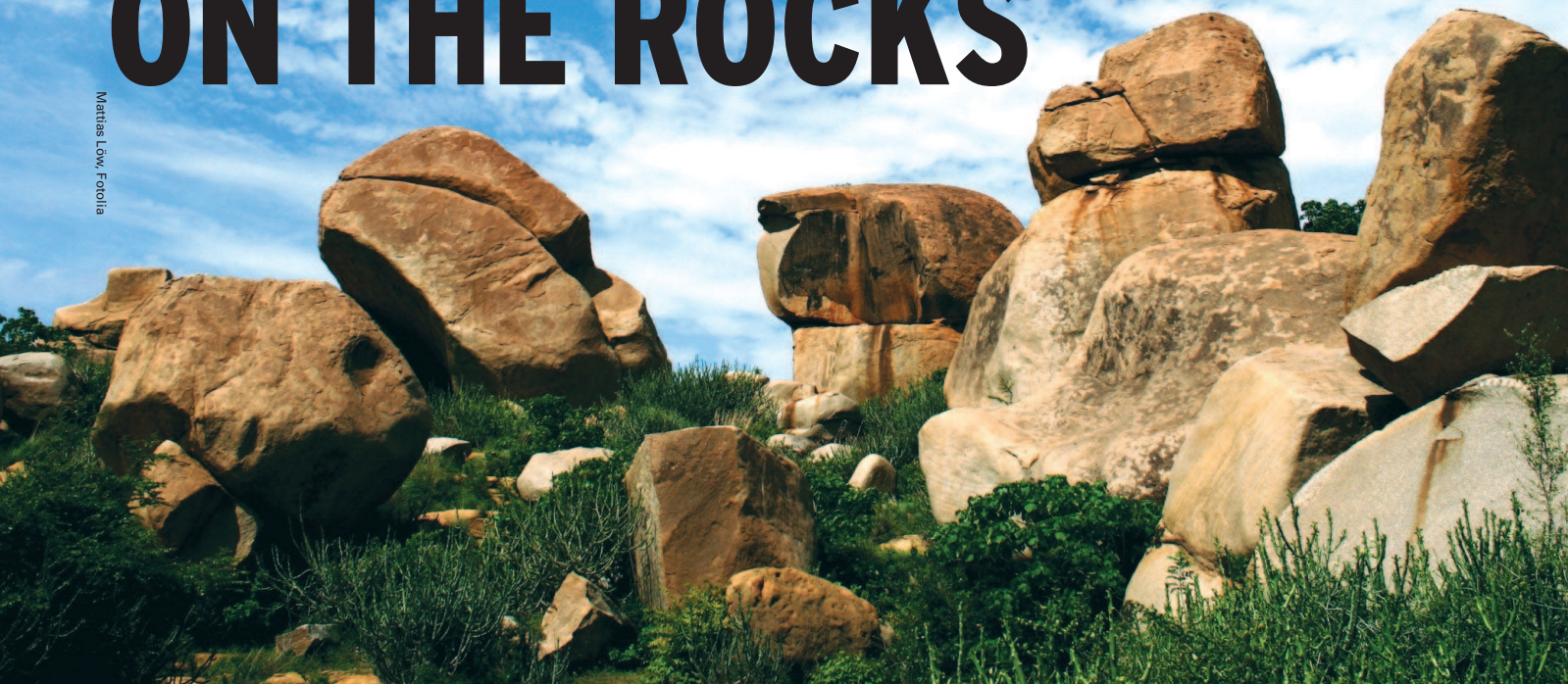


Creación de clusters virtuales con Rocks

# ON THE ROCKS

Martina Low, Fotolia



Rocks ofrece un sencilla solución de clústering con máquinas virtuales.

