



IPICYT
INSTITUTO POTOSINO DE
INVESTIGACIÓN CIENTÍFICA
Y TECNOLÓGICA, A.C.



CNS
CENTRO NACIONAL
DE SUPERCÓMPUTO
IPICYT

Guía de uso para la supercomputadora Thubat Kaal II



Antecedentes

Clúster Thubat-Kaal II (**TK II**), quedó fuera de línea por fallas técnicas que obligaron a detener el servicio. Este equipo fue reestructurado y sometido a un mantenimiento exhaustivo; reduciendo a la mitad el número de nodos de cálculo disponible. **TK II** ha sido puesto en línea nuevamente.

Tomando en cuenta los antecedentes; el servicio de HPC es otorgado al usuario bajo ciertas consideraciones y limitaciones, efecto de la reestructuración del equipo.

Objetivo

El usuario será informado a través de esta guía, sobre el uso correcto de **TK II** a modo que considere las adaptaciones a consecuencia de la reestructuración del equipo.

Distribución de los recursos de cómputo del CNS (infraestructura, características y capacidades).

Thubat Kaal II es un clúster que alberga a una serie de usuarios y realiza distintas funciones. Su infraestructura está dividida en tres secciones de suma importancia para garantizar la ejecución de las tareas y la seguridad máxima en la privacidad de la información: Procesamiento, interconexión y almacenamiento.

Como primera parte tenemos el procesamiento, la cual está clasificada en los siguientes tipos de nodos: cálculo estándar, cálculo Fat y cálculo GPU, como se describe a continuación.

Nodos de cálculo estándar

Cada nodo tipo estándar, tiene 2 procesadores de 16 cores físicos c/u haciendo en total una suma de 32 cores físicos; cada core físico puede realizar 2 operaciones simultáneas, llamando a esto trabajar a 2 hilos por core. Esto nos ofrece el equivalente de 64 cores disponibles por nodo para ejecución. Cada nodo cuenta con una capacidad de memoria RAM de 192GB, esta configuración debe de ser considerada en la planeación de ejecución. Actualmente se cuenta con 40 nodos de este tipo.

Nodos de cálculo tipo Fat

Los nodos de cálculo tipo Fat, tendrán la ventaja de poder correr aplicaciones en modo serial, aprovechando los recursos de memoria RAM de 384 GB. Cada nodo tipo Fat; al igual que los nodos estándar tienen 64 cores disponibles por nodo para ejecución. Esta configuración debe de ser considerada en la planeación de ejecución. Actualmente se cuenta con 8 nodos de este tipo.

Nodos de cálculo tipo GPU

Cada nodo tipo GPU, tiene 2 procesadores con 14 cores físicos, dando un total de 28 cores; contiene 4 tarjetas gráficas NVIDIA P100 con una capacidad de 3,584 cuda-cores cada una, dando un total de 14,334 cuda-cores por nodo. Actualmente se cuenta con 4 nodos de este tipo.

Almacenamiento scratch

A través de éste, se realizan la mayoría de las transacciones de los usuarios. Dicho almacenamiento cuenta con una capacidad de 60TB de espacio disponible.

El usuario cuenta con un espacio de 1 TB de almacenamiento para guardar su información.

Acceso al clúster Thubat Kaal II

Configuración VPN

El primer paso para realizar la conexión al clúster, es la instalación de un cliente de VPN en el equipo, donde iniciara sesión cuando requiera conectarse a la supercomputadora. Las credenciales de acceso y el manual para la instalación y configuración serán proporcionadas por medio de correo electrónico por parte del personal del departamento de servicios de supercómputo.

Acceso a través del Nodo Maestro

Una vez que el usuario ha realizado la instalación y configuración de su cliente VPN en su equipo, deberá ingresar al clúster conectándose a través del servidor maestro bajo el protocolo de seguridad ssh.

Por medio de la terminal (Linux o MAC) o aplicación para conexión tipo SSH (Windows), podrá iniciar sesión con el siguiente comando:

```
ssh -X hpc-usuario@192.168.91.241
```

Al ingresar por primera vez, le solicitará el cambio de contraseña, es importante que recuerde su nuevo password.

Una vez iniciada la sesión, tendrá un límite de 100 GB de espacio en su /home, por lo que solo deberá guardar archivos esenciales. Por esta razón deberá utilizar de manera obligatoria su segundo directorio localizado en /mnt/scratch_huawei/INTERNOS/usuario (usuarios IPICYT) o /mnt/scratch_huawei/EXTERNOS/institución/usuario (usuarios externos) para guardar sus datos y realizar sus ejecuciones.

