

# Cluster usando ganeti

## Instalación

### Bookmarks

- <http://docs.ganeti.org/ganeti/current/html/install.html>
- <http://code.google.com/p/ganeti/downloads/list>
- <http://docs.ganeti.org/ganeti/2.2/html/walkthrough.html>
- [http://groups.google.com/group/ganeti/browse\\_thread/thread/1357a84a3304a724/a325f4890df56fd?pli=1](http://groups.google.com/group/ganeti/browse_thread/thread/1357a84a3304a724/a325f4890df56fd?pli=1)
- <http://docs.ganeti.org/ganeti/current/html/admin.html>
- <http://code.google.com/p/ganeti/issues/detail?id=66>
- [http://wiki.osuosl.org/public/ganeti/common\\_commands](http://wiki.osuosl.org/public/ganeti/common_commands)

### Instalación de paquetes

Se modifican los archivos:

- /etc/xen/xend-config.sxp
- /etc/default/xendomains

```
mv /etc/grub.d/10_linux /etc/grub.d/50_linux
```

de acuerdo a las indicaciones de instalacion. Luego hace falta reconfigurar grub ejecutando **update-grub**

**CUIDADO:** La opción nosmp que indica la guía de instalación está obsoleta (deprecated) y además si se actualiza /etc/default/grub más un update-grub se modifican los parámetros del kernel y esto NO funciona. Lo que hay que hacer es agregar directamente los parametros dom0\_mem=256M maxcpus=1 en la línea que dice multiboot.

En el /etc/host hace falta definir ip para el nombre del cluster, de los nodos y de las instancias, por ejemplo:

```
10.4.0.101 e1.rec.unicen.edu.ar e1
10.4.0.101 eln1.rec.unicen.edu.ar eln1
10.4.0.102 eln2.rec.unicen.edu.ar eln2
10.4.0.104 elm1.rec.unicen.edu.ar elm1
10.4.0.105 elm2.rec.unicen.edu.ar elm2
10.4.0.106 elm3.rec.unicen.edu.ar elm3
```

Se instala

- drbd8-utils

y se edita /etc/modules para agregar:

```
drbd minor_count=128 usermode_helper=/bin/true
```

y se comenta el contenido del archivo `/etc/drbd.conf` ya que ganeti se ocupa de la configuracion en adelante.

Se instalan los paquetes:

- lvm2 ssh bridge-utils iproute iputils-arping python python-twisted-core python-pyopenssl openssl
- mdadm python-pyparsing python-simplejson debootstrap dump kpartx
- pandoc graphviz pylint socat python-sphinx python-pyinotify python-pycurl python-paramiko make

NOTA: Antes de hacer esto, si se configuró Xen como en la otra wiki, hay que comentar el bridge de xen en `/etc/xen/xend-config.sxp`

Se convierten la `eth0` en bridge, reemplazandola por la sig configuracion:

```
auto xen-br0
iface xen-br0 inet static
    address 10.4.0.101
    netmask 255.255.255.0
    network 10.4.0.0
    broadcast 10.4.0.255
    gateway 10.4.0.1
    bridge_ports eth2
    bridge_stp off
    bridge_fd 0
    dns-nameservers 10.4.0.1
    dns-search rec.unicen.edu.ar
```

Se crea el volumen para el almacenamiento usando LVM. Ya existian 3 file systems e cada equipo (5,6 y 7)

```
pvccreate /dev/sda5          # uno por cada dis
vgcreate xenvg /dev/sda5 /dev/sda6 /dev/sda7
```

Se agrega el filtro a `/etc/lvm/lvm.conf`

```
#GANETI# filter = [ "a/*/" ]
filter = ["r|/dev/cdrom|", "r|/dev/drbd[0-9]+|" ]
```

Se instalan los paquetes que se bajan desde google

```
cd /opt
wget http://ganeti.googlecode.com/files/ganeti-2.4.1.tar.gz
tar -xvzf ganeti-2.4.1.tar.gz
cd ganeti-2.4.1/
./configure --localstatedir=/var --sysconfdir=/etc
make
make install
```

```
mkdir -p /srv/ganeti/ /srv/ganeti/os /srv/ganeti/export
cp doc/examples/ganeti.initd /etc/init.d/ganeti
chmod 755 /etc/init.d/ganeti
update-rc.d ganeti defaults 20 80
cd /usr/local
cd src
wget
http://ganeti.googlecode.com/files/ganeti-instance-debootstrap-0.9.tar.gz
tar xzf ganeti-instance-debootstrap-0.9.tar.gz
cd ganeti-instance-debootstrap-0.9
./configure
make
make install
```

Se detecta un problema con los sistemas operativos disponibles, que están vacíos:

```
gnt-os list
gnt-os diagnose
```

y se corrige:

```
cd /usr/local/share/ganeti/os
cp -r debootstrap /srv/ganeti/os/
```

## Creación de cluster, usando dos interfaces

Las direcciones secundarias serán utilizadas para el tráfico interno del cluster. Es obligatorio, que si el cluster se crea con esta modalidad, **todos** los nodos también se creen de la misma manera.

```
gnt-cluster init -s 10.4.11.1 e1
```

El agregado de un nodo se hace con

```
gnt-node add -s 10.4.11.2 eln2.rec.unicen.edu.ar
gnt-cluster modify --hypervisor-parameter xen-pvm:root_path='/dev/xvda1'
```

Los comandos “gnt-cluster info” y “gnt-node info” se pueden usar para corroborar que se ha creado usando direcciones secundarias.

## Creación de instancias con diferentes sisop y diferentes arquitecturas

Es necesario agregar los kernels y los file systems de boot para las diferentes versiones y arquitecturas. Esto es en **todos los nodos** Copio al /boot lo necesario de versiones de debian y creo links simbólicos para que sea mas claro:

```
lrwxrwxrwx 1 root root          29 May 18 10:02 initrd-d5-686-xenU ->
```

```

initrd.img-2.6.26-2-xen-686
lrwxrwxrwx 1 root root      29 May 18 10:02 initrd-d5-amd64-xenU ->
initrd.img-2.6.26-2-xen-amd64
lrwxrwxrwx 1 root root      35 May 12 11:29 initrd-d6-amd64-xenU ->
/boot/initrd.img-2.6.32-5-xen-amd64
-rw-r--r-- 1 root root 6135019 May 18 09:19 initrd.img-2.6.26-2-xen-686
-rw-r--r-- 1 root root 6964969 May 18 10:01 initrd.img-2.6.26-2-xen-amd64
-rw-r--r-- 1 root root 10380246 May 12 11:50 initrd.img-2.6.32-5-xen-amd64
-rw-r--r-- 1 root root 1485914 May 18 09:19 vmlinuz-2.6.26-2-xen-686
-rw-r--r-- 1 root root 1706059 May 18 10:01 vmlinuz-2.6.26-2-xen-amd64
-rw-r--r-- 1 root root 2475808 Mar  7 21:44 vmlinuz-2.6.32-5-xen-amd64
lrwxrwxrwx 1 root root      24 May 18 09:20 vmlinuz-d5-686-xenU ->
vmlinuz-2.6.26-2-xen-686
lrwxrwxrwx 1 root root      26 May 18 10:01 vmlinuz-d5-amd64-xenU ->
vmlinuz-2.6.26-2-xen-amd64
lrwxrwxrwx 1 root root      32 May 12 11:29 vmlinuz-d6-amd64-xenU ->
/boot/vmlinuz-2.6.32-5-xen-amd64

```

Modifico el /etc/default/ganeti-instance-debootstrap para explicitamente fijar parametros para definir que tipo de instancia quiero crear (algo parecido al /etc/xen-tools). Creo instancia inicialmente SIN drdb ya que será mas facil manimular el contenido del disco. Luego se puede cambiar el tipo.

