib/xen/images/  
[root@molly ~]# tail /var/log/messages  
Jun  3 00:19:14 molly kernel: Joined cluster. Now mounting FS...  
Jun  3 00:19:14 molly kernel: GFS: fsid=dom0-cluster:gfs.0: jid=0: Trying to acq  
uire journal lock...  
Jun  3 00:19:14 molly kernel: GFS: fsid=dom0-cluster:gfs.0: jid=0: Looking at jo  
urnal...  
Jun  3 00:19:14 molly kernel: GFS: fsid=dom0-cluster:gfs.0: jid=0: Done  
Jun  3 00:19:14 molly kernel: GFS: fsid=dom0-cluster:gfs.0: jid=1: Trying to acq  
uire journal lock...  
Jun  3 00:19:14 molly kernel: GFS: fsid=dom0-cluster:gfs.0: jid=1: Looking at jo  
urnal...  
Jun  3 00:19:14 molly kernel: GFS: fsid=dom0-cluster:gfs.0: jid=1: Done  
Jun  3 00:19:14 molly kernel: GFS: fsid=dom0-cluster:gfs.0: jid=2: Trying to acq  
uire journal lock...  
Jun  3 00:19:14 molly kernel: GFS: fsid=dom0-cluster:gfs.0: jid=2: Looking at jo  
urnal...  
Jun  3 00:19:14 molly kernel: GFS: fsid=dom0-cluster:gfs.0: jid=2: Done  
[root@molly ~]#
```

Setting up clustered logical volumes manually

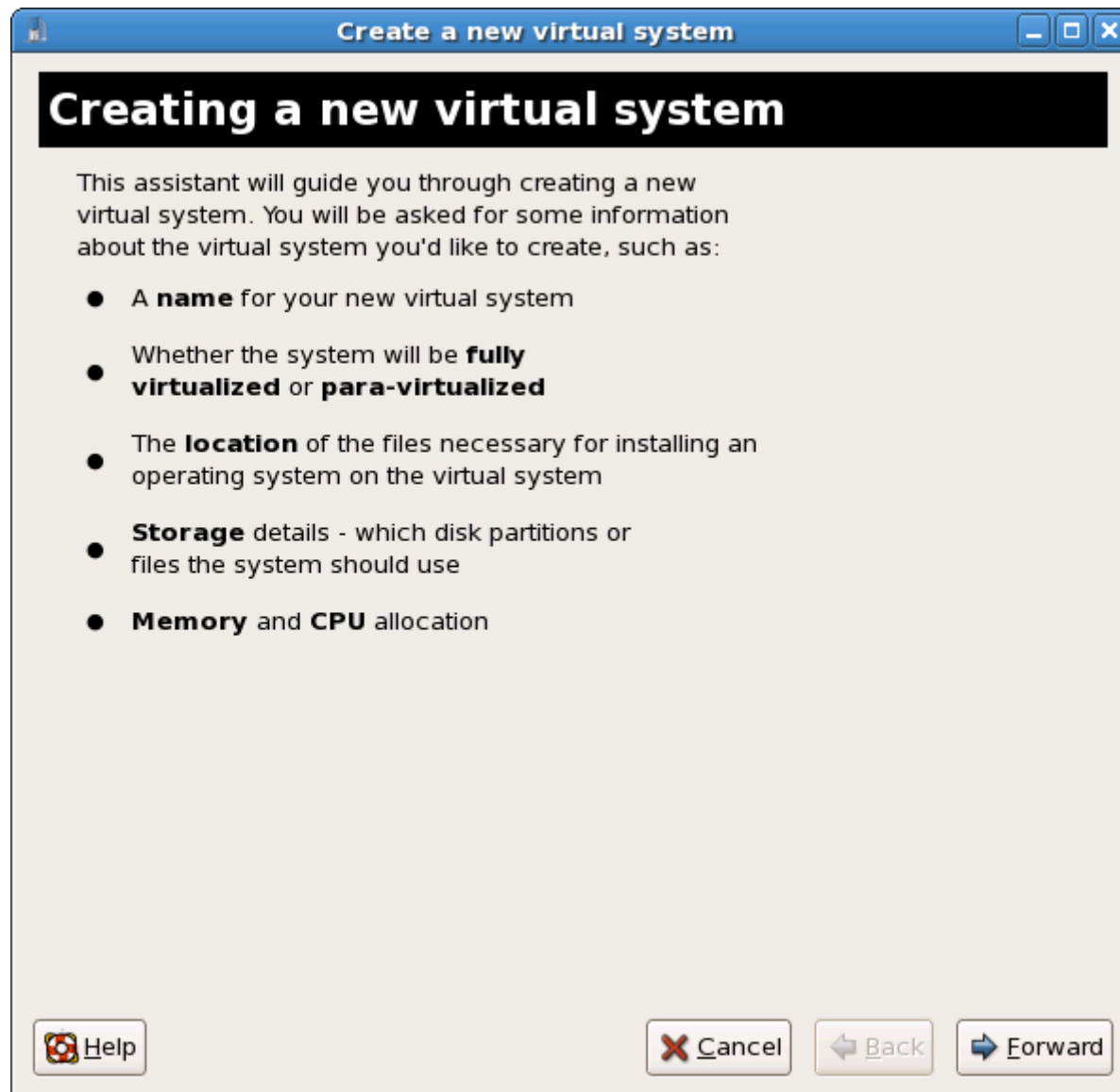
- Add the mount point to /etc/fstab on both hosts

Setting up clustered logical volumes manually