**POR MATTHEW SACKS**

**R**ocks es una distribución de Linux y un sistema de gestión de clusters que permite el despliegue rápido de clusters Linux sobre hardware físico o contenedores virtuales Xen. Un clúster de Rocks [1] es fácil de implementar y ofrece todas las ventajas de la virtualización de los nodos del clúster. Con un mínimo de dos máquinas físicas, Rocks permite un despliegue y una gestión de clusters simple y rápida, liberando al administrador para que se centre en el apoyo a la red informática y a las aplicaciones distribuidas que hacen del clúster una opción atractiva.

Incluidas en la distribución estándar de Rocks existen diferentes herramientas de código abierto de computación paralela y distribuida de alto rendimiento, tales como Sun Grid Engine [2], OpenMPI [3] y Condor. Esta poderosa colección de funciones avanzadas es una de las razones por las que la NASA, la NSA, el Laboratorio de Investigación de IBM en Austin, la Marina de los EE.UU., el MIT, Harvard y la Universidad Johns Hopkins están todos utilizando Rocks por alguna de sus aplicaciones más intensivas.

## ¿Por Qué Virtualizar un Clúster?

Las razones para el despliegue de clusters virtuales son los mismos que justifican cualquier solución de virtualización: flexibilidad, facilidad de gestión y la utilización eficiente de los recursos hardware. Por ejemplo, en un entorno en el que se deben ejecutar al mismo tiempo sistemas operativos de 32 bits y de 64 bits, la virtualización es una solución mucho más eficiente que intentar dar soporte a dos plataformas de hardware distintas en un único clúster.

## Tareas de Pre-Instalación

Antes de la instalación del clúster, asegúrese de que están disponibles todos los componentes necesarios. Un clúster de Rocks puede ser configurado de muchas formas diferentes y con distintas configuraciones de red. Rocks puede instalarse dentro de contenedores virtuales (contenedores VM) o directamente sobre hardware físico. El ejemplo proporcionado en este artículo supone que tenemos al menos dos máquinas físicas para desplegar un nodo front-end y al menos un con-

tenedor VM. El nodo front-end necesita un mínimo de 1Gb de RAM, mientras que el contenedor VM debe tener al menos 4GB (Rocks necesita un mínimo de 1 GB).

Es primordial asegurarse de que el hardware esté soportado por la distribución Rocks. El sistema operativo Rocks se basa en CentOS, de modo que es imprescindible cerciorarse de que el hardware cumple con la lista de compatibilidad del hardware CentOS/Red Hat. Como norma general, se aconseja utilizar hardware soportado ampliamente, en especial al seleccionar adaptadores de red y de gráficos.

La configuración de red básica de Rocks supone la presencia de una red pública y una red privada para el contenedor VM y su nodos. El nodo frontal debe tener dos tarjetas de red, y los nodos de computación requieren por lo menos una tarjeta para conectarse a la red privada de los nodos de computación. Para conectar los distintos contenedores VM al nodo frontal se precisa además un switch. Véase la Figura 1 para un ejemplo de configuración de red en Rocks.

## Preparación de la Instalación

Inserte el DVD de Rocks (o el CD de inicio) y arranque el sistema desde la uni-

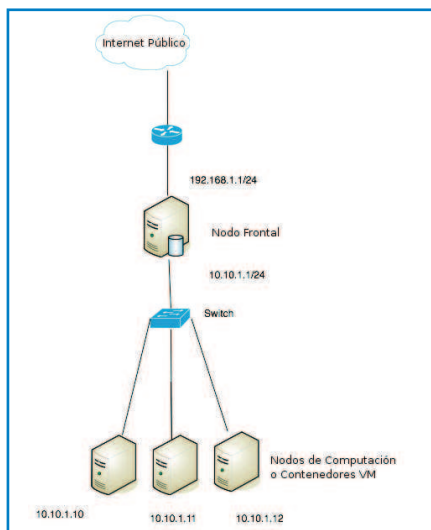


Figura 1: Un ejemplo de configuración de Rocks.

dad de CD/DVD. Si utiliza CDs, inserte primero el CD Kernel/Boot de Rocks. Rocks sugerirá varios *rolls*. En su jerga, un *roll* es una colección de programas informáticos destinados a una tarea específica. Una configuración básica requiere el *roll* Kernel/Boot, el *roll* Base, el *roll* del servidor Web y los *rolls* del SO 1 y 2, así como el *roll* Xen para el soporte de virtualización del clúster.