## Nuevos repos de Debian

```

#deb http://debian.unicen.edu.ar:9999/debian lenny main contrib non-free
#deb http://debian.unicen.edu.ar:9999/security lenny/updates main contrib
non-free
#deb http://debian.unicen.edu.ar:9999/backports lenny-backports main contrib
non-free
deb http://debian.unicen.edu.ar:9999/archive lenny main contrib non-free
deb http://debian.unicen.edu.ar:9999/secarchive lenny/updates main contrib
non-free

```

## Creación de instancia con Debian 5 de 64 bits

```

PROXY="http://proxy.unicen.edu.ar:8080/"
MIRROR="http://debian.unicen.edu.ar:9999/debian"
ARCH="amd64"
SUITE="lenny"

gnt-instance add -t plain -o debootstrap+default -s 20g -n
elml.rec.unicen.edu.ar -B memory=512 -H xen-pvm:kernel_path=/boot/vmlinuz-
d5-amd64-xenU,initrd_path=/boot/initrd-d5-amd64-xenU,root_path=/dev/sda1
elml.rec.unicen.edu.ar

```

que da como resultado una maquina operativa sin mas con las siguientes características

```
elml:~# cat /proc/version
```

```
Linux version 2.6.26-2-xen-amd64 (Debian 2.6.26-25) (dannf@debian.org) (gcc
version 4.1.3 20080704 (prerelease) (Debian 4.1.2-25)) #1 SMP Tue Aug 31
11:17:26 UTC 2010
elm1:~# cat /etc/debian_version
5.0.8
```

## Creación de instancia con Debian 5 de 32 bits

```
PROXY="http://proxy.unicen.edu.ar:8080/"
MIRROR="http://debian.unicen.edu.ar:9999/debian"
ARCH="i386"
SUITE="lenny"

gnt-instance add -t plain -o debootstrap+default -s 20g -n
eln1.rec.unicen.edu.ar -B memory=512 -H xen-pvm:kernel_path=/boot/vmlinuz-
d5-686-xenU,initrd_path=/boot/initrd-d5-686-xenU,root_path=/dev/sda1
elm2.rec.unicen.edu.ar
```

que da como resultado una maquina operativa sin mas con las siguientes características

```
elm2:~# cat /proc/version
Linux version 2.6.26-2-xen-686 (Debian 2.6.26-24) (dannf@debian.org) (gcc
version 4.1.3 20080704 (prerelease) (Debian 4.1.2-25)) #1 SMP Mon Jun 21
10:37:05 UTC 2010
elm2:~# cat /etc/debian_version
5.0.8
```

## Creación de instancia con Debian 6 de 64 bits:

En este caso, la creamos con 2 discos, para poder usar uno de swap area. Ojo, que a diferencia del concepto de particion, nuestra máquina ahora tendrá el disco xvda (con xvda1 como particion /) y el disco xvdb. El mismo queda disponible, pero no formateado.

```
PROXY="http://proxy.unicen.edu.ar:8080/"
MIRROR="http://debian.unicen.edu.ar:9999/debian"
ARCH="amd64"
SUITE="squeeze"

gnt-instance add -t plain -o debootstrap+default --disk 0:size=10G --disk
1:size=1G -n eln1.rec.unicen.edu.ar -B memory=512 -H xen-
pvm:kernel_path=/boot/vmlinuz-d6-amd64-xenU,initrd_path=/boot/initrd-d6-
amd64-xenU,root_path=/dev/xvda1 elm3.rec.unicen.edu.ar
```

que da como resultado una maquina operativa sin mas con las siguientes características

```
root@elm3:~# cat /proc/version
Linux version 2.6.32-5-xen-amd64 (Debian 2.6.32-31) (ben@decadent.org.uk)
(gcc version 4.3.5 (Debian 4.3.5-4) ) #1 SMP Tue Mar 8 00:01:30 UTC 2011
```

```
root@elm3:~# cat /etc/debian_version
6.0.1
```

## Migración de máquinas existentes

El camino será:

- Crear instancia con la misma arquitectura que el DomU en xen

```
root@eln1:/# gnt-instance add --no-ip-check --no-start --no-install -t drbd
-o debootstrap+default --disk 0:size=10G -n eln1:eln2 -B memory=512 -H xen-
pvm:kernel_path=/boot/vmlinuz.lenny.64,initrd_path=/boot/initrd.lenny.64,roo
t_path=/dev/xvda1 kolla.rec.unicen.edu.ar
```

- Activar, y preparar el disco nuevo

```
root@eln1:/# gnt-instance activate-disks kolla
eln1.rec.unicen.edu.ar:disk/0:/dev/drbd0
root@eln1:/# fdisk /dev/drbd0
Device contains neither a valid DOS partition table, nor Sun, SGI or OSF
disklabel
root@eln1:/# kpartx -av /dev/drbd0
add map drbd0p1 (253:15): 0 20964762 linear /dev/drbd0 63
root@eln1:/# mkfs.ext4 /dev/mapper/drbd0p1
mke2fs 1.41.12 (17-May-2010)
```

- Montar el disco nuevo

```
root@eln1:/# mount /dev/mapper/drbd0p1 /mnt
root@eln1:/# ls -l /mnt
total 16K
drwx----- 2 root root 16K Aug 31 11:25 lost+found
```

- Sincronizarlos vía “rsync -delete”
- Desmontar los discos

```
root@eln1:/# umount /mnt
root@eln1:/# kpartx -d /dev/drbd0
```

## Actualizacion de D5 a D6

Cambiar repos, y actualizar. Cambiar en master:

```
gnt-instance modify -H initrd_path='/boot/initrd.squeeze.64' diaguita-nueva
gnt-instance modify -H kernel_path='/boot/vmlinuz.squeeze.64' diaguita-nueva
gnt-instance modify -H kernel_args='ro' diaguita-nueva
gnt-instance modify -H root_path=/dev/xvda1 diaguita-nueva
```

## Montado de disco de instancia

```
root@eln1:/boot# gnt-instance activate-disks elm3
eln1.rec.unicen.edu.ar:None:/dev/xenvg/ce7e0c83-caf7-4e5c-
b126-44b597c095cc.disk0_data
```

Cuidado, que las sentencias siguientes se tienen que ejecutar en el host que tiene el disco físico que se está montando

```
root@eln1:/boot# kpartx -av /dev/xenvg/ce7e0c83-caf7-4e5c-
b126-44b597c095cc.disk0_data
add map xenvg-ce7e0c83--caf7--4e5c--b126--44b597c095cc.disk0_data1 (253:3):
0 41929649 linear /dev/xenvg/ce7e0c83-caf7-4e5c-b126-44b597c095cc.disk0_data
1
root@eln1:/boot# mount /dev/mapper/xenvg-ce7e0c83--caf7--4e5c--b126-
-44b597c095cc.disk0_data1 /mnt
```

## Sincronización

En este caso, usé una máquina existente y encendida. Considerar que en este caso se está copiando también la definición de la red. O bien se excluye también /etc/network, o bien se modifica antes de desmontar. Otra cuestión estará en /etc/fstab. Estas máquinas arrancan sin swap area. Hay que crearlo a mano, como un archivo.

```
rsync -azvH --numeric-ids --exclude=/proc --exclude=/tmp --exclude=/sys --
exclude=/mnt --exclude=/media --delete / root@10.4.0.101:/mnt/
```

Luego, hay que crear los directorios faltantes con los permisos adecuados:

```
cd /mnt
mkdir mnt media tmp proc sys
chmod 777 tmp
chmod o+t tmp
```

## Instancias con Kernel en el propio Dom0

### Cambios en la instancia

```
mkdir /boot/grub;
apt-get -y install linux-image-amd64 linux-headers-amd64 firmware-linux-free
firmware-linux-nonfree grub-legacy;
grub-set-default 1;
update-grub;
```

En ubuntu al parecer alcanza con instalar el paquete grub-legacy-ec2

Editar con

```
vi /boot/grub/menu.lst
```

y realizar las siguientes modificaciones.