```
root@case:~  
File Edit View Terminal Tabs Help  
[root@case ~]# cat /etc/fstab  
LABEL=/ / ext3 defaults 1 1  
LABEL=/boot1 /boot ext3 defaults 1 2  
tmpfs /dev/shm tmpfs defaults 0 0  
devpts /dev/pts devpts gid=5,mode=620 0 0  
sysfs /sys sysfs defaults 0 0  
proc /proc proc defaults 0 0  
LABEL=SWAP-sda3 swap swap defaults 0 0  
/dev/mapper/XenVol-XenLV /var/lib/xen/images gfs2 rw,noatime 0 0  
[root@case ~]#
```

Installing the domU machines

- Use your favorite installer (`/usr/sbin/virt-manager` if you're a Windows™ admin or `/usr/sbin/virt-install` if you're a real sysadmin)



Create a new virtual system

Naming your virtual system

Please choose a name for your virtual system:

System Name:

 **Example:** system1

 Help

 Cancel  Back  Forward

Create a new virtual system

Choosing a virtualization method

You will need to choose a virtualization method for your new system:

☒ **Paravirtualized:**
Lightweight method of virtualizing machines. Limits operating system choices because the OS must be specially modified to support paravirtualization. Better performance than fully virtualized systems.

☐ **Fully Virtualized:**
Involves hardware simulation, allowing for a greater range of operating systems (does not require OS modification). Slower than paravirtualized systems.

CPU architecture:

☐ Enable kernel / hardware acceleration

 Help

 Cancel  Back  Forward

Create a new virtual system

Locating installation media

Please indicate where installation media is available for the operating system you would like to install on this **paravirtualized** virtual system. Optionally you can provide the URL for a kickstart file that describes your system:

Install Media URL:

i **Example:** http://servername.example.com/distro/i386/tree

Kickstart URL:

i **Example:** ftp://hostname.example.com/ks/ks.cfg

Kernel parameters:

i **Example:** updates=http://hostname.example.com/updates.img

 Help

 Cancel

 Back

 Forward

Create a new virtual system

Assigning storage space

Please indicate how you'd like to assign space on this physical host system for your new virtual system. This space will be used to install the virtual system's operating system.

☐ Normal Disk Partition:

Partition:

 **Example:** /dev/hdc2

☒ Simple File:

File Location:

File Size:

▲

▼

 MB

☒ Allocate entire virtual disk now?

 **Warning:** If you do not allocate the entire disk at VM creation, space will be allocated as needed while the guest is running. If sufficient free space is not available on the host, this may result in data corruption on the guest.

 **Tip:** You may add additional storage, including network-mounted storage, to your virtual system after it has been created using the same tools you would on a physical system.

 Help

