

Qué es SLURM

SLURM es un sistema de planificador de trabajos y administración de clústeres de código abierto y altamente escalable para clústeres grandes como es el TeideHPC.

Este planificador tiene tres funciones clave.

- Asigna acceso exclusivo y/o no exclusivo a los recursos (nodos de cómputo) a los usuarios durante un período de tiempo para que puedan realizar el trabajo.
- Proporciona un marco de trabajo para iniciar, ejecutar y monitorear los trabajos (normalmente un trabajos paralelos) en el conjunto de nodos asignados.
- Arbitra la disputa por recursos administrando una cola de trabajo pendiente.

Actualmente en el cluster se dispone de la **versión 19** de SLURM.

Los conceptos más importantes dentro de slurm son:

- Nodos de cómputo.
- Particiones.
- Trabajos.
- Tareas que representan un proceso dentro de un trabajo.

Nodos de login

Desde los nodos de login es desde donde el usuario interactúa con Slurm y desde donde lanza y monitoriza sus trabajos. Desde aquí, el usuario accede a sus datos y a los resultados de las ejecuciones.

Info

Hay que recordar que los **nodos de login** son de uso compartido por todos los usuarios, por lo que **está prohibida la ejecución de software en dichos nodos**. Para ello se debe usar el planificador Slurm.

Particiones

Las particiones se pueden considerar *colas de trabajos*, cada una de las cuales pueden tener una variedad de restricciones como límite de tamaño del trabajo, límite de

tiempo del trabajo, usuarios autorizados a usarlo, etc. Al lanzarse los trabajos en una partición concreta, éstos son ordenados por prioridad dentro de una partición hasta que los recursos (nodos, procesadores, memoria, etc.) dentro de esa partición están agotados pudiendo quedar a la espera.

Las particiones definidas en nuestro cluster son las siguientes y sus únicas restricciones son el tiempo y los usuarios que puede usarla:

Partición	Tiempo máximo	
express	3 horas	<i>todos los usuarios</i>
batch	24 horas	<i>todos los usuarios</i>
long	72 horas	<i>usuarios autorizados</i>
fatnodes	--	<i>previa solicitud a los administradores</i>
gpu	--	<i>previa solicitud a los administradores</i>

Si no se especifica ninguna, la partición por defecto es la `batch`.

Si su trabajo va a durar más del tiempo máximo que establecido por la partición donde lo ha lanzado contacte con support@hpc.iter.es para pedir una ampliación de tiempo límite.

Comandos más comunes en SLURM

Para familiarizarse con el planificador a continuación se detallan los comandos más comunes de slurm. Para obtener más información acerca del comando como opciones siempre puede ejecutar

```
man <command>  
<command> --help
```

- `sbatch <script file>` : lanzar un script al gestor de colas
- `squeue` : verificar el estado de los trabajos en las colas
- `scancel <job_id list>` : cancelar un trabajo
- `scontrol show job <job_id>` : obtener información del estado de un trabajo
- `sinfo` : ver estado de las colas del sistema

- `salloc <opciones>` : iniciar sesión interactiva (obtener un nodo para ejecutar)
- `srun <aplicacion>` : enviar un trabajo a ejecutar o inicia los pasos del trabajo en tiempo real
- `sacct` : consultar el accounting de la propia cuenta
- `sstat` : obtener información sobre los recursos utilizados por un trabajo en ejecución

Puede ver una guía resumen más detallada en nuestra sección [Comandos útiles en slurm](#) o en la [página oficial de slurm](#).