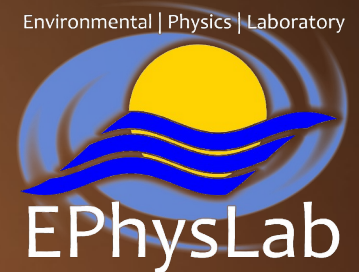


Creación de un cluster HPC con GNU/Linux

Jornadas sobre HPC y Software Libre (5-7 de Octubre 2011, Campus de Ourense)

Orlando García Feal <orlando@uvigo.es>
Ephyslab, Universidade de Vigo



Instalación CentOS



CentOS-5
Community ENTERprise Operating System



- To install or upgrade in graphical mode, press the **<ENTER>** key.
- To install or upgrade in text mode, type: **linux text <ENTER>**.
- Use the function keys listed below for more information.

[F1-Main] [F2-Options] [F3-General] [F4-Kernel] [F5-Rescue]
boot: _

Instalación CentOS



Instalación CentOS



Bienvenido a CentOS

Tipo de teclado

¿Qué tipo de teclado tiene?

- cz-us-qwertz
- de
- de-latin1
- de-latin1-nodeadkeys
- dk
- dk-latin1
- dvorak
- es**

Aceptar Anterior

<Tab>/<Alt-Tab> entre elementos ; <Espacio> seleccionar ; <F12> siguiente

Instalación CentOS



Bienvenido a CentOS

Método de instalación

¿Qué tipo de dispositivo contiene los paquetes para ser instalados?

CDROM Local
Disco duro
Imagen NFS
FTP
HTTP

AceptarAnterior

<Tab>/<Alt-Tab> entre elementos ; <Espacio> seleccionar ; <F12> siguiente

Instalación CentOS



Bienvenido a CentOS

Configurar TCP/IP

☒ Activar soporte IPv4

☐ Dynamic IP configuration (DHCP)

☒ Manual configuration

☐ Activar soporte IPv6

☒ Automatic neighbor discovery (RFC 2461)

☐ Dynamic IP configuration (DHCP)

☐ Manual configuration

Aceptar

Anterior

<Tab>/<Alt-Tab> entre elementos ; <Espacio> seleccionar ; <F12> siguiente

Instalación CentOS



Bienvenido a CentOS

Configuración manual de TCP/IP

Introduzca la dirección y prefijo (dirección / prefijo) IPv4 y/o IPv6. Para IPv4 se acepta la máscara de red de cuatro dígitos separados por comas o el prefijo CIDR. Los campos para la puerta de enlace y el servidor de nombres deben ser direcciones IPv4 o IPv6 válidas.

Dirección IPv4: 10.0.5.1 / 16
Puerta de enlace: 10.0.0.1
Servidor de nombres: 8.8.8.8

Aceptar

Anterior

<Tab>/<Alt-Tab> entre elementos ; <Espacio> seleccionar ; <F12> siguiente

Instalación CentOS



P.ej. : http://sunsite.rediris.es/mirror/CentOS/5.7/os/x86_64/

Nombre del sitio: sunsite.rediris.es

Directorio CentOS: mirror/CentOS/5.7/os/x86_64/

Bienvenido a CentOS

Configuración HTTP

Por favor introduzca la siguiente información:

- o el nombre o número IP de su servidor Web
- o el directorio del servidor que contiene CentOS para su arquitectura

Nombre del sitio web:

Directorio CentOS:

Aceptar

Anterior

<Tab>/<Alt-Tab> entre elementos | <Espacio> seleccionar | <F12> siguiente

Instalación CentOS



Instalación CentOS



La instalación requiere la partición de su disco duro. Por defecto, se escoge una capa de partición razonable para la mayoría de usuarios, ésta es suficiente en la mayoría de los casos. Usted puede escoger esta partición predeterminada o crearla usted mismo.

Suprimir particiones de Linux en dispositivos seleccionados y crear diseño predeterminado.

☐ Sistema de cifrado

Seleccione la(s) unidad(es) a usar para esta instalación.

☒ hda 9994 MB QEMU HARDDISK

[+ Configuración Avanzada de almacenamiento](#)

☐ Revise y modifique la capa de particiones


 [Notas de lanzamiento](#)

[← Atrás](#)

[→ Siguiente](#)

Instalación CentOS



 **CentOS**

Dispositivos de red

Activar al inicio	Dispositivo	IPv4/Máscara de red	IPv6/Prefijo
<input checked="" type="checkbox"/>	eth0	10.0.5.1/16	Desactivado

Modificar

Nombre del Host

Configurar el nombre del host:

☐ de forma automática a través de DHCP

☒ manualmente (ej. "mipc.dominio.com.ar")

Configuración miscelánea

Puerta de enlace:

DNS Primario:

DNS Secundario:

Notas de lanzamiento

Atrás

Siguiente

Instalación CentOS



 **CentOS**


Por favor haga click en el mapa para seleccionar una región



Europa/Madrid

mainland

☒ El reloj del sistema utiliza UTC

 [Notas de lanzamiento](#)

[← Atrás](#)

[→ Siguiente](#)

Instalación CentOS



CentOS



La cuenta root se utiliza para la administración del sistema. Introduzca una contraseña para el usuario root.