Create a new virtual system

Connect to host network

Please indicate how you'd like to connect your new virtual system to the host network.

☐ Virtual network

Network:

Tip: Choose this option if your host is disconnected, connected via wireless, or dynamically configured with NetworkManager.

☒ Shared physical device

Device:

Tip: Choose this option if your host is statically connected to wired ethernet, to gain the ability to migrate the virtual system.

☐ Set fixed MAC address for your virtual system?

MAC address:

Help

Cancel

Back

Forward

Create a new virtual system

Allocate memory and CPU

Memory:

Please enter the memory configuration for this VM. You can specify the maximum amount of memory the VM should be able to use, and optionally a lower amount to grab on startup. Warning: setting VM memory too high will cause out-of-memory errors in your host domain!

Total memory on host machine: 2.99 GB

VM Max Memory (MB): 384

VM Startup Memory (MB): 384

CPU:

Please enter the number of virtual CPUs this VM should start up with.

Logical host CPUs: 2

VCPUs: 1

Tip:

For best performance, the number of virtual CPUs should be less than (or equal to) the number of logical CPUs on the host system.

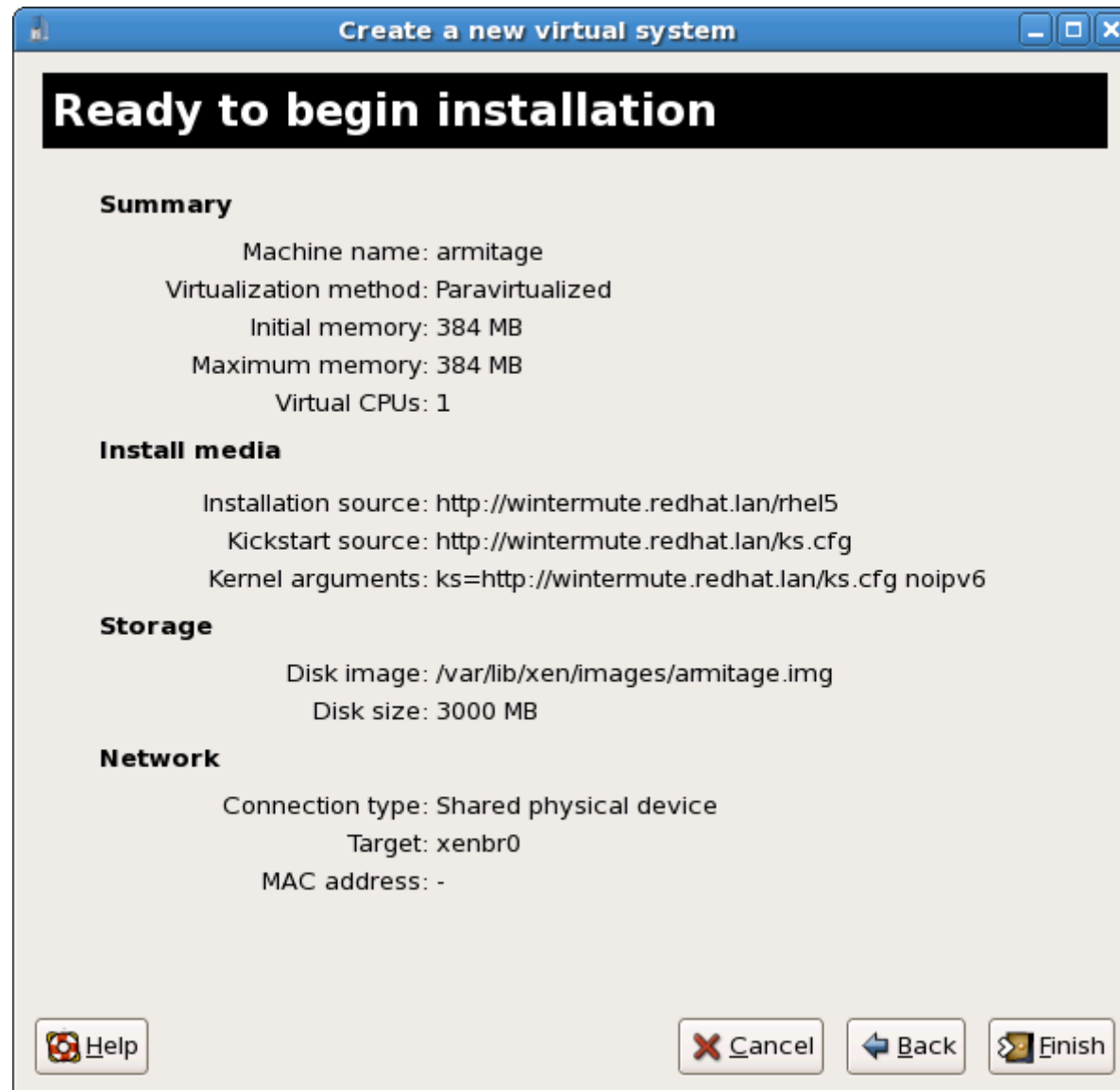
Help

Cancel

Back

Forward

THE FOURTH ANNUAL
RED HAT SUMMIT

A screenshot of a window titled "Create a new virtual system". The window has a blue title bar with standard window controls (minimize, maximize, close). Below the title bar is a black header with the text "Ready to begin installation" in white. The main content area is light gray and contains several sections: "Summary", "Install media", "Storage", and "Network". Each section lists configuration details for a virtual machine named "armitage". At the bottom of the window are four buttons: "Help" (with a question mark icon), "Cancel" (with a red X icon), "Back" (with a left arrow icon), and "Finish" (with a right arrow icon).

Ready to begin installation

Summary

Machine name: armitage
Virtualization method: Paravirtualized
Initial memory: 384 MB
Maximum memory: 384 MB
Virtual CPUs: 1

Install media

Installation source: <http://wintermute.redhat.lan/rhel5>
Kickstart source: <http://wintermute.redhat.lan/ks.cfg>
Kernel arguments: `ks=http://wintermute.redhat.lan/ks.cfg noipv6`

Storage

Disk image: `/var/lib/xen/images/armitage.img`
Disk size: 3000 MB

Network

Connection type: Shared physical device
Target: xenbr0
MAC address: -

 Help  Cancel  Back  Finish



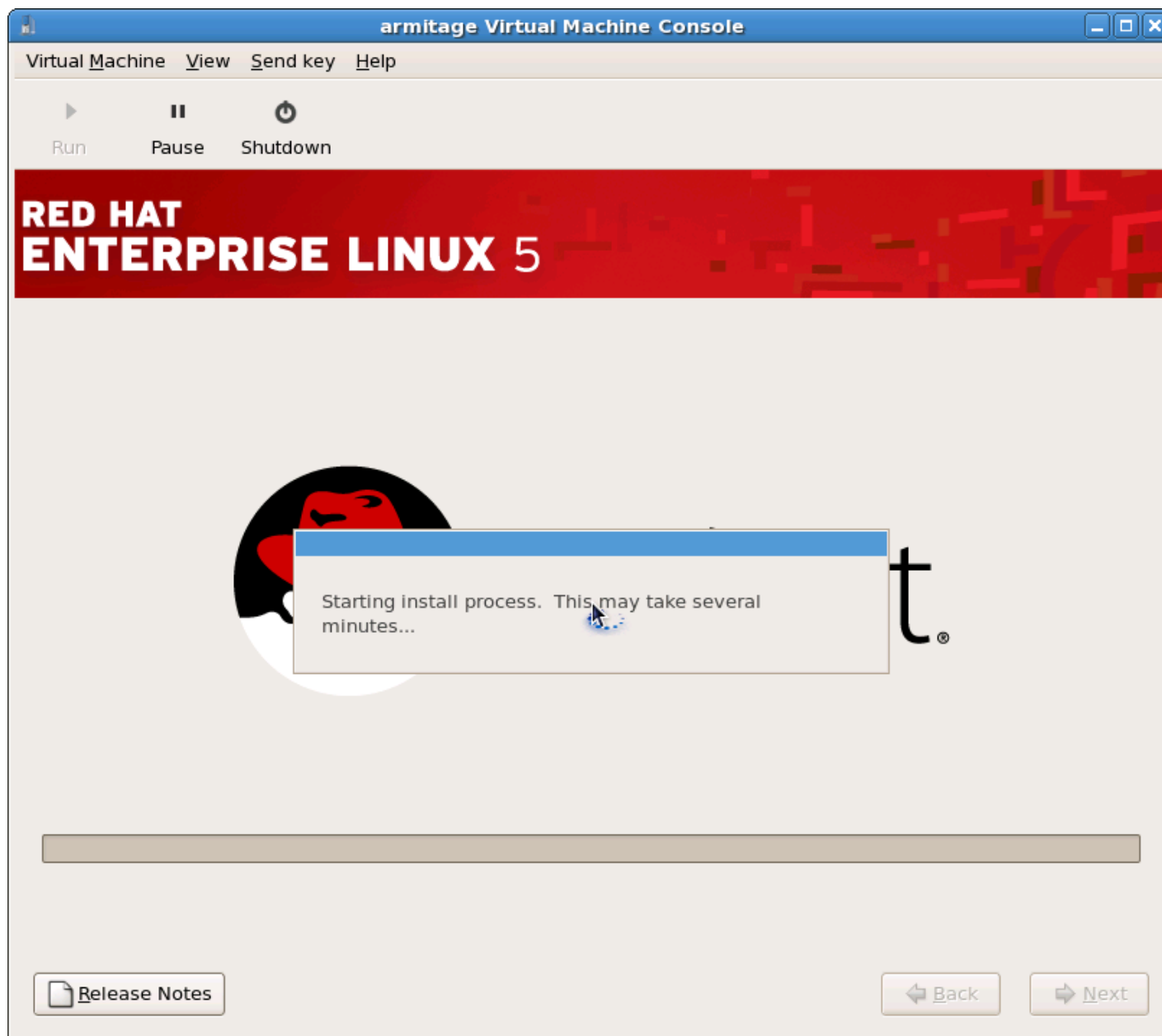
armitage Virtual Machine Console

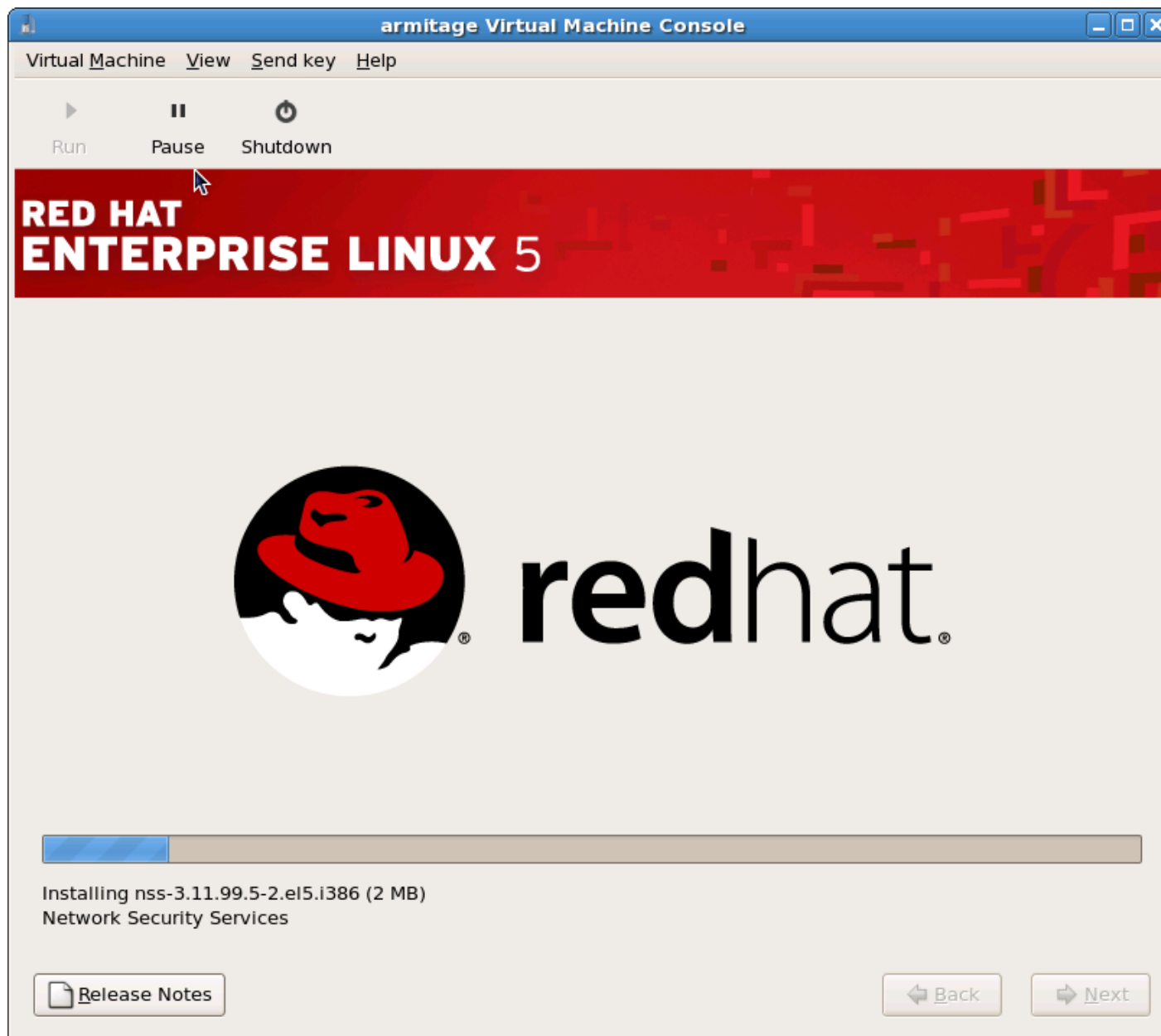
Virtual Machine View Send key Help

Run Pause Shutdown