1. En la línea "timeout 5" cambiar valor 5 por valor 10
2. La línea **comentada** que contiene: "# kopt=root" por "# kopt=root=UUID=xxxxxxxxxxxxxxxxxxxxxx ro console=hvc0 quiet"
3. La línea **comentada** que contiene: "# groot=" por "# groot=(hd0)"

Y ejecutar update-grub

UBUNTU: al parecer alcanza con instalar el paquete grub-legacy-ec2

SQUEEZE 32 bits: los paquetes a instalar son: linux-image-2.6.32-5-686 linux-headers-2.6.32-5-686 firmware-linux-free firmware-linux-nonfree grub-legacy

SQUEEZE 64 bits los paquetes de kernel a instalar son: linux-image-2.6-amd64 linux-headers-2.6-amd64 firmware-linux-free firmware-linux-nonfree grub-legacy

LENNY 32 bits: los paquetes de kernel a instalar son: linux-image-2.6.26-2-686 linux-headers-2.6.26-2-686. DE MOMENTO NO LOGRO BOOT LOCAL

## Cambios en el cluster

Hay que cambiar para indicar el path del bootloader y el uso de un bootloader

```
gnt-instance modify -H  
bootloader_path=/usr/lib/xen-4.1/bin/pygrub,use_bootloader=true,kernel_path=  
/boot/null,initrd_path=/boot/null nombre_instancia
```

Si hay que hacer vuelta atrás:

```
gnt-instance modify -H bootloader_path= nombre_instancia  
gnt-instance modify -H use_bootloader=False nombre_instancia  
gnt-instance modify -H kernel_path='/boot/path_kernel' nombre_instancia  
gnt-instance modify -H initrd_path='/boot/path_initrd' nombre_instancia
```

## Manipulación del cluster

### Redistribuir configuracion

```
gnt-cluster redist-conf
```



## Borrado de instancias y cluster

```
gnt-instance list
gnt-instance shutdown elm1
gnt-instance remove elm1
gnt-node list
gnt-node remove eln2.rec.unicen.edu.ar
gnt-cluster info
gnt-cluster destroy --yes-do-it
```

## Listar los locks

```
gnt-debug locks
```

## Listar y cancelar procesos

```
gnt-job list
gnt-job cancel nro
gnt-job info nro
```

## Verificación y Ajustes

- “gnt-cluster verify” Lista los problemas. Lista de errores y soluciones en: <http://docs.ganeti.org/ganeti/2.2/html/walkthrough.html>
- “hbal -C -m el.rec.unicen.edu.ar” Estudia el balanceo y recomienda soluciones.

Parametrizado así evita poner en el mismo equipo ciertas instancias y evita usar ciertos nodos

```
hbal --exclusion-tags=ntp,mail,dns -C -p -O nodo05 -m cluster.unicen.edu.ar
```

nueva versión de hbal

```
/usr/local/bin/hbal --exclusion-tags=ntp,mail,dns -C -L -p -G default
```

Para Rectorado, que no hay grupos, es más sencillo:

```
hbal -C -L
```

## Manipulación de los nodos

### Reboot de un nodo

- Localizar todos las instancias que tienen secundarios en ese nodo. Si puedo pasar todos los secundarios a un nodo donde no exista un primario de los mismos

(porque sino quedarían primario y secundario en el mismo nodo) puedo usar de evacuación, sino debo usar el segundo para hacerlo instancia por instancia para cada uno de ellos pasando el secundario a otro nodo.

```
gnt-node evacuate --new-secondary e1n1 e1n2
```

o bien

```
gnt-instance replace-disks --new-secondary e1n2 wichip
```

- Localizar todas las instancias que estan corriendo en el ese nodo. Migrar sin apagar a su secundario, con un comando parecido a:

```
gnt-instance migrate pentaho-desas
```

- Reiniciar el nodo.
- Volver a colocar los secundarios sobre el nodo original

```
gnt-instance replace-disks --new-secondary e1n3 wichip
```

- Recrear los primarios de las instancias que se migraron, o solamente activar los discos si ya se recuperaron.

```
gnt-instance replace-disks -p pentaho-desas o bien
```

```
gnt-instance activate-disks pentaho-desas
```

- Volver a correr las instancias sobre su primario

```
gnt-instance migrate pentaho-desas
```

## Manipulación de las instancias

Estas modificaciones requieren el reinicio de la instancia, o el apagado previo a la acción.

### Cambiar la cantidad de memoria y o cpus

```
gnt-instance modify -B memory=1024,vcpus=2 e1m3
```

### Cambiar disco de plano a replicado

Con instancia apagada. Se indica el nodo que será secundario.

```
gnt-instance modify -t drbd -n e1n2 e1m3
```

## Cambiar kernel

```
gnt-instance modify -H  
kernel_path=/boot/vmlinuz.wheezy.64,initrd_path=/boot/initrd.wheezy.64  
nombredelamaquina
```

## Agregado de swap a las VMs

```
dd if=/dev/zero of=/media/swapfile bs=1k count=524288
```

o bien

```
fallocate -l 1G /media/swapfile
```

```
mkswap /media/swapfile  
swapon /media/swapfile  
echo "/media/swapfile swap swap defaults 0 0" >> /etc/fstab
```

## Clonado de una instancia

Para clonar (hacer una copia con otro nombre) una instancia sobre el propio cluster se utiliza export/import.

- Se realiza el export de la instancia. Para esto la maquina debe estar apagada. En rectorado los backups se están realizando en e1n5, por lo que se usaría: "-n e1n5". Un ejemplo de nombre de máquina sería pilaga-prod. Todas las semanas se hace un backup de las máquinas, así que si ya existe este paso se puede saltar.

```
gnt-backup export -n nodo05 web.ihlla
```

- Mover o copiar lo exportado con el nombre nuevo en el nodo de exportacion. Esa tarea se realiza logueado como root en el nodo de destino (e1n5 en rectorado). Los nombres de directorio tienen FQDN. Para el caso anterior, el origen sería pilaga.rec.unicen.edu.ar y el destino un nombre que debemos agregar en el DNS.

```
cd /var/lib/ganeti/export  
tar -xvzf web.ihlla.unicen.edu.ar.tgz  
mv web.ihlla.unicen.edu.ar ftp.ihlla.unicen.edu.ar
```

- Editar el archivo config.ini y cambiar los parametros indicados: name: web por ftp

```
cd ftp.ihlla.unicen.edu.ar  
vi config.ini
```

- Hay que agregar el nuevo nombre e ip en el DNS que corresponda. Para el caso de rectorado hay que ingresar al host dns.rec.unicen.edu.ar (no olvidar de cambiar la fecha en el serial)

```
vi /var/cache/bind/named.hosts
```

```
/etc/init.d/bind9 restart
```

- Si la instancia ya existe, para evitar el rebalanceo del cluster es preferible crearla en los mismos nodos. Primero se obtiene la info de la misma ("Nodes"), y luego se borra.

```
gnt-instance info nombredelainstancia  
gnt-instance remove nombredelainstancia
```