Contraseña de root:

.....

Confirmar:

.....



Notas de lanzamiento

← Atrás

Siguiente →

Instalación CentOS



CentOS

La instalación por defecto de CentOS incluye un grupo de aplicaciones para el uso general de Internet. ¿Qué tareas adicionales le gustaría poder realizar en su sistema?

☒ Desktop - KDE

☒ Server

☐ Server - GUI

☐ Virtualización

Por favor, seleccione cualquier repositorio adicional que quiera usar para la instalación de software.

[+ Agregar repositorios de software adicionales](#)

La selección de software se puede personalizar ahora o después de la instalación mediante la aplicación de administración del sistema.

☒ Personalizar luego ☐ Personalizar ahora

 [Notas de lanzamiento](#)

[← Atrás](#)

[→ Siguiente](#)


Instalación CentOS





Pulse en Siguiente para iniciar la instalación de CentOS.

El registro completo de la instalación puede encontrarse en el archivo '`/root/install.log`' luego de reiniciar su sistema.

Podrá encontrar un archivo kickstart con las opciones de instalación seleccionadas en el archivo '`/root/anaconda-ks.cfg`' luego de reiniciar el sistema.

 [Notas de lanzamiento](#)

 [Atrás](#)

 [Siguiente](#)

Instalación CentOS



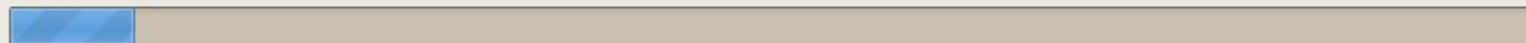


Repositorios de software

Los siguientes repositorios existen en CentOS desde los cuales se pueden instalar aplicaciones:

- [base]** (alias **[os]**) - RPMs presentes en los ficheros ISO de CentOS.
- [updates]** - Actualizaciones al repositorio **[base]**.
- [extras]** - Paquetes que no están presentes en el proveedor (no mejoran paquetes de **[base]**).
- [centosplus]** - Paquetes que no están presentes en el proveedor (mejoran paquetes de **[base]**).

<http://wiki.centos.org/AdditionalResources/Repositories>



Instalando glibc-2.5-65.x86_64 (11 MB)
Las bibliotecas libc de GNU.