```
Uniform Multi-Platform E-IDE driver Revision: 7.00alpha2
ide: Assuming 50MHz system bus speed for PIO modes; override with idebus=xx
ide-floppy driver 0.99.newide
usbcore: registered new driver hiddev
usbcore: registered new driver usbhid
drivers/usb/input/hid-core.c: v2.6:USB HID core driver
PNP: No PS/2 controller found. Probing ports directly.
i8042.c: No controller found.
mice: PS/2 mouse device common for all mice
md: md driver 0.90.3 MAX_MD_DEVS=256, MD_SB_DISKS=27
md: bitmap version 4.39
TCP bic registered
Initializing IPsec netlink socket
NET: Registered protocol family 1
NET: Registered protocol family 17
Using IPI No-Shortcut mode
XENBUS: Device with no driver: device/vbd/51712
XENBUS: Device with no driver: device/vif/0
Freeing unused kernel memory: 176k freed
Write protecting the kernel read-only data: 379k

Greetings.
anaconda installer init version 11.1.2.113 starting
mounting /proc filesystem... done
creating /dev filesystem... done
mounting /dev/pts (unix98 pty) filesystem... done
mounting /sys filesystem... done
trying to remount root filesystem read write... done
mounting /tmp as ramfs... done
running install...
running /sbin/loader
```





Installing the domU machines

- `virt-install -n armitage -r 384 \`
 `-f /var/lib/xen/images/armitage.disk \`
 `-s 3 --nonsparse -m 00:16:3e:30:9c:e3 \`
 `--nographics --os-type linux --os-variant \`
 `rhel5 --paravirt --location \`
 `http://wintermute.redhat.lan/rhel5 \`

Installing the domU machines

```
-x "ip 172.31.100.10 netmask 255.255.255.0 \  
gateway 172.31.100.1 dns 172.31.100.9 \  
noipv6 text \  
ks=http://wintermute.redhat.lan/ks.cfg"
```

- Make sure that the Clustering and Cluster Storage groups are part of your installation

Installing the domU machines

```
root@case:~  
File Edit View Terminal Tabs Help  
[root@case ~]# virt-install -n armitage -r 384 -f /var/lib/xen/images/armitage.d  
isk -s 3 --nonsparse -m 00:16:3e:30:9c:e3 --nographics --os-type linux --os-vari  
ant rhel5 --paravirt --location http://wintermute.redhat.lan/rhel5 -x "ip 172.31  
.100.10 netmask 255.255.255.0 gateway 172.31.100.1 dns 172.31.100.9 noipv6 text  
ks=http://wintermute.redhat.lan/ks.cfg"  
  
Starting install...  
Retrieving file .treeinfo 100% |=====| 436 B 00:00  
Retrieving file vmlinuz.. 100% |=====| 2.1 MB 00:00  
Retrieving file initrd.im 100% |=====| 5.4 MB 00:00  
Creating storage file... 14% |===| 454 MB 00:07 ETA
```

Installing the domU machines

- Verify that all hosts are up

```
root@armitage:~
File Edit View Terminal Tabs Help
[tcameron@case ~]$ ssh root@armitage
The authenticity of host 'armitage (172.31.100.10)' can't be established.
RSA key fingerprint is ac:c5:0f:11:01:71:af:0c:1e:f1:f4:d7:e5:58:ff:c7.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'armitage,172.31.100.10' (RSA) to the list of known hosts.
root@armitage's password:
Last login: Thu Jun 5 16:35:26 2008
[root@armitage ~]#
```