La configuración base no es muy emocionante, pero se puede investigar entre los diferentes *rolls* que están disponibles [4] e incluir tantos *rolls* de computación distribuida en red y Grid como se desee para divertirse con Rocks. Sun Grid Engine (SGE), Torque, y el *roll* de computación de alto rendimiento (HPC) son buenos puntos de partida para aprovechar un clúster de Rocks al máximo.

Una pantalla de inicio solicitará el modo de arranque. Para iniciar la instalación del frontal teclee *frontend* y pulse Enter. Si no hace esto tras unos segundos, el instalador de Rocks iniciará la instalación de un nodo de *computación*. Si esto ocurre, reinicie el sistema y teclee *frontend* en el cursor antes de que se inicie automáticamente de nuevo.

Una vez que haya arrancado el CD de instalación de Rocks, intentará conectar con un servidor DHCP, pero si no encuentra ningún servidor DHCP en las dos interfaces de red, solicitará la configuración de la red. Lo más probable es que *eth0* lo consiga, pero *eth1* (la red privada del clúster) no tendrá un servidor DHCP. En este caso, o tiene un servidor DHCP en la red privada, o seleccione

*manual configuration* e introduzca manualmente la dirección IPv4, la puerta de enlace y el nombre del servidor. Una vez que se establece la conectividad de red, seleccione OK para continuar con la instalación del nodo frontal.

Aparecerá una pantalla que dice “Welcome to Rocks” para indicar que puede lanzar la instalación desde el DVD, los CDs o la red. La solución más simple es descargar el DVD de antemano e instalar desde el DVD, ya que contiene la mayoría de los *rolls* o paquetes de software que se ofrecen en la web de Rocks.

Con el DVD de instalación de Rocks en la unidad, pulse *CD/DVD Based Roll*, y seleccione entonces los *rolls* que desea instalar. Un sistema base de Rocks consiste en el kernel, el sistema operativo, un servidor web y los *rolls* de base. Para configurar un clúster virtual también se necesita el *roll* de Xen (Figura 2).

Ahora seleccione los *rolls* recomendados y pulse *Submit*. Éstos aparecerán a la izquierda de la pantalla de instalación. Pulsando *Next* comienza la instalación.

Si se introduce la información del clúster proporcionará su identificación si está registrado en [rocksclusters.org](http://rocksclusters.org). Diversas indicaciones solicitarán la información de configuración, tales como la configuración de red para *eth0* y *eth1*, la contraseña de root, la zona horaria y el esquema de particiones.

## Ejecución de la Instalación.

El instalador arrancará y podremos ver un cuadro de mensaje que dice que el guión de pre-instalación se está ejecutando. A continuación el instalador formatea el sistema de archivos, apareciendo un instalador gráfico muy similar a la instalación gráfica de Red Hat Enterprise Linux con imágenes de las pirámides y el Tío Sam incitando a los entusiastas de los clusters a registrar los suyos en nombre de la National Science Foundation (Figura 3).

**Nota:** Si ha seleccionado una instalación de red, los *rolls* se descarga-

rán de Internet antes de la instalación del producto. En un momento dado durante el proceso de instalación, los medios de la instalación deberían expulsarse automáticamente y deberíamos retirarlos, de modo que vigile cuidadosamente el proceso de instalación, porque a veces, por no estar vigilados, los medios de instalación no se retiran y el ordenador entra en otro ciclo de instalación.

Una vez que la instalación se ha completado, el sistema se reiniciará automáticamente y aparecerá una ventana azul de inicio de sesión de CentOS 5.

Ahora que se ha instalado el controlador frontal, puede comenzar la instalación de varios contenedores VM y nodos de computación.