- Se importa la instancia. Esto se realiza en el nodo master del cluster. En rectorado esto es en e1n1. Los argumentos de "-n" son donde se va a crear el disco nuevo. Si no usa disco espejado, tendrá un solo nombre de host y en vez de -t drbd será -t plain. En rectorado los valores posibles son de e1n1 a e1n6.

```
gnt-backup import -n nodo04:nodo01 --net 0:mac=generate --src-node=nodo05 -t  
drbd ftp.ihlla.unicen.edu.ar
```

- La máquina nueva queda apagada. Cuando se encienda tendrá la ip vieja, causando temporariamente un conflicto. Entonces hay que encender la máquina, realizar algunos cambios y reiniciarla.

```
gnt-instance start nombre-nueva-maquina  
gnt-instance console nombre-nueva-maquina
```

Ahora se ingresa como usuario root y se realizan cambios en 3 archivos. El primero por la ip, el segundo por el nombre del host y el tercero por la finalidad del host.

```
vi /etc/network/interfaces  
vi /etc/hostname  
init 6
```

- Si no se puede optar por la opción del conflicto, habrá que montar el disco de la instancia en un anfitrión y realizar los cambios. Algunas tareas se realizan en el anfitrión del nodo primario, en este caso nodo04

En el nodo master

```
gnt-instance activate-disks ftp.ihlla.unicen.edu.ar
```

que retorna algo parecido a:

```
nodo04.unicen.edu.ar:disk/0:/dev/drbd3
```

En el anfitrión del primario (nodo04)

```
kpartx -av /dev/drbd3
```

que retorna

```
add map drbd3p1 (253:24): 0 104855552 linear /dev/drbd3 2048
```

Entonces montar localmente el file system con:

```
mount /dev/mapper/drbd3p1 /mnt
```

Editar

```
vi /mnt/etc/network/interfaces  
vi /mnt/etc/hostname
```

Desmontar y retornar

```
umount /mnt  
kpartx -dv /dev/drbd3
```

- y borrar el directorio con el backup descompactado de la vm

## Manipulación de los discos

### Agregar espacio en disco

#### Alternativa 1 : Grow Disk

Es un proceso en 3 etapas:

1) Agregar 10 GB al disco 1 (en general /dev/xvdb) de la instancia:

```
gnt-instance grow-disk <instancia> 1 10G
```

2) Recrear la tabla de particiones del disco (asumiendo /dev/xvdb con 1 particion /dev/xvdb1)

```
umount /dev/xvdb1 - Se debe desmontar la particion  
fdisk /dev/xvdb - Se utiliza fdisk para manipular la tabla  
comando "d" - Elimina la particion  
comando "n" - Crea nuevamente la particion (dejarla igual, si era primaria,  
id 1)  
comando "w" - Escribe y sale
```

3) Redimensionar el filesystem

```
fsck -f /dev/xvdb1 - Primero se debe chequear la particion  
resize2fs /dev/xvdb1 - Redimensiona la particion  
mount /dev/xvdb1 - Se vuelve a montar
```

**Nota:** No hay alternativas para achicar el disco.

#### Alternativa 2 : Export/Import

Una alternativa para agrandar el disco exportar e importar indicando el nuevo tamaño:

```
gnt-instance stop instance
gnt-backup export -n nodoXX instance
gnt-instance remove base.ihlla
gnt-backup import --disk 0:size=24G -n nodoPP:nodoSS --src-node=nodoXX -t
drbd instance
```

## Preparar el disco para acceso

```
# gnt-instance activate-disks mailer
eln1.rec.unicen.edu.ar:disk/0:/dev/drbd7
# kpartx -av /dev/drbd7
add map drbd0p1 (253:15): 0 20964762 linear /dev/drbd7 63
```

Con fdisk agregar una particion. Luego

```
# kpartx -d /dev/drbd7
# kpartx -av /dev/drbd7
```

Ahora aparecen las 2. Formatear y usuario como usualmente.

## Activación y montado de un disco replicado

Para activar un disco replicado, es decir, dejarlo disponible para un montado

```
gnt-instance activate-disks elm1.rec.unicen.edu.ar
kpartx -av /dev/drbd0
mount /dev/mapper/drbd0p1 /mnt
cd /mnt
cd /
umount /mnt
kpartx -d /dev/drbd0
gnt-instance deactivate-disks elm1.rec.unicen.edu.ar
```

En lo anterior, fallo la desactivación del disco. Y al ver la informacion del estado de la instancia lo daba como degradado (porque modifique en uno solo). Al levantar la instancia, lo arreglo, y se pudo entrar perfectamente a la consola. Igual, por lo que he visto, conviene cambiar el tipo de almacenamiento a plain, operar y luego volverlo a replicado.

## Borrado manual de un disco

Obtener los datos del disco con

```
lvdisplay
```

Eliminarlo con:

```
lvremove /dev/xenvg/xxxxxxx
```

## Agregar y quitar particiones a un LVM

Sumar:

```
vgextend my_volume_group /dev/sda7
```

Restar:

```
vgreduce my_volume_group /dev/sda7
```

La cuestión es que el drive debe estar vacío (Available). Eso lo podemos verificar usando:

```
pvdisk /dev/sda7
```

En la línea de status debe figurar como disponible. Si no lo está, entonces hay que mover lo que tiene a cualquier otro lugar disponible del volumen:

```
pvmove -v /dev/sda7
```

## Esconder volumen LVM a ganeti

Si se crean nuevas particiones, gnt-cluster verify falla. Hay que indicar a ganeti que ese volumen NO está administrado por ganeti.

```
Mon Mar 31 13:18:16 2014 * Verifying orphan volumes
Mon Mar 31 13:18:16 2014 - ERROR: node eln5.rec.unicen.edu.ar: volume
bkpvg/ganeti_export is unknown
```

```
gnt-cluster modify --reserved-lvs='bkpvg/ganeti_export'
```

## Instancias con windows u otros sistemas sobre Full Virtualización

Por todos lados se puede leer que el kvm es mejor para full virtualización porque aprovecha las instrucciones de virtualización que provee el hardware, cosa contraria a xen. A pesar de esto es importante saber que **NO** se puede utilizar en simultáneo el hipervisor xen-pvm con el kvm.

Hay que usar xen-hvm para esto es necesario hacer:

Habilitar el uso de VNC para ver la consola:

```
touch /etc/ganeti/vnc-cluster-password
editar /etc/xen/xend-config.sxp habilite el uso de vnc (vnc-listen
'0.0.0.0') (vncpasswd '')
/etc/init.d/xen restart
```

Habilitar el hipervisor en el ganeti

```
gnt-cluster modify --enabled-hypervisors xen-pvm,xen-hvm
```

Agregar la virtual como se muestra a continuación y conectarse con un escritorio remoto al puerto del vnc (netstat -nat antes y despues de crear la virtual para detectar el puerto)

```
gnt-instance add -t plain --disk=0:size=12g -B memory=1024 -H xen-
hvm:cdrom_image_path=/home/usina/Elastix-2.5.0-BETA3-x86_64-
bin-29ene2014.iso,boot_order=dc,vnc_bind_address=0.0.0.0 -n upe522 -o
debootstrap+default elastix
```

Cuando haya sido correctamente instalada hay que cambiar el parámetro boot\_order en la configuración de la virtual

## Alta disponibilidad

Para conocer el estado de los servidores se utiliza heartbeet. Este demonio se encuentra corriendo en cada nodo, constantemente evalua el estado de cada servidor, esto lo hace utilizando paquetes udp. Cuando detecta que un nodo se cayo, le da un start a un script en solo un nodo. El mismo, tomo el rol de master, sino lo fuera y realiaa un failover automatico sobre el nodo caido, luego envia un mail para dar aviso de la situación con las instrucciones para agregar nuevamente el nodo caido.