```
root@hideo:~
File Edit View Terminal Tabs Help
[tcameron@case ~]$ ssh -Y root@hideo
The authenticity of host 'hideo (172.31.100.12)' can't be established.
RSA key fingerprint is ce:04:27:76:13:54:d3:b3:4c:b5:5f:7e:91:4c:34:fb.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'hideo,172.31.100.12' (RSA) to the list of known hosts.
root@hideo's password:
[root@hideo ~]#
```

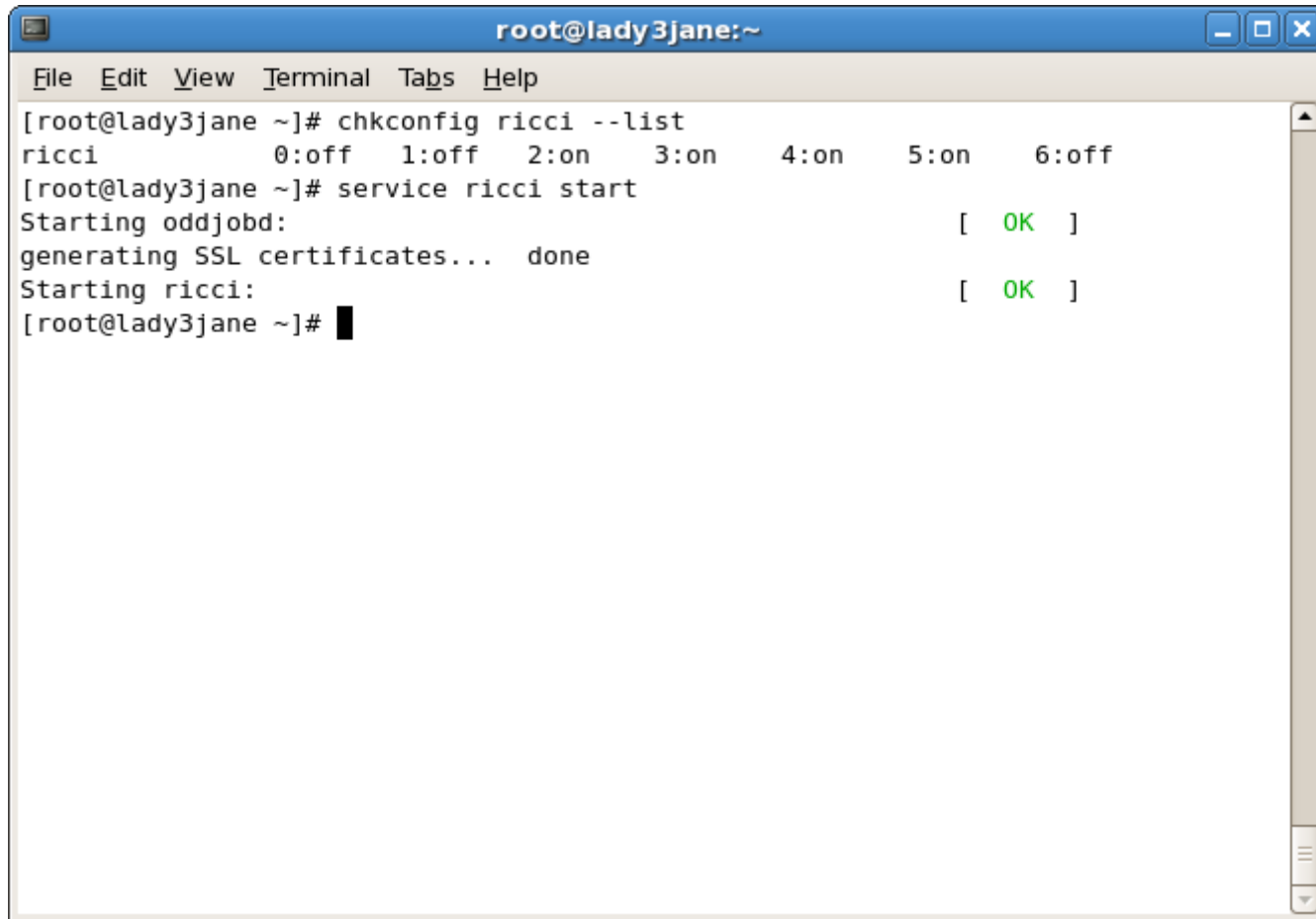
```
root@lady3jane:~
File Edit View Terminal Tabs Help
[tcameron@case ~]$ ssh -Y root@lady3jane
The authenticity of host 'lady3jane (172.31.100.11)' can't be established.
RSA key fingerprint is 2a:2a:a4:8f:1c:dc:9c:58:b2:47:4e:40:c4:66:f0:97.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'lady3jane,172.31.100.11' (RSA) to the list of known hosts.
root@lady3jane's password:
Last login: Thu Jun 5 16:55:39 2008
[root@lady3jane ~]#
```

```
root@thefinn:~
File Edit View Terminal Tabs Help
[tcameron@case ~]$ ssh -Y root@thefinn
The authenticity of host 'thefinn (172.31.100.13)' can't be established.
RSA key fingerprint is ff:2f:8e:ae:26:cb:e1:b4:ff:99:9c:fd:5e:92:31:6d.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'thefinn,172.31.100.13' (RSA) to the list of known hosts.
root@thefinn's password:
[root@thefinn ~]#
```

Installing the domU machines

- Make sure that `ricci` is set to start at bootup
- Make sure `ricci` is running

Installing the domU machines

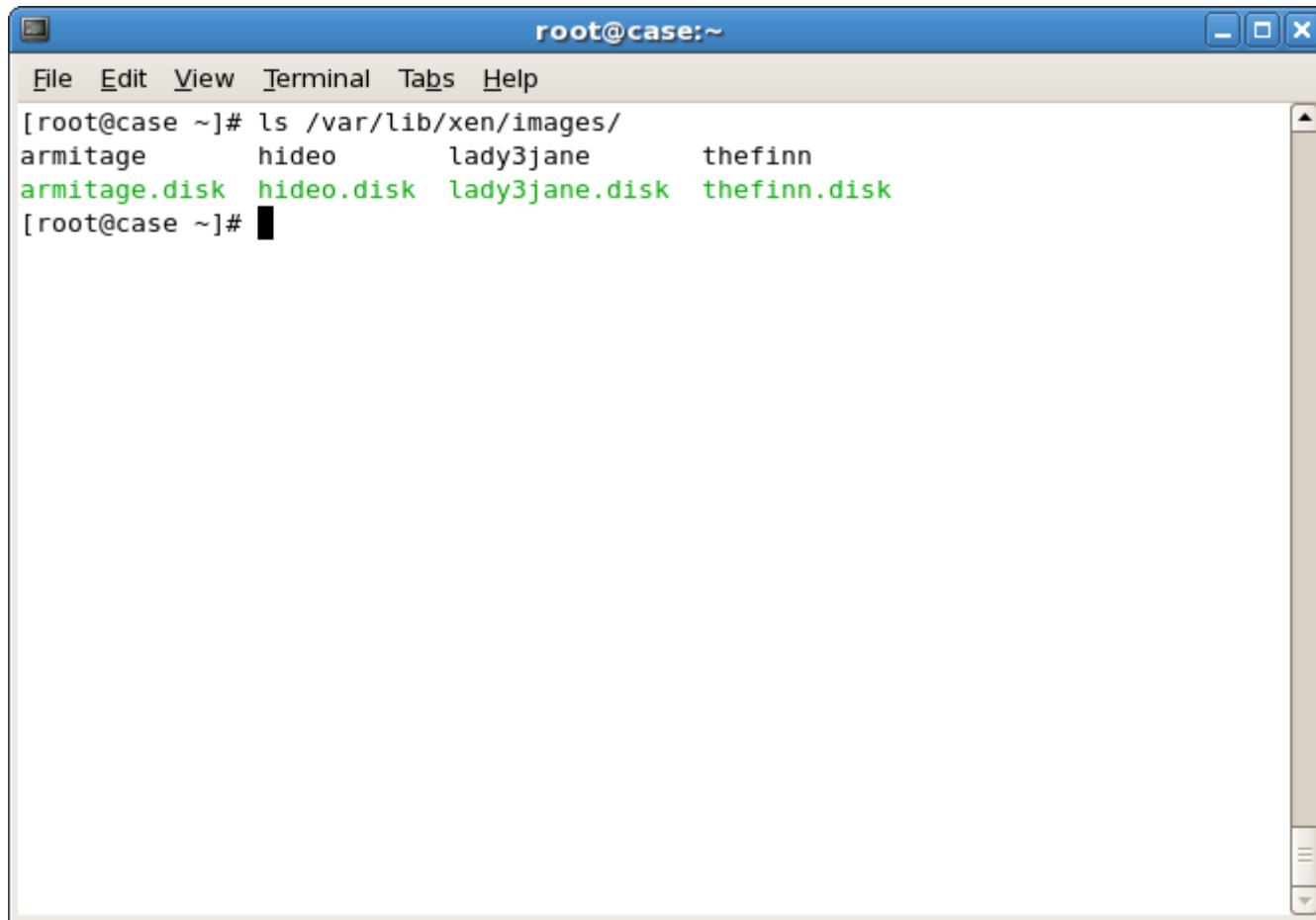
A terminal window titled 'root@lady3jane:~' with a menu bar (File, Edit, View, Terminal, Tabs, Help). The terminal shows the following commands and output:

```
[root@lady3jane ~]# chkconfig ricci --list
ricci          0:off  1:off  2:on   3:on   4:on   5:on   6:off
[root@lady3jane ~]# service ricci start
Starting oddjobd:           [ OK ]
generating SSL certificates... done
Starting ricci:             [ OK ]
[root@lady3jane ~]#
```

Define the domUs as cluster resources

- Since the domU cluster will be a clustered service, disable the per-dom0 xendomains service
 - `chkconfig xendomains off`
 - `service xendomains stop`
- Move the config files for each domU from `/etc/xen` to `/var/lib/xen/images`

Create the domU Cluster

A terminal window titled 'root@case:~' with a menu bar (File, Edit, View, Terminal, Tabs, Help). The terminal shows the command 'ls /var/lib/xen/images/' and its output: 'armitage hideo lady3jane thefinn' on the first line and 'armitage.disk hideo.disk lady3jane.disk thefinn.disk' on the second line. The prompt '[root@case ~]#' is followed by a cursor.

```
root@case:~  
File Edit View Terminal Tabs Help  
[root@case ~]# ls /var/lib/xen/images/  
armitage      hideo      lady3jane    thefinn  
armitage.disk hideo.disk lady3jane.disk thefinn.disk  
[root@case ~]#
```

Define the domUs as cluster resources

- Define each dom0 as a failover domain
 - prefer_case
 - prefer_molly
- Choose the dom0 cluster, go to “Failover Domains,” then “Add a Failover Domain,” then create the domains
- Set prioritization so that the preferred machine has a lower number

cluster — luci - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index_html?pagetype=44&domname=pre Google

Most Visited Smart Bookmarks Red Hat, Inc. Red Hat Network Support Shop Products Training

 **redhat**  **CLUSTER AND STORAGE SYSTEMS**

homebase cluster storage help log out

clusters

- Cluster List
- Create a New Cluster
- Configure

dom0-cluster

- Nodes
- Services
- Resources
- Failover Domains
 - Add a Failover Domain
 - Configure a Failover Domain
 - prefer_molly**
 - prefer_case
- Shared Fence Devices

dom0-cluster

Failover Domain Form

Failover Domain Name

Prioritized ☒

Restrict failover to this domain's members ☐

Do not fail back services in this domain ☐

Failover domain membership

Node	Member	Priority
case.redhat.lan	☒	10
molly.redhat.lan	☒	1

The Conga Cluster and Storage Management System is Copyright © 2000–2008 Red Hat, Inc.
Distributed under the GNU GPL license.

Done wintermute.redhat.lan:8084

cluster — luci - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index_html?pagetype=44&fdomname=pr

Most Visited Smart Bookmarks Red Hat, Inc. Red Hat Network Support Shop Products Training

redhat CLUSTER AND STORAGE SYSTEMS

homebase cluster storage help log out

clusters

- Cluster List
- Create a New Cluster
- Configure

dom0-cluster

- Nodes
- Services
- Resources
- Failover Domains
 - Add a Failover Domain
 - Configure a Failover Domain
 - prefer_molly
 - prefer_case**
- Shared Fence Devices

dom0-cluster

Failover Domain Form

Failover Domain Name prefer_case

Prioritized ☒

Restrict failover to this domain's members ☐

Do not fail back services in this domain ☐

Failover domain membership

Node	Member	Priority
case.redhat.lan	☒	1
molly.redhat.lan	☒	10

Submit

The Conga Cluster and Storage Management System is Copyright © 2000–2008 Red Hat, Inc.
Distributed under the GNU GPL license.

Done wintermute.redhat.lan:8084

Luci — cluster — failover domains - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index_html?pagetype=40&clustername=d

Most Visited Smart Bookmarks Red Hat, Inc. Red Hat Network Support Shop Products Training

redhat CLUSTER AND STORAGE SYSTEMS

homebase cluster storage help log out

clusters

- Cluster List
- Create a New Cluster
- Configure

dom0-cluster

- Nodes
- Services
- Resources
- Failover Domains
- Add a Failover Domain
- Configure a Failover Domain
- Shared Fence Devices

dom0-cluster

dom0-cluster

 Failover Domain Name: [prefer_molly](#)

This Failover Domain is Ordered

This Failover Domain is Unrestricted

Members of this Failover Domain

-  [case.redhat.lan](#)
-  [molly.redhat.lan](#)

Services employing this Failover Domain:

- No Services Defined

dom0-cluster

 Failover Domain Name: [prefer_case](#)

This Failover Domain is Ordered

This Failover Domain is Unrestricted

Members of this Failover Domain

-  [case.redhat.lan](#)
-  [molly.redhat.lan](#)

Services employing this Failover Domain:

- No Services Defined

Done wintermute.redhat.lan:8084

Create the domU Cluster

- Enable live migration by setting the following three settings in `/etc/xen/xend-config.sxp`:
 - `(xend-relocation-server yes)`
 - `(xend-relocation-port 8002)`
 - `(xend-relocation-address '')`
- The `' '` in `xend-relocation-address` is two single quotes, not a double quote
- Be aware of the security implications of having this wide open, you can also specify a space separated list of regular expressions for hostnames or ip addresses
- Restart the xend service

Create the domU Cluster

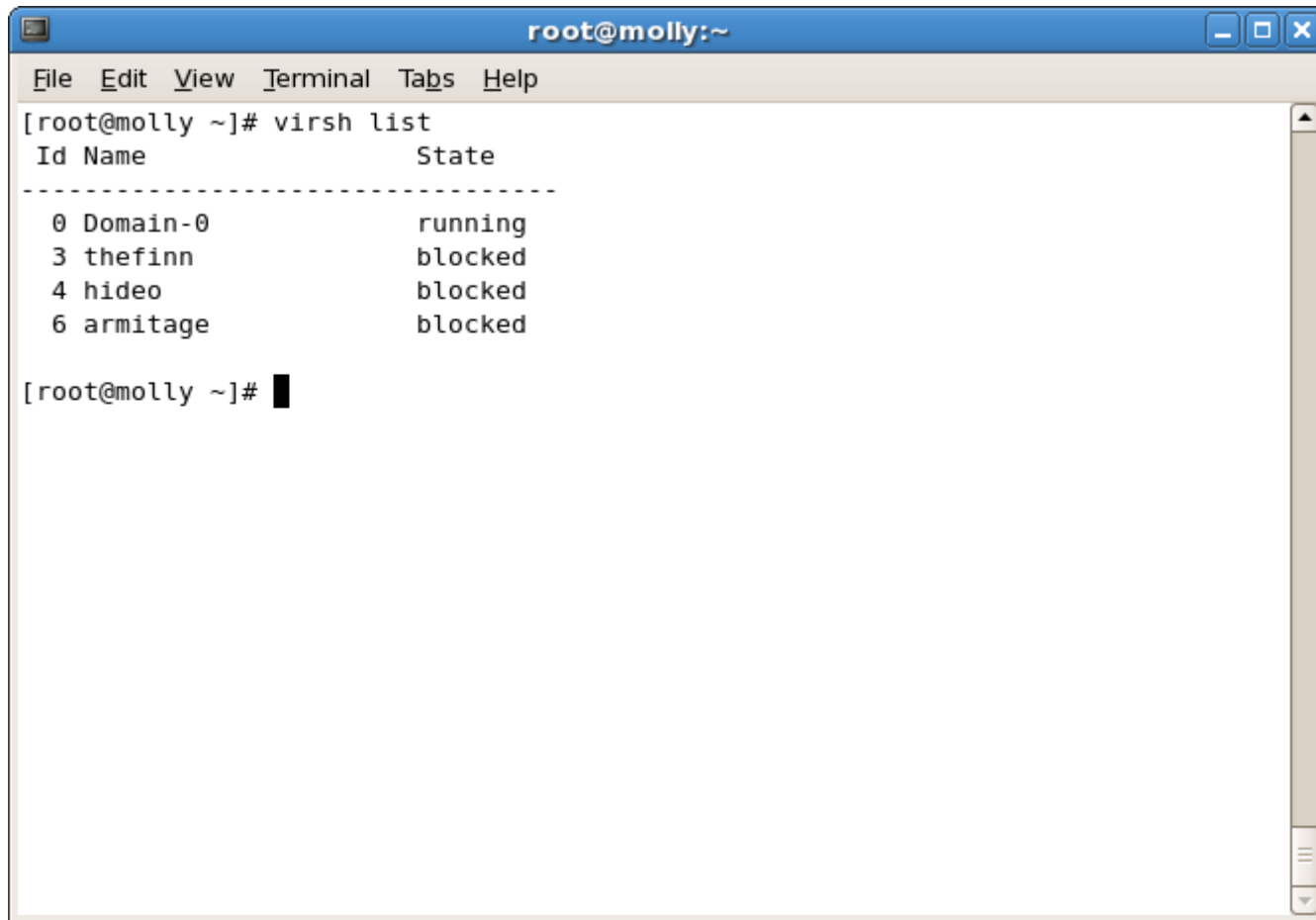
- Test live migration.
- From case, issue the command:
 - `xm migrate -l armitage molly`

Create the domU Cluster

A terminal window titled 'root@case:~' with a menu bar (File, Edit, View, Terminal, Tabs, Help). The terminal shows the execution of 'virsh list' and 'xm migrate' commands. The first 'virsh list' command shows three domains: 'Domain-0' (running), 'lady3jane' (blocked), and 'armitage' (blocked). The 'xm migrate armitage molly' command is executed. A second 'virsh list' command shows two domains: 'Domain-0' (running) and 'lady3jane' (blocked). The 'armitage' domain is no longer listed.