## Instalación del Contenedor VM

Los nodos de computación realizan todo el trabajo y lo sirven como sistemas individuales dentro del “supercomputador” que estamos construyendo; podemos configurar nodos de computación en cualquier número de máquinas físicas individuales. Un clúster de computación puede tener 500 servidores en un enorme centro de datos, o podemos tener dos máquinas colocadas bajo la mesa. Para aprovechar los beneficios de la virtualización en un clúster se utilizará un contenedor VM para desplegar nodos. Desplegar nodos de computación en un contenedor VM Xen permitirá un despliegue rápido y la gestión de los nodos de computación.

Para iniciar el despliegue de nodos de computación del clúster, inicie sesión en el frontal y abra una terminal. La primera vez que se abre una ventana de terminal,

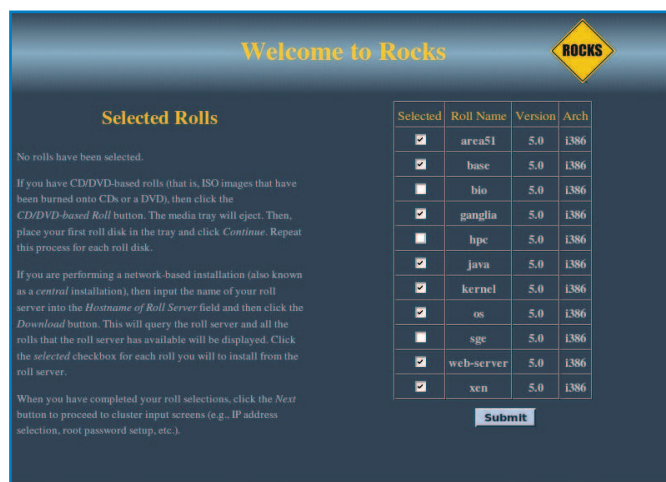


Figura 2: Seleccionando *rolls* en Rocks.



Figura 3: El Tío Sam reclama su voto para un clústering más eficiente.

una indicación le pedirá que cree un archivo de claves SSH. Para aceptar los valores predeterminados, pulse [Enter], a menos que prefiera introducir una frase de paso para la clave SSH.

Seguidamente teclee la orden `insert-ethers` en la terminal y seleccione `VM Container` bajo `Choose Application Type`. Aparecerá un cuadro de mensaje que dice `Inserted Appliances`, debiendo aparecer la dirección MAC del contenedor VM (que se corresponde con la de la máquina física) una vez que se ha arrancado.

Asegúrese de que el contenedor VM soporta el arranque por red (PXE) y establece el arranque por red antes de cualquier otro método de inicio. Si el sistema no soporta arranque por PXE, será necesario que iniciemos manualmente el contenedor VM con el CD del roll del Kernel o el DVD de Rocks en el contenedor VM simplemente insertando el DVD/CD y dejando que arranque desde los medios de instalación.

Téngase en cuenta que si no hay servidor DHCP en la red de los nodos de computación, al arrancarlos agotarán el tiempo de espera cuando intenten solicitar la asignación de dirección IP a través de DHCP. En el caso de un gran clúster abastecido con varios segmentos de red necesitaremos un servidor DHCP, ya que los timeouts de los arranques tardarán demasiado.

**Nota:** Si tiene un conmutador gestionado, unidad de potencia o un dispositivo NAS en la red privada del clúster de computación, recuerde configurar estos dispositivos antes de instalar contenedores VM o nodos de computación. Verifique las direcciones MAC en la consola `insert-`

`ethers` como dispositivos arrancables por la instalación.

El contenedor VM aparecerá en la sección `Inserted Appliances` con la dirección MAC del contenedor. Una vez que el nodo frontal solicita una configuración Kickstart, se mostrará entre paréntesis un asterisco (\*) que incluirá los rolls predeter-

minados incluidos en la instalación inicial del frontal.

