

Uso del Cluster Rigel

Entorno de usuario

Todos los sistemas de RIGEL (Cluster General, Cluster Memoria Compartida y Cluster GPUs) tienen instalada la misma versión de Sistema Operativo, CentOS 6, y la misma configuración, tanto del sistema operativo como librerías y utilidades.

El gestor de colas es SGE, y todo trabajo debe ser enviado a través de colas para su ejecución. Cuando un usuario se da de alta en RIGEL, se le indicará el modo de enviar sus trabajos a través de manuales y documentación en línea.

Las opciones especificadas en el envío de trabajos permiten al sistema SGE asignar a las diferentes colas los trabajos según requerimientos especificados.

El esquema de autenticación es por LDAP y la contraseña se valida con Active Directory de la UPV.

Los usuarios no pueden acceder de modo interactivo a los nodos de cálculo, sólo trabajan en los nodos de cabecera, a los que pueden acceder por ssh. El frontend de usuario es un cluster en alta disponibilidad de dos sistemas, y su dirección es rigel.cc.upv.es

El directorio home del usuario está montado en /home/grupo/usuario en todos los nodos de todos los clusters.

Sistema de colas

Para que el uso del sistema sea lo mas efectivo posible, todos los trabajos de cálculo deben enviarse siempre por medio del gestor de colas.

En el siguiente resumen se muestran las ordenes mas importantes para enviar un trabajo al gestor de colas.

El gestor de colas le asignara recursos cuando le llegue su turno.

El gestor de colas instalado en en Rigel es SGE ([SUN Grid Engine](http://www.sun.com)).

Ordenes del sistema de colas

Para enviar a ejecutar un script utilizar qsub script:

El sistema responderá con información sobre el trabajo enviado así como un identificador.

```
rigel$ qsub run.sh
cpus=8 h_rt=0h:10m:0s h_vmem=1g cputime=1h:20m:0s
Your job 7283 ("run.sh") has been submitted
```

En el script se deben incluir ciertas directivas que indican al sistema de colas los recursos (memoria, procesadores, tiempo) que se requieren. Debe indicarse el entorno paralelo y numero de cpus (nodes), la cantidad de memoria por core (h_vmem), y el tiempo maximo de ejecucion (h_rt).

```
# Ejemplo de cabecera de un script qsub.
# Las lineas que comienzan por #$ son indicaciones al sistema de colas
# Solicitar 8 tareas paralelas en entorno MPI
# Solicitar 2 GB de memoria por core y 30 minutos de tiempo de ejecucion.
#$ -pe mpi 8
#$ -l h_vmem=2gb,h_rt=00:30:00
```

Si se omite el parametro -pe se asigna 1 core.

Si se omite el parametro h_vmem se establece un limite de 1 GB.

Si se omite el parametro h_rt se establece un limite de 1 hora.

Pueden verse algunos ejemplos en el directorio /ayuda/ejemplos de rigel.

Por defecto SGE deposita la salida de los trabajos en ficheros que se crean en el directorio de trabajo. El nombre de estos ficheros tiene el formato

```
run.sh.e7283
run.sh.o7283
```

donde la e significa *error* y la o *salida*. Es decir contienen los errores y el resultado de la ejecución respectivamente.

Para consultar el estado de los trabajos encolados qstat:

Sin parámetros devuelve los trabajos del propio usuario.

```
rigel$ qstat
job-ID prior name user state submit/start at queue slots ja-task-ID
-----
7284 0.00000 run.sh ico qw 07/11/2013 14:07:03 8
```

Donde se puede observar que el trabajo está en cola y esperando, *qw*.

Cuando el trabajo entra ejecución la salida nos muestra en qué cola (y nodo) se está ejecutando.

```
rigel$ qstat
job-ID prior name user state submit/start at queue slots ja-task-ID
-----
7284 0.56000 run.sh ico r 07/11/2013 14:07:12 general@hpc033.cc.upv.es 8
```

Con el parámetro *-g t* mostrará todas las tareas en paralelo para el trabajo.

Se pueden obtener los detalles de un trabajo concreto con *-j identificador*.

Para cancelar un trabajo:

```
qdel identificador
```

Notas

- Las colas definidas en este momento limitan el tiempo de ejecución de un trabajo a un máximo de 144 horas y el tiempo total de cpu (numero de procesadores * tiempo de ejecución) a un máximo de 1000 horas. Estos parámetros se pueden ir modificando para optimizar el rendimiento del sistema.

Ejemplos de parámetros de qsub:

```
-pe mp 9 -l exclusive=true,h_rt=2:00:00,h_vmem=2g
```

(trabajo OpenMP que ocupa 1 cores de un mismo nodo durante 2 horas)

```
-pe mpi 9 -l exclusive=true,h_rt=2:00:00,h_vmem=2g
```

(trabajo MPI que ocupa 4 cores de diversos nodos durante 2 horas)

```
-cwd
```

(iniciar trabajo en directorio actual en lugar de \$HOME)

Más información sobre el uso del sistema de colas en las páginas de manual de *qsub*, *qstat*, *qdel*.

SGE USeR Guide



SGEUserGuide.pdf