**IMPORTANTE** Todo lo que se indica a continuación deberá ser realizado en todos los nodos del cluster, y la configuración es exactamente la misma para cada uno.

### Instalación de heartbeat

En cada nodo ejecutar:

```
aptitude install heartbeat
```

### Configuración de heartbeat

**ha.cf**

El archivo /etc/ha.d/ha.cf es el de configuración principal, a continuación un ejemplo:

```
logfile /var/log/ha-log
logfacility local0
```



```
keepalive 2
deadtime 10
warntime 5
initdead 60
udpport 694
bcast eth1
auto_failback off
node clcom1.unicen.edu.ar
node clcom2.unicen.edu.ar
```

Aquí indica que hará dos comprobaciones del nodo (**keepalive 2**). Dándole 10 seg de tiempo para responder satisfactoriamente(**deadtime 10**). El archivo de log se encuentra **/var/log/ha-log**. Luego de una caída si se detecta que “revive” el nodo muerto no va a reagregarlo al cluster (**auto\_failback off**). El cluster esta compuesto por todas las definiciones de **node** que se encuentren, en este caso solo dos. Dichas definiciones tienen que tener el mismo nombre que responde `uname -n`.

## haresources

El archivo se encuentra en `/etc/ha.d/haresources`, el siguiente es un ejemplo:

```
clcom1.unicen.edu.ar 192.168.171.100 recuperarnodocaido.sh
```

El primer campo (**clcom1.unicen.edu.ar**) es el nombre del nodo master. Aunque no importa que nodo sea en realidad, es necesario que haya alguno especificado. El segundo campo (**192.168.171.100**) es la ip virtual que esta asociada al master del cluster. y el tercero (**recuperarnodocaido.sh**) Es el script que realizará el failover automático.

## Script de recuperación

Este script debe estar ubicado en `/etc/init.d/` Debe ser modificado solo las variables declaradas: **ARCH\_LOG** y **EMAIL** que indican la ubicación del archivo de log y la dirección de mail de notificación respectivamente

[recuperarnodocaido.sh.zip](#)

## Optimización del cluster

**Nota:** Para simplicidad en la administración y la ejecución de los comandos, primero me aseguré que en todos los nodos la placa de servicio sea `eth0` y la de backplane la `eth1`. Esto lo hice utilizando el archivo `/etc/udev/rules.d/70-persistent-net.rules`

- Asegurarse que todas las placas sincronicen a 1 Gbit

```
gnt-cluster command "mii-tool eth1" // El resultado debiera ser: eth1:
negotiated **1000**baseT-FD flow-control, link ok
```

- Tema MTU: La primera recomendación para mejorar el throughput es aumentar el mtu de las

eht1, sin embargo haciendo esto el drdb se vuelve loco y no funciona nada. Supongo que es tema del switch que no los oporta, reintentaremos más adelante con un switch como la gente.

- Aumento de tamaños de buffers del SO y de las colas de las placas. Se ejecutaron los siguientes comandos:

```
gnt-cluster command "sysctl -w net.core.wmem_max=8388608"
gnt-cluster command "sysctl -w net.core.rmem_max=8388608"
gnt-cluster command "sysctl -w net.core.wmem_default=1048576"
gnt-cluster command "sysctl -w net.core.rmem_default=1048576"
gnt-cluster command "sysctl -w net.ipv4.tcp_mem=1048576\ 10485760\ 125376"
gnt-cluster command "sysctl -w net.ipv4.tcp_wmem=8192\ 262144\ 10485760"
gnt-cluster command "sysctl -w net.ipv4.tcp_rmem=8192\ 262144\ 10485760"
gnt-cluster command "sysctl -w net.core.netdev_max_backlog=10000"
gnt-cluster command "sysctl -w net.ipv4.route.flush=1"
gnt-cluster command "ifconfig eth1 txqueuelen 10000"
```

- Para dejar los cambios persistentes. Editar el archivo /etc/sysctl.conf en el nodo master

```
net.core.wmem_max = 8388608
net.core.rmem_max = 8388608
net.core.wmem_default = 1048576
net.core.rmem_default = 1048576
net.ipv4.tcp_mem = 1048576 10485760 125376
net.ipv4.tcp_wmem = 8192 262144 10485760
net.ipv4.tcp_rmem = 8192 262144 10485760
net.core.netdev_max_backlog = 10000
```

- Impactar cambios en el cluster

```
gnt-cluster copyfile /etc/sysctl.conf
gnt-cluster command "sysctl -p"
```

- Dejar permanente txqueuelen: Editar /etc/network/interfaces en cada nodo dejando la siguiente sintaxis:

```
auto eth1
iface eth1 inet static
    address 10.253.2.21
    netmask 255.255.255.0
up ifconfig eth1 txqueuelen 10000
```

## Resolución de problemas

### Freezar fecha (pilaga)

1. En crontab comentar sincronización de la hora
2. En crontab agregar \*/30 \* \* \* \* /bin/date -s "12/30/20XX \$(/bin/date +%R)" > /tmp/freeze-hora.log 2>&1
3. En /etc/rc.local agregar: /bin/date -s "12/30/20XX \$(date +%R)" > /tmp/freeze-hora.log 2>&1

## Volumenes huérfanos

Uno de los problemas que “ponen en rojo” la verificación del cluster es cuando quedan volúmenes huérfanos. Hay 2 casos posibles:

- El volumen queda por un error en la migración de disco secundario. En este caso se debe ir al nodo que presenta el problema y hacer

```
lvremove xenvg - Esto intenta eliminar todo el volume group, pero como hay virtuales andando, solo eliminará el volumen huérfano
```

- El volumen queda por un error en la generación de la imagen semanal, el problema es que en este caso el volumen queda marcado como “en uso” y por lo tanto se requiere un paso previo:

```
dmsetup ls - Muestra los volúmenes, detectar los que terminen en ".snap-1"  
dmsetup remove <volumen detectado>  
lvremove xenvg
```

## Se cuelga el drbd al migrar una virtual

Si al migrar una virtual queda esperando la sincronización, puede suceder que haciendo `/etc/init.d/drbd status`, diga stalled y el volumen esté a un 100% pero aun así le falten algunos KB por copiar.

La solución en ese caso es ir al que quedó como master y en estado consistente, y hacer

```
drbdsetup N disconnect
```

donde N es el menor del lado del primario

## Se rompe el secundario de un disco

Por ejemplo porque se rompe el disco de un nodo. La mejor alternativa es pasarla a tipo plain para luego restablecer el espejo en otro nodo.

```
gnt-instance stop www  
gnt-instance modify -t plain www  
gnt-instance start www
```

## Se rompe el primario de un disco

```
gnt-instance migrate -f www  
gnt-instance stop www  
gnt-instance modify -t
```

## Actualizacion de 2.4.1 a 2.5.2

- No tiene que haber corriendo jobs. Se verifica con “gnt-job list”
- Detener ganeti en todos los nodos
- Realizar un backup de la configuración actual. En esta versión el directorio es “/var/lib/ganeti”. Cuando terminemos de actualizar esto va a cambiar de lugar!!
- Descompactar “ganeti-2.5.2.tar.gz” en el directorio “/opt”.
- En el directorio con la nueva versión, construir desde cero de manera usual “./configure; make; make install”
- Para actualizar la configuración hay que usar una aplicación que viene con ganeti. Primero se corre con parámetro “-dry-run” para que no grabe nada. Solo simula. Si esto no da error entonces podemos hacer la corrida definitiva

```
/usr/local/lib/ganeti/tools/cfgupgrade --verbose --dry-run --  
path=/var/lib/ganeti  
/usr/local/lib/ganeti/tools/cfgupgrade --verbose --path=/var/lib/ganeti
```

- Ahora hay que realizar dos cambios no previstos/documentados

```
mv /var/lib/ganeti /usr/local/var/lib  
ln -s /var/lock /usr/local/var/
```

- Levantar ganeti en todos los nodos (ultimo en el master). Nada debería fallar.
- Redistribuir la configuracion

```
gnt-cluster redist-conf
```

- Reiniciar ganeti en todos los nodos.
- Correr verificación en el cluster.