Ahora relájese y deje que Rocks haga todo el trabajo. Lo mejor de Rocks es que hace muy simple el aprovisionamiento de clusters – siempre que se realice una instalación automatizada del nodo frontal basada en Kickstart. La instalación Kickstart desplegará los paquetes de instalación en el contenedor VM. Ahora es un buen momento para un café. Cuando regrese, deberá encontrar un nuevo contenedor VM listo para el despliegue de nodos basados en Xen. El nombre predeterminado para el primer contenedor VM es `vm-container-0-0`.

## Creación e Instalación de Nodos de Computación Virtuales

Ahora que el contenedor VM se ha instalado, necesita crear los nodos de computación Xen en el contenedor VM para realizar el trabajo para el grupo. Se puede crear cualquier número de contenedores VM y mezclarlos con nodos de computación no virtuales. Para pasar un nodo de computación a un estado activo y utilizable, solamente se deben ejecutar unas cuantas órdenes desde el front-end para iniciar el nodo de computación.

Para crear un nodo de computación de máquina virtual, ejecute la siguiente orden en el front-end:

```
#rocks add host vm ↗
vm-container-0-0 membership=↗
"Compute"
```

Para establecer los parámetros de configuración de varias máquinas virtuales, como tamaño de memoria, tamaño

de disco e información de la red, puede pasar argumentos opcionales a la orden `add`. El tamaño predeterminado de memoria para una máquina virtual creada con esta orden es de 1.024 megabytes.

Una vez ejecutado el comando, se añade la configuración de un nuevo nodo de computación virtual a la base de datos de Rocks, apareciendo la siguiente salida:

```
Added VM on node ↗
"vm-container-0-0" slice "1" ↗
with vm_name "compute-0-0-1"
```

Al intentar crear la máquina virtual Xen con la orden anterior, podría aparecer un mensaje de error indicando que no hay suficiente memoria en la máquina. En tales casos, asegúrese de tener suficiente memoria física en el contenedor VM, así como para las máquinas virtuales huésped. Como último recurso, la cantidad de memoria asignada por VM puede reducirse mediante el parámetro de la configuración del dominio de configuración de memoria de Xen `dom0-min-mem` en el archivo de configuración `/etc/xen/xend-config.sxp`.

Ahora que el contenedor VM ha sido asignado al contenedor VM físico, la instalación de Xen VM en el contenedor de destino se ejecuta escribiendo `rocks create host vm compute-0-0-1`. Esta orden lee la configuración de la base de datos de Rocks y comienza automáticamente la provisión de máquinas virtuales de Xen en el contenedor de destino. Escribiendo `rocks-console compute-0-0-1` puede ver el progreso de la instalación.

## Referencia Rápida de Comandos

Parar un ordenador nodo de un clúster:

```
# rocks stop host vm ↗
compute-x-x-x
```

Iniciar un ordenador nodo de un clúster:

```
# rocks start host vm ↗
compute-x-x-x
```

Solicitar información de todos los nodos:

```
# rocks list host
```

Solicitar información de todos los nodos virtuales de un ordenador:

```
#rocks list host vm
```

Una vez que ha creado la máquina virtual, puede iniciarla ejecutando esta orden desde el front-end:

```
rocks start host VM ↗
compute-0-0-1.
```

Si arranca satisfactoriamente, el terminal mostrará lo siguiente:

```
Using config file ↗
'/etc/xen/rocks/compute-0-0-1'
Started domain compute-0-0-1
```

Ahora repita el proceso de asociar a la máquina virtual con una máquina física, creando y arrancando la VM tantas veces como desee o hasta agotar la memoria física libre del contenedor VM. El clúster está ahora listo para realizar las órdenes del científico loco que lo concibió y le dio la vida.