```
root@case:~  
File Edit View Terminal Tabs Help  
[root@case ~]# virsh list  
Id Name State  
-----  
0 Domain-0 running  
3 lady3jane blocked  
5 armitage blocked  
  
[root@case ~]# xm migrate armitage molly  
[root@case ~]# virsh list  
Id Name State  
-----  
0 Domain-0 running  
3 lady3jane blocked  
  
[root@case ~]#
```

Create the domU Cluster

A terminal window titled 'root@molly:~' with a menu bar (File, Edit, View, Terminal, Tabs, Help). The terminal shows the command '[root@molly ~]# virsh list' and its output, which is a table of virtual machines. The table has columns 'Id', 'Name', and 'State'. The first row is '0 Domain-0' with state 'running'. The next three rows are '3 thefinn', '4 hideo', and '6 armitage', all with state 'blocked'. The prompt '[root@molly ~]#' is followed by a cursor.

```
root@molly:~  
File Edit View Terminal Tabs Help  
[root@molly ~]# virsh list  
Id Name                State  
-----  
0  Domain-0             running  
3  thefinn               blocked  
4  hideo                 blocked  
6  armitage              blocked  
[root@molly ~]#
```

Define the domUs as cluster resources

- Define the domUs as resources for the dom0 to which they are assigned
- Choose the cluster, then “Services,” then “Add a Virtual Service.”
 - Name of domU
 - Path to config
 - Automatically start
 - Failover domain
 - Recovery policy
 - Migration type
- Do this for all domUs

cluster — luci - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index_html?pagetype=18&clustername=d

Most Visited Smart Bookmarks Red Hat, Inc. Red Hat Network Support Shop Products Training

 **redhat**  **CLUSTER AND STORAGE SYSTEMS**

homebase cluster storage help log out

clusters

- Cluster List
- Create a New Cluster
- Configure

dom0-cluster

- Nodes
- Services
 - Add a Service
 - Add a Virtual Service**
 - Configure a Service
- Resources
- Failover Domains
- Shared Fence
- Devices

dom0-cluster

Create a Virtual Machine Service

Virtual machine name:

Path to VM configuration files:

Automatically start this service: ☒

Run exclusive: ☐

Failover Domain:

Recovery policy:

Migration type:

The Conga Cluster and Storage Management System is Copyright © 2000–2008 Red Hat, Inc.
Distributed under the GNU GPL license.

Done wintermute.redhat.lan:8084

Luci — cluster — services - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index_html?pagetype=20&clustername=d

Most Visited Smart Bookmarks Red Hat, Inc. Red Hat Network Support Shop Products Training

redhat CLUSTER AND STORAGE SYSTEMS

homebase cluster storage help log out

clusters

- Cluster List
- Create a New Cluster
- Configure

dom0-cluster

- Nodes
- Services
 - Add a Service
 - Add a Virtual Service
 - Configure a Service
- Resources
- Failover Domains
- Shared Fence
- Devices

dom0-cluster

Service Name armitage (virtual service) Choose a Task... Go

Status: Running on case.redhat.lan
Autostart is enabled for this service

Failover Domain Association: prefer_case

Service Name lady3jane (virtual service) Choose a Task... Go

Status: Running on case.redhat.lan
Autostart is enabled for this service

Failover Domain Association: prefer_case

Service Name hideo (virtual service) Choose a Task... Go

Status: Running on molly.redhat.lan
Autostart is enabled for this service

Failover Domain Association: prefer_molly

Service Name thefinn (virtual service) Choose a Task... Go

Status: Running on molly.redhat.lan
Autostart is enabled for this service

Failover Domain Association: prefer_molly

Done wintermute.redhat.lan:8084

Define the domUs as cluster resources

- Test that Conga will migrate a domU to a different dom0

Luci — cluster — services - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index_html?pagetype=20&clustername=d

Most Visited Smart Bookmarks Red Hat, Inc. Red Hat Network Support Shop Products Training

 **redhat**  **CLUSTER AND STORAGE SYSTEMS**

homebase cluster storage help log out

clusters

- Cluster List
- Create a New Cluster
- Configure

dom0-cluster

- Nodes
- Services**
 - Add a Service
 - Add a Virtual Service
 - Configure a Service
- Resources
- Failover Domains
- Shared Fence
- Devices

dom0-cluster

Service Name armitage (virtual service)

Status: Running on case.redhat.lan
Autostart is enabled for this service

Failover Domain Association: prefer_case

Service Name lady3jane (virtual service)

Status: Running on case.redhat.lan
Autostart is enabled for this service

Failover Domain Association: prefer_case

Service Name hideo (virtual service)

Status: Running on molly.redhat.lan
Autostart is enabled for this service

Failover Domain Association: prefer_molly

Service Name thefinn (virtual service)

Status: Running on molly.redhat.lan
Autostart is enabled for this service

Failover Domain Association: prefer_molly

Done wintermute.redhat.lan:8084

Luci — cluster — services - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index_html?pagetype=22&clustername=d

Most Visited Smart Bookmarks Red Hat, Inc. Red Hat Network Support Shop Products Training

 **redhat**  CLUSTER AND STORAGE SYSTEMS

homebase cluster storage help log out

clusters

- Cluster List
- Create a New Cluster
- Configure

dom0-cluster

- Nodes
- Services**
 - Add a Service
 - Add a Virtual Service
 - Configure a Service
- Resources
- Failover Domains
- Shared Fence
- Devices

dom0-cluster

Service Name lady3jane (virtual service) Choose a Task... Go

Status: Running on case.redhat.lan
Autostart is enabled for this service

Failover Domain Association: prefer_case

Service Name armitage (virtual service) Choose a Task... Go

Status: Running on molly.redhat.lan
Autostart is enabled for this service

Failover Domain Association: prefer_case

Service Name hideo (virtual service) Choose a Task... Go

Status: Running on molly.redhat.lan
Autostart is enabled for this service

Failover Domain Association: prefer_molly

Service Name thefinn (virtual service) Choose a Task... Go

Status: Running on molly.redhat.lan
Autostart is enabled for this service

Failover Domain Association: prefer_molly

Done wintermute.redhat.lan:8084

Set up the Apache cluster

- Using similar steps as the dom0 cluster, set up a cluster of domUs
- Set up a prioritized failover domain which sets the service to prefer armitage, then hideo, then lady3jane, then thefinn

Luci — cluster — Deploy a cluster - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index_html?pagetype=6

Most Visited Smart Bookmarks Red Hat, Inc. Red Hat Network Support Shop Products Training

redhat CLUSTER AND STORAGE SYSTEMS

homebase cluster storage help log out

clusters

- Cluster List
- Create a New Cluster
- Configure

Create a new cluster

Cluster Name

Node Hostname	Root Password	Key ID
armitage.redhat.lan		 
lady3jane.redhat.lan		 
hideo.redhat.lan		 
thefinn.redhat.lan		 

☒ Download packages
☐ Use locally installed packages.

☒ Enable Shared Storage Support
☒ Reboot nodes before joining cluster
☒ Check if node passwords are identical.

The Conga Cluster and Storage Management System is Copyright © 2000–2008 Red Hat, Inc.
Distributed under the GNU GPL license

Done wintermute.redhat.lan:8084

Luci — cluster — cluster list - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index_html?pagetype=3

Most Visited Smart Bookmarks Red Hat, Inc. Red Hat Network Support Shop Products Training

redhat CLUSTER AND STORAGE SYSTEMS

homebase cluster storage help log out

clusters

- Cluster List
- Create a New Cluster
- Configure

Choose a cluster to administer

Cluster Name: dom0-cluster Restart this cluster Go

- Status: Quorate
- Total Cluster Votes: 2
- Minimum Required Quorum: 1

Nodes

- case.redhat.lan
- molly.redhat.lan

Services

- armitage
- lady3jane
- hideo
- thefinn

Cluster Name: apache-cluster Restart this cluster Go

- Status: Quorate
- Total Cluster Votes: 4
- Minimum Required Quorum: 3

Nodes

- armitage.redhat.lan
- hideo.redhat.lan
- lady3jane.redhat.lan
- thefinn.redhat.lan

Services

- No Services Defined

Done wintermute.redhat.lan:8084

Luci — cluster — failover domains — Add a failover domain - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index_html?pagetype=41&clustername=a

Google

Most Visited

Smart Bookmarks

Red Hat, Inc.

Red Hat Network

Support

Shop

Products

Training

redhat

CLUSTER AND STORAGE SYSTEMS

homebase

cluster

storage

help log out

clusters

Cluster List

Create a New Cluster

Configure