### Problemas detectados

Las maquinas que estaban corriendo no se vieron afectadas por la actualizacion. Pero la primer migración que hice se rompio, aunque antes iba bien. Todas las siguientes que hice después de reiniciar la VM fueron perfecto. La conclusión es que hasta que las maquinas virtuales no se reinician en el nuevo cluster, algunas cosas no están del todo bien.

### Otras verificaciones

Creacion de maquinas desde cero, rebalanceo del cluster, apagado, prendido, upgrade de la cantidad de memoria, procesadores, migraciones (después del reinicio de la VM) no fallo nada.

## Actualización de 2.5.2 a 2.9.2

De los repositorios de squeeze instalar los paquetes:

```
apt-get install python-bitarray python-ipaddr python-yaml r-cran-design
```

Luego colocar temporariamente los repositorios de debian 7, solo para instalar

```
deb http://debian.unicen.edu.ar:9999/debian wheezy main contrib non-free
deb http://debian.unicen.edu.ar:9999/security wheezy/updates main contrib
non-free
```

y redistribuir la configuración

```
gnt-cluster copyfile /etc/apt/sources.list
gnt-cluster command "apt-get update"
gnt-cluster command "apt-get -y install python-sphinx libghc6-hslogger-dev
libghc-utf8-string-dev libghc-parallel-dev libghc-curl-dev"
```

Finalmente, comentar los repos de debian 7 y redistribuir configuración. A continuación:

- Descompactar "ganeti-2.9.2.tar.gz" en el directorio "/opt".
- En el directorio con la nueva versión, construir desde cero de manera usual "./configure; make;"

## Estado de VMs en el cluster

La pagina se visualiza a través de la wiki. Tiene que tener instalada la extensión phpwikifi, y disponibles los includes de jquery. Una cuestión algo complicada, es encontrar exactamente cual es la OID a través de la cual está disponible la información que se busca. Lo que me resultó mas comodo hacer es buscarla haciendo

```
snmpwalk -v 1 -c public ip_host .1.3.6.1.4.1.8072
```

ya que es a partir de esa OID desde donde están disponibles las extesiones a SNMP

La página se crea con el siguiente bloque de codigo php.

```
~~NOCACHE~~
==== Configuración de servidores de rectorado en cluster ====
Esta información se obtiene on-line. Muestra exactamente la configuración
actual. En caso de que el host esté no disponible, solamente se mostrarán
las 4 primeras columnas.
<HTML>
    <script src="lib/tpl/default/jquery.min.js"></script>
    <script src="lib/tpl/default/jquery.peity.js"></script>
</HTML>

<phpwikify>
    $x = snmpwalk('eln1.rec.unicen.edu.ar', 'public',
'.1.3.6.1.4.1.8072.1.3.2.4.1.2.12.115.110.109.112.95.71.78.84.76.73.83.84');
    foreach ($x as $val) {
        $y = preg_replace('/^...../', '', $val);
        echo "$y\n";
    }
```

```

}
</phpwikify>

<HTML>
  <script type="text/javascript">
    // El orden de los colores es el inverso a la notación. Si se desea
    graficar 3/10, el primer color corresponde al 10 (el total del gráfico) y el
    segundo al 3 (lo ocupado)
    $.fn.peity.defaults.pie = {
                                colours: ['#D0D0D0', '#006666'],
    delimiter: '/',           diameter: 20
                                };
    // El orden de los colores es el inverso a la notación. Si se desea
    graficar 3/10, el primer color corresponde al 10 (el total del gráfico) y el
    segundo al 3 (lo ocupado)
    $.fn.peity.defaults.pie2 = {
                                colours: ['#D0D0D0', '#FF0000'],
    delimiter: '/',           diameter: 20
                                };

    // Invocación estandar.
    $("span.pie").peity("pie")
    $("span.pie2").peity("pie2")
  </script>
</HTML>

```

El monitoreo se realiza a través de snmp. Debe permitirse la consulta a las oid en forma publica para lectura. Al menos debe incluirse la siguiente configuracion en /etc/snmp/snmpd.conf:

```

view all      included   .1                                80
view system included   .iso.org.dod.internet.mgmt.mib-2.system
# OID: .1.3.6.1.4.1.2021.8.1.101.1
exec snmp_GNTLIST /bin/cat /etc/snmp/snmp_GNTLIST.sh.cache

```

El archivo que se muestra snmp\_GNTLIST.sh.cache se regenera cada 5 minutos desde el cron, con la siguiente linea en la crontab:

```

2,7,12,17,22,27,32,37,42,47,52,57 * * * * /bin/sh /etc/snmp/cron.sh
>/tmp/cron.log 2>&1

```

el script "/etc/snmp/cron.sh" contiene:

```

#!/bin/sh
cd /etc/snmp
a=`ls snmp_*.sh`
for b in $a
do
  /bin/bash $b > $b.cache
done

```

El script "/etc/snmp/snmp\_GNTLIST.sh" es el que itera sobre la lista de las maquinas virtuales en el

cluster, haciendo a su vez una consulta snmp a cada una de ellas y contiene:

```
#!/bin/sh
echo "^Host ^Nodo ^Estado ^ Disco ^Tipo ^IP ^Version Debian ^#Proc
^Mem(KB) ^Disk(GB) ^Servicio ^Resp. ^Monitor ^URLs de servicio ^"
/usr/local/sbin/gnt-instance list --no-headers -o
name,pnode,disk_template,status | sed -e
's/(e[0-9])\.\.rec\.unicen\.edu\.ar/1/g' | while true
do
    read m p d s
    if [ -z "$m" ]
    then
        exit 0
    fi
    # Siempre se coloca el nombre de la maquina, disk_template y estado
    if [ "$d" = "plain" ]
    then
        dd="<color black/yellow>plain</color>"
    else
        dd="$d"
    fi
    /bin/echo -n "|$m|$p|$s|$dd| "
    # Si se consigue algo vía snmp de la misma, se agrega.
    snmpwalk -v 1 -r 1 -O n -c public $m
.1.3.6.1.4.1.2021.9.4.1.2.8.104.111.115.116.105.110.102.111 2>/dev/null |
sed -e 's/^.* = STRING: //' -e 's/"//g' | awk '{printf "%s |", $0}' | sed -
e 's/End of MIB//' 2>/dev/null
    /bin/echo
done
```

En cada una de las máquinas virtuales también hace falta que esté instalado snmp, y configurado en snmpd.conf con:

```
extend .1.3.6.1.4.1.2021.9 hostinfo /bin/cat
/etc/snmp/snmp_HOSTINFO.sh.cache
```