## Administración de Nodos de Computación Virtuales

Tras crear un nodo de computación virtual, posiblemente desee consultar el estado de la VM desde el nodo frontal para asegurarse de que las máquinas virtuales han arrancado satisfactoriamente. Para consultar el estado de los nodos, ejecute la orden `rocks list host compute-0-0-1`; un resultado similar al Listado 1 proporcionará información acerca de las máquinas virtuales instaladas actualmente y sus estados en todos los contenedores VM. (La condición de salida está en blanco

porque los procesos no se han programado para correr en la recién creada VM). Así mismo, puede comprobar el estado de las VMs en el mismo contenedor VM con la orden `Xen xm` (Listado 2).

La orden `Rocks` es la orden de administración principal para todo el sistema de gestión de Rocks. La estructura básica de la línea de órdenes es `rocks <orden> <argumentos>`. Hay disponible una lista completa de los comandos de Rocks en el sitio web [rocksclusters.org](http://rocksclusters.org) o escribiendo `rocks list help` en la línea de comandos. La orden `rocks` permite arrancar y parar los nodos de computación, cambiar las configuraciones y consultar las entradas de la configuración.

Rocks incluye utilidades de administración distribuidas para la ejecución de órdenes contra todo el clúster o un grupo de nodos del clúster. El comando predefinido para la ejecución distribuida de una orden es `tentakel`, que se describe como un programa para ejecutar la misma orden en muchos hosts en paralelo. `tentakel` es fácil de utilizar, y Rocks añade automáticamente todos los nodos del clúster a `/etc/tentakel.conf` en varias agrupaciones. Para ejecutar la orden contra todos los nodos en `/etc/tentakel.conf` escriba `tentakel <orden>`. Para ejecutarla contra un grupo de nodos introduzca `tentakel -g <nombre_grupo> <orden>`.

## Aplicaciones Clusterizadas

Para utilizar todos los nodos del cluster en una sola aplicación debe estar dise-

ñada para el clúster. OpenMPI es una librería de software comúnmente utilizada para diseñar aplicaciones distribuidas que se ejecutan sobre clusters. Sun Grid Engine y Torque son sistemas de colas para distribuir trabajos entre los elementos de un clúster. Aunque el problema de diseñar e implementar aplicaciones para que se usen todos los recursos del clúster es un tema amplio, encontrará recursos en línea sobre el uso de Sun Grid Engine, Torque y Condor para estos fines.

## Conclusión

La facilidad de uso y el apoyo de las aplicaciones distribuidas más comunes de Rocks hacen que sea la favorita de los científicos que trabajan en instalaciones de supercomputación para desbloquear los misterios del universo. ■

## RECURSOS

- [1] Rocks clusters: <http://www.rocksclusters.org/>
- [2] Sun Grid Engine: <http://gridengine.sunsource.net/>
- [3] OpenMPI: <http://www.open-mpi.org/>
- [4] Guía de usuario del Sun Grid Engine Roll: <http://www.rocksclusters.org/roll-documentation/sge/4.3/roll-sge-usersguide.pdf>
- [5] Página de documentación de Rocks: [http://www.rocksclusters.org/wordpress/?page\\_id=4](http://www.rocksclusters.org/wordpress/?page_id=4)

### Listado 1: Verificación de Estado de Máquinas Virtuales

```
01 [root@clúster ~]# rocks list host vm
02
03 VM-HOST      SLICE  MEM    CPUS  MAC                HOST                STATUS
04 compute-0-0-1: 1      512    1     00:16:3e:00:00:02  vm-container-0-0   active
05 compute-0-0-2: 2      512    1     00:16:3e:00:00:03  vm-container-0-0   active
06 compute-0-0-3: 3      512    1     00:16:3e:00:00:04  vm-container-0-0   active
```

### Listado 2: La Orden xm

```
01 [root@vm-container-0-0 ~] xm list
02 Name          ID      Mem(MiB)  VCPUs  State  Time(s)
03 Domain-0     0       1024      1      r-----  4035.5
04 Compute-0-0-1 2       512       1      -b----  23174.3
05 Compute-0-0-2 3       512       1      r-----  3228.3
06 Compute-0-0-3 4       512       1      -b----  2114.
```