Notas de lanzamiento

← Atrás

→ Siguiente

Instalación CentOS



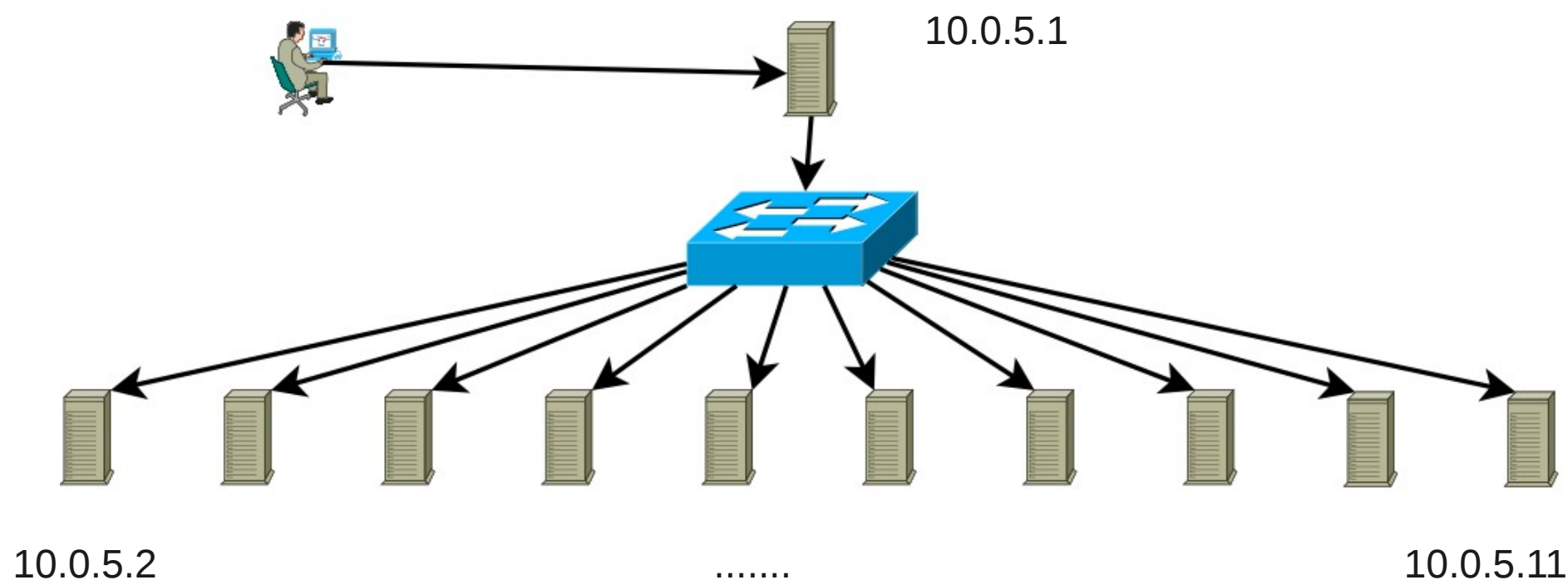
Configuración cluster



Frontend: 2 interfaces de red (uno red externa, otro red interna).

Nodos de cálculo: 1 interfaz de red (red interna).

Instalar a continuación los nodos de cálculo por el método preferido (uno a uno, clonado de disco, pxe...).



Configuración cluster



Creación del almacenamiento compartido (en este caso centralizado en el frontend y compartido mediante nfs).

Suponemos que tenemos nfs instalado y el correspondiente demonio inicializado.

En el frontend compartiremos el directorio “/home” editando /etc/exports y añadiendo la línea :

```
/home 10.0.5.0/24(sync,no_wdelay,subtree_check,rw,root_squash)
```

Y reiniciamos nfs ejecutando “/etc/init.d/nfs restart”.

En el resto de nodos añadimos el punto de montaje del directorio home compartido añadiendo la siguiente línea al fichero /etc/fstab :

```
10.0.5.1:/home /home nfs vers=3,noac,defaults,auto 0 0
```

Configuración cluster



Para la creación del cluster necesitamos instalar:

- Un planificador de procesos: en este caso **MAUI**
- Un gestor de colas de trabajos: en este caso **TORQUE**

En este caso los compilaremos e instalaremos a partir del código fuente:

1.- Descargamos el código fuente de:

<http://www.clusterresources.com/downloads/torque/>

<http://www.clusterresources.com/downloads/maui/>

2.- Descomprimos los paquetes:

```
# tar -xzf *.tar.gz
```

3.- Compilamos TORQUE y creamos los paquetes instalables:

```
# ./configure  
# make  
# make packages
```

Configuración cluster



4.- Instalamos los paquetes en el frontend:

```
# ./torque-package-server-linux-x86_64.sh --install
# ./torque-package-mom-linux-x86_64.sh --install
# ./torque-package-doc-linux-x86_64.sh --install
# ./torque-package-devel-linux-x86_64.sh --install
# ./torque-package-clients-linux-x86_64.sh --install
```

5.- Configuramos torque (luego para iniciar el server ejecutar “pbs_server”):

```
# ./torque.setup root
```

6.- En cada uno de los nodos instalaremos los paquetes:

```
# ./torque-package-mom-linux-x86_64.sh --install
# ./torque-package-clients-linux-x86_64.sh --install
```

8.- Ejecutar “pbs_mom” para incializar cada uno de los nodos.

Configuración cluster



9.- Añadir la lista de nodos en el frontend (recuerda tener en /etc/hosts la lista de nodos del cluster), esto se puede hacer editando el fichero “/var/spool/torque/server_priv/nodes” o bien mediante la consola de administración “qmgr”.

```
nodo1 np=2 batch
nodo2 np=2 batch
...
```

10.- Editar el fichero /var/spool/torque/mom_priv/config de cada uno de los nodos e insertar el contenido:

```
$logevent          255
$usecp              *:/home /home
```

11.- Instalar MAUI. Para ello entramos en el directorio y lo compilamos:

```
# ./configure --with-pbs --with-spooldir=/var/spool/maui
# make
# make install
```