Misma consideración que para el snmp del nodo master del cluster. Tienen que tener en el cron la ejecución del cron.sh. El script snmp\_HOSTINFO.sh toma información de dos fuentes. Una es la configuración del propio equipo (memoria, disco, etc.) y la otra es del contenido del archivo /etc/hostinfo cuyo contenido se detalla más adelante. El código es:

```
#!/bin/sh

# Se leen todas las líneas necesarias del /etc/hostinfo
for x in 1 2 3 4 5
do
    read y
    case $x in
        1) DESC="$y";;
        2) RESP="$y";;
        3) MONI="$y";;
```

```
4) URLS="$y";;
5) TIPO="$y";;
esac
done < /etc/hostinfo

# En primer lugar va el tipo Prod/Desa, coloreado solo para Prod
if [ "$TIPO" = "Prod" ]
then
    echo "<color white/green>$y</color>"
else
    echo "$TIPO"
fi

# Direccion ip
/sbin/ifconfig eth0|grep "inet addr" | awk '{print $2}'|sed -e 's/^addr:/'
# Version de Debian
/bin/cat /etc/debian_version
# Cantidad de procesadores
cat /proc/cpuinfo |grep "^processor" | wc -l
# Memoria ocupado/total. >= 75 en rojo
MEMC=$(/usr/bin/free | grep "^Mem" | awk '{printf "%2d", 100*($3-$7)/$2}')
MEM=$(/usr/bin/free | grep "^Mem" | awk '{printf "%2d/%2d\n", ($3-$7)/1024, $2/1024}')
if [ $MEMC -gt 74 ]
then
    GR="pie2"
else
    GR="pie"
fi
echo "<html><span class=\"$GR\">$MEM</span><br>${MEM}</html>"
# Disco ocupado/total
DISC=$(/bin/df -B M|sed -e 's/M//g'|grep "\/$"|awk '{printf "%2d", 100*$3/$2}')
DIS=$(/bin/df -B M |sed -e 's/M//g'| grep "\/$"|awk '{print $3 "/" $2}'|sed -e 's/G//g')
DISD=$(/bin/df -B M |sed -e 's/M//g'| grep "\/$"|awk '{printf "%.1f/%.1f", $3/1024, $2/1024}'|sed -e 's/G//g')
if [ $DISC -gt 74 ]
then
    GR="pie2"
else
    GR="pie"
fi
echo "<html><span class=\"$GR\">$DIS</span><br>${DISD}</html>"
# Descripcion del host
echo $DESC
# Responsable
echo $RESP
# Monitoreos
echo $MONI
# Urls de servicio
```



## echo \$URLS

El archivo /etc/hostinfo contiene información personalizada. Cada renglón contiene una información particular:

1. Descripción de la finalidad o servicio del host
2. Persona responsable del host
3. Puertos que se deben monitorear vía nagios
4. Url/Urls a través de las que brinda servicios
5. Cadena "Prod" o "Desa" dependiendo de que sea host de producción o de desarrollo.

Un ejemplo de tal archivo es:

```
Servidor de Moodle para Unicen
Sergio
tcp/22,80
http://moodle.rec.unicen.edu.ar
Prod
```

## Migración a ganeti 2.10 sobre debian 7.4

No se puede actualizar ganeti "paso a paso". Está documentado que tiene que estar en la misma versión en **todos** los nodos. El procedimiento que funcionó ok implica:

1. Actualizar el sistema operativo nodo a nodo, recompilando y reinstalando la versión actual. Requiere reinicio del anfitrión
2. Actualizar ganeti desde repositorio en todos los nodos a la vez. No requiere reinicio.

### Verificar en todos los nodos

1. Instalado y operativo ntp y ntpdate.
2. En /etc/default/grub: GRUB\_CMDLINE\_XEN\_DEFAULT="dom0\_mem=1024M" y GRUB\_DISABLE\_OS\_PROBER=true
3. En /etc/default/xendomains: XENDOMAINS\_SAVE=
4. Reemplazar /etc/xen/xend-config por:

```
(xend-http-server yes)
(xend-relocation-server yes)
(xend-relocation-port 8002)
(xend-address '')
(xend-relocation-address '')
(xend-relocation-hosts-allow '')
(vif-script vif-bridge)
(dom0-min-mem 256)
(enable-dom0-ballooning no)
(total_available_memory 0)
(dom0-cpus 0)
(vncpasswd '')
```

Tiene que estar booteando Linux 2.6.32-5-xen-amd64 and XEN 4.0-amd64

Verificar que se encuentren instalados:

- drbd8-utils lvm2 pandoc python-sphinx python-pygraphviz pylint pep8 socat ghc libghc6-json-dev libghc6-network-dev libghc6-quickcheck-dev python-openssl python-simplejson python-pyparsing python-pyinotify python-pycurl python-paramiko dump iputils-arping kpartx libsysfs2 libghc6-parallel-dev libghc6-curl-dev ganeti-instance-debootstrap

Que /etc/modules contenga:

- drbd minor\_count=128 usermode\_helper=/bin/true

Que /etc/drbd.conf esté vacío, y que los valores de /etc/default/ganeti-instance-debootstrap sean correctos

## Compilación de ganeti desde fuentes

```
make clean
./configure --localstatedir=/var --sysconfdir=/etc
make
make install
cp doc/examples/ganeti.initd /etc/init.d/ganeti
mkdir /etc/ganeti
mkdir /srv/ganeti
mkdir /srv/ganeti/os
mkdir /srv/ganeti/export
chmod +x /etc/init.d/ganeti
update-rc.d ganeti defaults 20 80
cd /usr/local/src
wget
http://ganeti.googlecode.com/files/ganeti-instance-debootstrap-0.14.tar.gz
tar xzf ganeti-instance-debootstrap-0.14.tar.gz
cd ganeti-instance-debootstrap-0.14/
./configure --with-os-dir=/srv/ganeti/os
make
make install
apt-get install debootstrap dump kpartx
apt-get install ganeti-instance-debootstrap
```

## Crear cluster de prueba

1. Instalación de Debian 6 y compilación y configuración de ganeti 2.5.2
2. Creación de dos maquinas virtuales. Backup de las virtuales.

## Hacer backup de la configuración actual del cluster

1. Salvar /var/lib/ganeti. Por ejemplo: tar czf /var/lib/ganeti-\$(date +%FT%T).tar.gz -C /var/lib

ganeti

## Actualización de cada nodo a Debian 7

1. Failover del master si el nodo a actualizar es el master
2. Failover para deslojar completo
3. actualización de d6 a d7 en nodo (preservar archivos de configuración cambiados). Reiniciar. Limpiar paquetes no usados.
4. Sacar xen
5. Colocar xen 4.1: apt-get install xen-linux-system-amd64 xen-tools xen-utils-4.1 xen-hypervisor-4.1-amd64
6. Cambiar boot a xen 4.1 (/etc/default/grub). Revisar /etc/default/grub, /etc/grub.d y config XEN
7. Hace falta recompilar y reinstalar ganeti. En /opt/ganeti-2.5.2/htools pisar los archivos que adjunto en htools.tar. Fueron modificados manualmente para adaptarlos a la nueva versión de ghc. Si aun así no compila, agregar al configure el flag -disable-htools.
8. Si al iniciar da error indicando que /var/lib/ganeti no es un directorio, reemplazar el symlink por lo que apunta.
9. Editar SI HACE FALTA /etc/xen/scripts/vif-bridge. Cambiar invocación a **setup\_bridge\_port** por **setup\_virtual\_bridge\_port**. Hecho esto, cuando se ejecuta el gnt-cluster verify da advertencia porque el script vif-bridge es distinto en los dos nodos.
10. Activar los discos de las maquinas que estaban en el nodo desalojado para recuperar drbd.
11. Verificar que el cluster está ok.