apache-cluster

Nodes

Services

Resources

Failover Domains

Add a Failover Domain

Configure a Failover Domain

Shared Fence Devices

apache-cluster

Add a Failover Domain

Failover Domain Name

prefer_armitage

Prioritized

☒

Restrict failover to this domain's members

☐

Do not fail back services in this domain

☐

Failover domain membership

Node	Member	Priority
lady3jane.redhat.lan	☒	20
hideo.redhat.lan	☒	10
armitage.redhat.lan	☒	1
thefinn.redhat.lan	☒	30

Submit

The Conga Cluster and Storage Management System is Copyright © 2000–2008 Red Hat, Inc.

Distributed under the GNU GPL license.

Done

wintermute.redhat.lan:8084

THE FOURTH ANNUAL
RED HAT SUMMIT

Set up the Apache cluster

- Define the following resources in apache-cluster
 - A script resource
 - `/etc/rc.d/init.d/httpd`
 - An ip address resource
 - `172.31.100.20`

Luci — cluster — resources — Add a resource - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index_html?pagetype=31& refresh tag

Most Visited Smart Bookmarks Red Hat, Inc. Red Hat Network Support Shop Products Training

Luci — cluster — resour... Red Hat -- ThomasCameron...

 **redhat**  **CLUSTER AND STORAGE SYSTEMS**

homebase cluster storage help log out

clusters

- Cluster List
- Create a New Cluster
- Configure

apache-cluster

- Nodes
- Services
- Resources
- Add a Resource**
- Configure a Resource
- Failover Domains
- Shared Fence
- Devices

apache-cluster

Add a Resource

IP Address Resource Configuration

IP address

Monitor link ☒

x Find: Previous Next Highlight all Match case

Done wintermute.redhat.lan:8084

Luci — cluster — resources — Add a resource - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index_html?pagetype=31& refresh tag

Most Visited Smart Bookmarks Red Hat, Inc. Red Hat Network Support Shop Products Training

Luci — cluster — resour... Red Hat -- ThomasCameron...

 **redhat**  CLUSTER AND STORAGE SYSTEMS

homebase cluster storage help log out

clusters

- Cluster List
- Create a New Cluster
- Configure

apache-cluster

- Nodes
- Services
- Resources
- Add a Resource**
- Configure a Resource
- Failover Domains
- Shared Fence
- Devices

apache-cluster

Add a Resource

Script Resource Configuration

Name

Full path to script file

Find: Previous Next Highlight all Match case

Done wintermute.redhat.lan:8084

Luci — cluster — resources - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index_html?pagetype= refresh tag

Most Visited Smart Bookmarks Red Hat, Inc. Red Hat Network Support Shop Products Training

Luci — cluster — resour... x Red Hat -- ThomasCameron... x

 **redhat**  **CLUSTER AND STORAGE SYSTEMS**

homebase cluster storage help log out

clusters

- Cluster List
- Create a New Cluster
- Configure

apache-cluster

- Nodes
- Services
- Resources**
- Add a Resource
- Configure a Resource
- Failover Domains
- Shared Fence
- Devices

apache-cluster

Resources for apache-cluster

Resource Name	Type	Configure	Delete
172.31.100.20	IP Address	configure	delete
webserver	Script	configure	delete

x Find: paste Previous Next Highlight all Match case

Done wintermute.redhat.lan:8084

Set up the Apache cluster

- Assign those resources to a service
 - `web-cluster`
- The script resource should be a child of the ip address resource, as the ip address needs to be up before Apache will bind to it

Luci — cluster — services — Add a new service - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index_html?pagetype= refresh tag

Most Visited Smart Bookmarks Red Hat, Inc. Red Hat Network Support Shop Products Training

Luci — cluster — service... x Red Hat -- ThomasCameron... x

clusters

- Cluster List
- Create a New Cluster
- Configure

apache-cluster

- Nodes
- Services
- Add a Service
- Configure a Service
- Resources
- Failover Domains
- Shared Fence
- Devices

apache-cluster

Add a Service

Service name web-cluster

Automatically start this service ☒

Run exclusive ☐

Failover Domain prefer_armitage

Recovery policy Relocate

IP Address Resource Configuration

IP address 172.31.100.20

Monitor link ☒

This resource is an independent subtree ☐

Add a child Delete this resource

Hide Children

Script Resource Configuration

Name webserver

Full path to script file /etc/rc.d/init.d/htt

This resource is an independent subtree ☐

Add a child Delete this resource

Add a resource to this service Submit

x Find: paste Previous Next Highlight all Match case

Done wintermute.redhat.lan:8084

Luci — cluster — services - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index_html?pagetype= refresh tag

Most Visited Smart Bookmarks Red Hat, Inc. Red Hat Network Support Shop Products Training

 **redhat**  **CLUSTER AND STORAGE SYSTEMS**

homebase cluster storage help log out

clusters

- Cluster List
- Create a New Cluster
- Configure

apache-cluster

- Nodes
- Services**
 - Add a Service
 - Configure a Service
- Resources
- Failover Domains
- Shared Fence
- Devices

apache-cluster

Service Name web-cluster Choose a Task... Go

Status: Running on armitage.redhat.lan
Autostart is enabled for this service

Failover Domain Association: prefer_armitage

x Find: paste Previous Next Highlight all Match case

Done wintermute.redhat.lan:8084

Add a virtual fence device

- Go to the domU cluster
 - Shared fence devices
 - Add fence device
 - Virtual machine fencing
 - Choose a name

Luci — cluster — fence devices - Add a new fence device - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index_html?pagetype=51&cl

Most Visited Smart Bookmarks Red Hat, Inc. Red Hat Network Support Shop Products Training

Luci — cluster — fence ... x http://172.31.100.20/ x

 **redhat**  **CLUSTER AND STORAGE SYSTEMS**

homebase cluster storage help log out

clusters

- Cluster List
- Create a New Cluster
- Configure

domu-cluster

- Nodes
- Services
- Resources
- Failover Domains
- Shared Fence Devices
- Add a Fence Device**
- Configure a Fence Device

domU-cluster

Add a Sharable Fence Device

Fencing Type

Virtual Machine Fencing

Fence Type Virtual Machine Fencing

Name

Done wintermute.redhat.lan:8084

Add a virtual fence device

- Go to a domU machine
- Fencing
- Add fence device to this level
- Select your fence device
- Type virtual domain name (domU name)
- Update main fence properties
- Repeat for each domU

Luci — cluster — nodes — properties - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index_html?pagetype=9&clus

Most Visited Smart Bookmarks Red Hat, Inc. Red Hat Network Support Shop Products Training

thefinn.redhat.lan

- Services
- Resources
- Failover Domains
- Shared Fence Devices

Services on this Node

web-cluster

Failover Domain Membership

prefer_armitage

Main Fencing Method

Fence Type Virtual Machine Fencing

Name

Domain

[Add a fence device to this level](#)

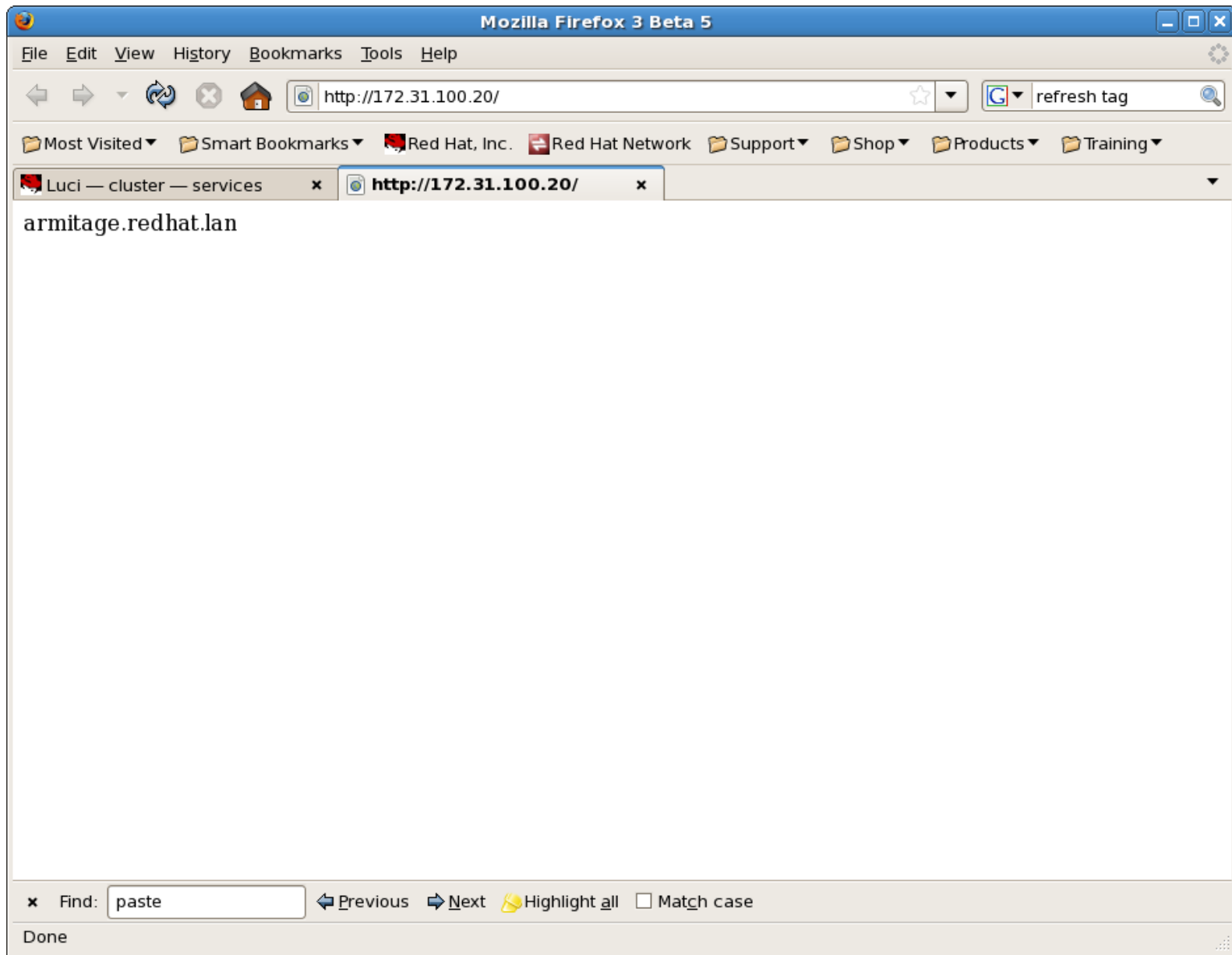
Backup Fencing Method

[Add a fence device to this level](#)

Done wintermute.redhat.lan:8084

Set up the Apache cluster

- To test, you can do something simple like
 - `hostname > /var/www/html/index.html`



Test the Apache cluster

- Fail the web-cluster service to another domU

Luci — cluster — services - Mozilla Firefox 3 Beta 5

File Edit View History Bookmarks Tools Help

https://wintermute.redhat.lan:8084/luci/cluster/index_html?pagetype= refresh tag

Most Visited Smart Bookmarks Red Hat, Inc. Red Hat Network Support Shop Products Training

Luci — cluster — services x http://172.31.100.20/ x

 **redhat**  **CLUSTER AND STORAGE SYSTEMS**

homebase cluster storage help log out

clusters

- Cluster List
- Create a New Cluster
- Configure

apache-cluster

- Nodes
- Services**
- Add a Service
- Configure a Service
- Resources
- Failover Domains
- Shared Fence
- Devices

apache-cluster

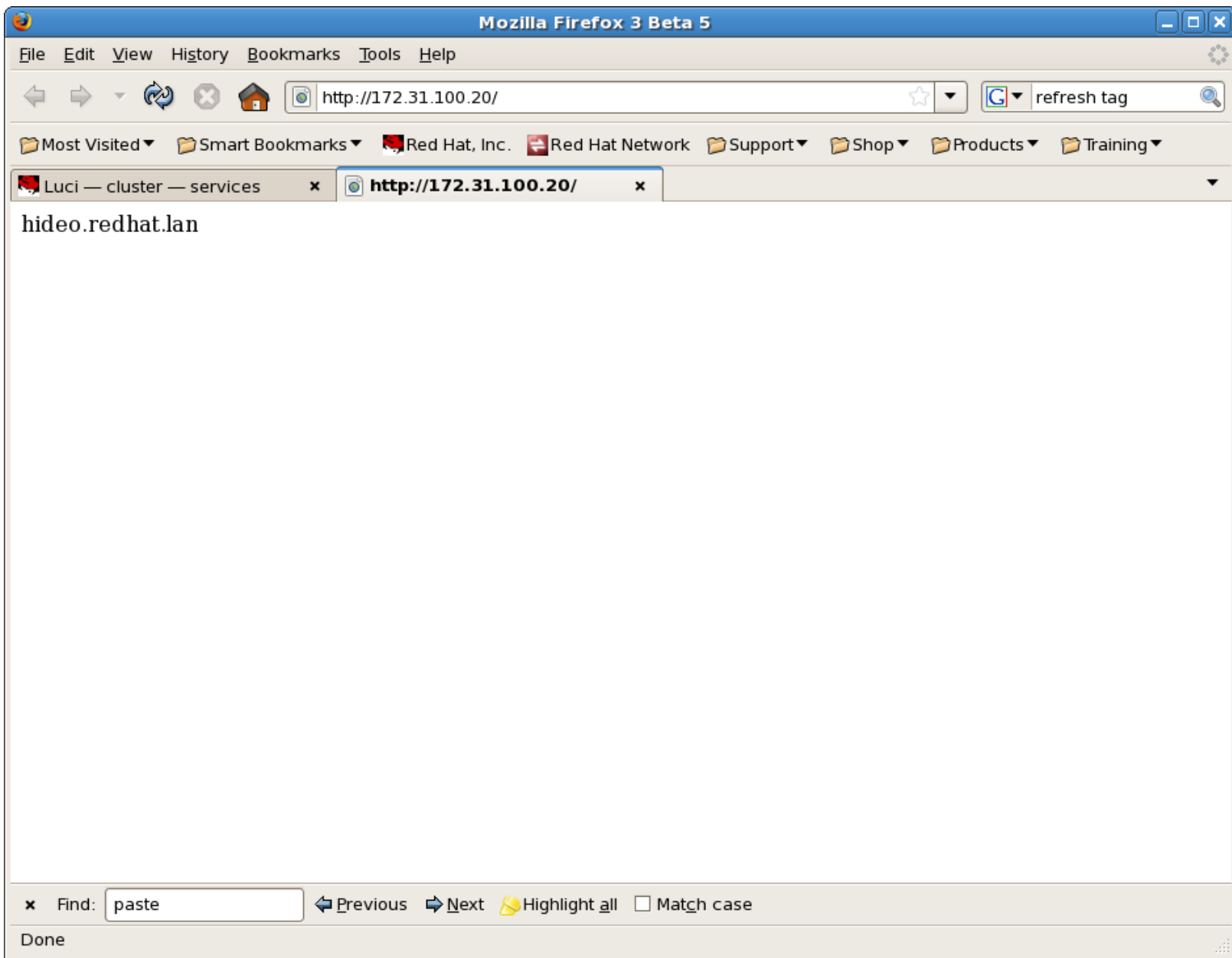
Service Name web-cluster Relocate this service to hideo.redhat.lan Go

Status: Running on armitage.redhat.lan
Autostart is enabled for this service

Failover Domain Association: prefer_armitage

x Find: paste Previous Next Highlight all Match case

Done wintermute.redhat.lan:8084



Test the Apache cluster

- Or you can “kick the plug” out of the domU
 - `virsh destroy hideo`
 - `echo s > /proc/sysrq-trigger ; echo b > /proc/sysrq-trigger`

This presentation is available online at
<http://people.redhat.com/tcameron>

- You can contact Thomas Cameron at tcameron@redhat.com
- You can contact Lon Hohberger at lh@redhat.com
- Thank you very much for attending!



Questions?