Configuración cluster



12.- Configurar fichero /usr/local/maui/maui.cfg

```
# maui.cfg 3.3.1

SERVERHOST          <nombre de host del frontend>
# primary admin must be first in list
ADMIN1              root
# Resource Manager Definition
RMCFG[SPECTRUM] TYPE=PBS
# Allocation Manager Definition
AMCFG[bank] TYPE=NONE
# full parameter docs at http://supercluster.org/mauidocs/a.fparameters.html
# use the 'schedctl -l' command to display current configuration
RMPOLLINTERVAL      00:00:30
SERVERPORT           42559
SERVERMODE           NORMAL
# Admin: http://supercluster.org/mauidocs/a.esecurity.html
LOGFILE              maui.log
LOGFILEMAXSIZE       10000000
LOGLEVEL             3
# Job Priority: http://supercluster.org/mauidocs/5.1jobprioritization.html
QUEUETIMEWEIGHT      1
# FairShare: http://supercluster.org/mauidocs/6.3fairshare.html
#FSPOLICY             PSDEDICATED
#FSDEPTH              7
#FSINTERVAL           86400
#FSDECAY              0.80
# Throttling Policies: http://supercluster.org/mauidocs/6.2throttlingpolicies.html
# NONE SPECIFIED
# Backfill: http://supercluster.org/mauidocs/8.2backfill.html
BACKFILLPOLICY        FIRSTFIT
RESERVATIONPOLICY     CURRENTHIGHEST
# Node Allocation: http://supercluster.org/mauidocs/5.2nodeallocation.html
NODEALLOCATIONPOLICY   MINRESOURCE
# QOS: http://supercluster.org/mauidocs/7.3qos.html
```

Configuración cluster



```
# QOSCFG[hi]  PRIORITY=100 XFTARGET=100 FLAGS=PREEMPTOR:IGNMAXJOB
# QOSCFG[low] PRIORITY=-1000 FLAGS=PREEMPTTEE

# Standing Reservations: http://supercluster.org/mauidocs/7.1.3standingreservations.html

# SRSTARTTIME[test] 8:00:00
# SRENDTIME[test]   17:00:00
# SRDAYS[test]      MON TUE WED THU FRI
# SRTASKCOUNT[test] 20
# SRMAXTIME[test]   0:30:00

# Creds: http://supercluster.org/mauidocs/6.1fairnessoverview.html

# USERCFG[DEFAULT]      FSTARGET=25.0
# USERCFG[john]          PRIORITY=100  FSTARGET=10.0-
# GROUPCFG[staff] PRIORITY=1000 QLIST=hi:low QDEF=hi
# CLASSCFG[batch]        FLAGS=PREEMPTTEE
# CLASSCFG[interactive]  FLAGS=PREEMPTOR
```


Configuración cluster



13.- Añadimos al path la ruta de MAUI y lo Iniciamos

```
# PATH=$PATH:/usr/local/maui/bin:/usr/local/maui/sbin  
# maui
```

14.- Los usuarios que envíen trabajos a la cola deben existir en los nodos, para ello lo ideal es disponer de un servidor LDAP.

Manejo del sistema de colas



Nuestro cluster ya debería estar funcionando. Para comprobar que los nodos están conectados correctamente ejecutamos “pbsnodes”.

Para comprobar que definitivamente funciona todo enviar un trabajo a la cola.

Comandos clave:

- pbs_mom : inicializa un nodo
- pbs_server : inicializa el servidor
- qterm : finaliza el proceso servidor
- maui: inicializa el gestor de recursos
- pbsnodes : muestra los nodos disponibles y su estado
- qstat : muestra los trabajos en cola
- qhowq: muestra el estado de las colas
- qsub : envía un trabajo a una cola

Manejo del sistema de colas



qsub

- > **qsub nombrescript.sh** #Envía el script a un nodo cualquiera
- > **qsub** # Envía a un nodo cualquiera un script leído por STDIN
- > **qsub -l nodes=5** # Envía un script exigiendo 5 nodos
- > **qsub -l nodes=1:ppn=8** # Envía un script exigiendo 1 nodo y 8 cores
- > **qsub -l mem=200mb** # Envía un script solicitando 200mb de memoria
- > **qsub -l nodes=nodo1:ppn=1+nodo2:ppn=2** # Solicita nodos y cores específicos
- > **qsub -l walltime=5:00:00** # Envía un script solicitando 5 horas para su ejecución

Manejo del sistema de colas



qstat

Qstat nos permite visualizar los trabajos que hemos enviado a la cola:

360.lechuck						
Job id	Name	User	Time Use	S	Queue	
-----	-----	-----	-----	-	-----	
360.lechuck	script	usuario		0 Q	batch	

Donde:

Job id: Identificador único del trabajo.

Name: Nombre del trabajo.

User: Usuario propietario del trabajo.

Time Use: Tiempo de uso de recursos del trabajo.

S: Estado actual del trabajo (Q= En espera, R= En ejecución, E= Error, C= Completado, H= parado).

Queue: Cola de trabajos actual (por defecto "batch").

Manejo del sistema de colas



qdel

Qstat nos permite borrar un trabajo de la cola de trabajos.

```
# qdel <id_trabajo>
```

Documentación



Torque: <http://www.clusterresources.com/torquedocs21/index.shtml>