## Paso final

1. gnt-cluster copyfile /etc/apt/sources.list
2. Si no quedó igual en todos lados: gnt-cluster copyfile /etc/xen/scripts/vif-bridge

## Actualización de ganeti

### En todos los nodos

1. Agregar "deb <http://debian.unicen.edu.ar:9999/debian> wheezy-backports main contrib non-free" a /etc/apt/sources.list
2. Actualizar todo el software.
3. Apagar ganeti
4. tar -cvf /root/ganeticonf.tar /var/lib/ganeti
5. cd /opt/ganeti-2.5.2
6. make uninstall
7. apt-get -t wheezy-backports install ganeti2 (pisar /etc/init.d/ganeti)
8. Pasos detallados en: <http://docs.ganeti.org/ganeti/2.5/html/upgrade.html>
9. hash -r (porque cambia el path de los ejecutables de ganeti)

### En el nodo master

1. /usr/lib/ganeti/tools/cfgupgrade -verbose -dry-run

2. `/usr/lib/ganeti/tools/cfgupgrade -verbose`
3. `gnt-cluster redist-conf`
4. Hecho esto, reconoce a todas las VMs corriendo sin problema. Da consola, etc.

#### En todos los nodos

1. `cd /usr/share/ganeti`
2. `mv os os-old`

#### Particularidades del cluster en rectorado

En todos los nodos, ya se trató de instalar ganeti-2.9. Por este motivo hay muchos paquetes instalados desde repositorios de wheezy. Es necesario volver atrás a paquetes de squeeze stable. Con las siguientes acciones:

- `apt-get remove ghc python-sphinx libcurl4-openssl-dev libghc6-quickcheck1-dev libcurl3-gnutls libgnutls26 libc6-dev`
- `apt-get clean; apt-get autoremove`
- `aptitude install binutils drbd8-utils lvm2 pandoc python-sphinx python-pygraphviz pylint pep8 socat ghc6 libghc6-json-dev libghc6-network-dev python-openssl python-simplejson python-pyparsing python-pyinotify python-pycurl python-paramiko dump iputils-arping kpartx libsysfs2 libghc6-parallel-dev libghc6-curl-dev libcurl4-gnutls-dev libgnutls-dev libghc6-quickcheck2-dev`
- Tomar en aptitude la segunda oferta, que hace un downgrade de paquetes.

## Instalacion sobre RAID 0

Habilitar RAID:

1. Entrar en la BIOS
2. On-Chip ATA Device
3. RAID mode: RAID (para ponerlos en raid)
4. Reiniciar

Control f mientras esta iniciando

1. 1
2. Raid 0
3. Stripe block 64kb - tamaño de bloque de archivo
4. Definir LD + enter
5. assign y y
6. ctrl y - guardar

Instalación de S.O (Debian 7.4)

1. Opción instalar + tab + `dmraid=true`
2. Seleccionar idioma, país, mapa de teclado, placa de red
3. Definir nombres, claves y dominio de nodo

## Seleccionar método de partición de disco: Manual

1. Crear tabla de partición
2. Partición nueva - 40 gb - primaria - principio - fichero : ext4 - punto de montaje /
3. Área de intercambio (swap) - 2 gb - primario - principio
4. Resto - crear partición - no utilizar-
5. Finalizar la partición
6. Escribir cambios en el disco
7. Repo de debian - introducir manualmente : debian.unicen.edu.ar:9999
8. Directorio de la replica de debian: /debian/
9. Sin proxy
10. Seleccionar utilidades estándar del sistema
11. Instalar cargador de arranque en /dev/mapper/nombre de disco (saber nombre de disco)
12. cntrl + alt + f2
13. ls /dev/mapper (ver nombre de dispositivo)
14. cntrl + alt + f1

## Reiniciar con cd de ubuntu live

1. Seleccionar idioma - Probar sin instalar
2. Entrar como superusuario sudo su
3. Montar la partición 1 - mount /dev/mapper/nombrededisco1 /mnt/
4. Montar archivos proc mount -bind /proc /mnt/proc -
5. Montar archivos sys mount -bind /sys /mnt/sys -
6. Montar archivos dev mount -bind /dev /mnt/dev -
7. Entrar como root - chroot /mnt/
8. Eliminar linea - force\_load dm-raid45 - de /usr/share/initramfs-tools/hooks/dmraid
9. Mover los archivos de inicio de raid a linux para el arranque- mv initrd.img-3.2.0-4-amd64 initrd.img-3.2.0-4-amd64.bkp - en /boot/
10. Actualizar archivo de arranque de raid - update-initramfs -c -k all -
11. Rebootear el dispositivo exit (para salir de chroot)- init 6 -

## Grub, presionar "e"

1. Editar nombre de partición de disco - borrar p -
2. F10 para guardar
3. Luego realizamos lo mismo en /boot/grub/grub.cfg para dejarlo definitivo. En la parte final del archivo, a partir de la linea

```
### BEGIN /etc/grub.d/10_linux ###
```

## Recuperar un disco en raid1 (en caso de fallar 1)

1. Volvemos a editar el nombre de la partición. En grub presionar "e" y editamos linux  
/boot/vmlinuz-2.6.32-23-generic root=**/dev/sdb1**

Continuar con la instalacion normal de Xen y Ganeti

## Utilización snf-image para crear maquinas virtuales

el snf-image es un software que crea virtuales al igual que debootstrap+default. Esta aplicación tiene la capacidad de levantar un RAW generado de otro disco y transformarlo en una máquina virtual.

## Ampliación de disco de LDAP.REC.UNICEN.EDU.AR

En eln1.rec.unicen.edu.ar:

```
clusterinfo | grep mailer
```

Obtendremos una linea del estilo:

```
mailer                    eln8 eln6 8 4.0G 100.1G running /usr/lib/xen-4.1/bin/pygrub
```

Esto muestra que el nodo primario es el 8 y el secundario el 6. Esta información es relevante para poder recrearla en las mismas condiciones. A continuación la apagamos y disparamos el backup.

```
gnt-instance stop mailer  
/etc/backup/backupuna.sh mailer.rec.unicen.edu.ar 4
```

En eln5.rec.unicen.edu.ar:

```
cd /var/lib/ganeti/export  
tar -xvf mailer.rec.unicen.edu.ar.tgz  
cd mailer.rec.unicen.edu.ar
```

Allí editar el archivo config.ini. Bajo el tag "[Instance]" se encuentra una linea con la indicación del tamaño:

```
disk0_size = 102400
```

Editarlo para colocar el nuevo tamaño y salvar.

En eln1.rec.unicen.edu.ar, borramos la VM original, e importamos el backup

```
gnt-instance remove --force mailer  
gnt-backup import -n eln8:eln6 --src-node=eln5 -t drbd  
mailer.rec.unicen.edu.ar  
gnt-instance start mailer
```

Y nuevamente en eln1.rec.unicen.edu.ar, borrar el directorio donde descompactamos el último backup

```
cd /var/lib/ganeti/export  
rm -rf mailer.rec.unicen.edu.ar
```

En cuanto a los tiempos, la exportación insume aproximadamente 4 minutos por cada 10GB y en la importación el doble. Así para esta intervención en la que se pasa de 100GB → 200GB deberíamos asumir una tarea de mas de 2 horas, si todo sale perfecto.

From:  
<https://wiki.rec.unicen.edu.ar/wiki/> - **Wiki UNICEN**

Permanent link:  
<https://wiki.rec.unicen.edu.ar/wiki/doku.php?id=informatica:ganeti&rev=1445548482>

Last update: **2017/10/10 16:08**