Al acceder por el nodo maestro, se requiere un mayor cuidado, es por eso que queda estrictamente prohibido la ejecución de trabajos dentro del mismo nodo; para todo cálculo se debe de utilizar el gestor de recursos Slurm y los nodos de procesamiento.

Revisión y carga de software disponible del CNS (módulos)

Para poder utilizar las aplicaciones se tienen que ocupar una serie de comandos para administrar los módulos que las contienen, a continuación, se muestra cuáles son y su función.

module avail: muestra una lista de los módulos (aplicaciones) disponibles dentro del clúster.

module load: carga el módulo de la aplicación a utilizar, por ejemplo:

```
module load apps/gromacs/2024.2-cpu
```

module list: lista los módulos cargados para el usuario.

module unload: quita un módulo cargado, por ejemplo:

```
module unload apps/gromacs/2024.2-cpu
```

module purge: quita todos los módulos cargados, para verificar este paso puede ejecutar posteriormente *module list*.

Compiladores GNU

Se realizó una nueva instalación de los compiladores GNU y para poder utilizarlos, se tiene que tomar en cuenta lo siguiente.

Al utilizar algún compilador como gcc, g++ o gfortran, se tiene que colocar el nombre del compilador seguido de guion medio y la versión, por ejemplo, para la versión 14.1.0 lo usaríamos de la siguiente forma:

gcc-14.1

g++-14.1

gfortran-14.1

Gestor de recursos Slurm

Para la ejecución de un job o trabajo, se requiere utilizar al administrador de recursos Slurm, el cual nos permitirá acceder a los nodos de cálculo, mediante un script de ejecución con el cual podremos solicitar los recursos que se necesiten, cargar los módulos y ejecutar las aplicaciones.

Estos comandos están disponibles en el nodo maestro, el cual es el nodo al que se les da acceso a los usuarios del clúster y desde donde deben de ser lanzados todos los trabajos.

A continuación, se listan las variables básicas que se pueden utilizar dentro de los scripts:

VARIABLE	DESCRIPCIÓN
#SBATCH --job-name=myjob	Nombre del trabajo.
#SBATCH --output=test_%j.log	Nombre del archivo de salida (%j se reemplaza por el ID del trabajo al mandar a ejecutar un trabajo).
#SBATCH --error=test_%j.err	Nombre del archivo de error.
#SBATCH -N 1	Número de nodos solicitados para la ejecución.
#SBATCH -n 32	Número de cores solicitados para la ejecución, sino se define, será igual a N*ntasks-per-node.
#SBATCH --ntasks-per-node=1	Número de tareas asignadas a cada nodo.
#SBATCH --time=0-00:05:00	Tiempo límite de ejecución permitidos para el trabajo.
#SBATCH --mem=1gb	Memoria para el trabajo.

#SBATCH --partition=partition-name	--partition=nodos_standard, se puede acceder a 40 nodos. --partition=nodos_fat, se puede acceder a 8 nodos. --partition=nodos_gpu, se puede acceder a 4 nodos gpu.
------------------------------------	--

Ejemplo de script de ejecución en Slurm:

```
#!/bin/bash
#SBATCH --partition=nodos_standard
#SBATCH -N 1
#SBATCH -n 32
#SBATCH --ntasks-per-node=64
#SBATCH --mem=0
#SBATCH --output mdtest.out
#SBATCH --error mdtest.err
#SBATCH --job-name="test_gromacs"

set -x

module load apps/gromacs/2024.2-cp

CPUS=32

#####
gmx_mpi grompp -f md.mdp -c md50ns.gro -t md50ns.cpt -p topol.top -o md100ns.tpr
mpirun -np $CPUS gmx_mpi mdrun -ntomp 2 -rethway -v -deffnm md100ns
```

Podrá visualizar si su trabajo está ejecutándose con el siguiente comando:

squeue -u hpc-usuario

y para revisar las particiones disponibles con el comando:

sinfo

```
[eva@cns-master-01 Test_GROMACS_redato]$ squeue -u eva
JOBID PARTITION NAME USER ST TIME NODES NODELIST(REASON)
21 nodos_sta test_gro eva R 0:13 1 cns-nodo-09
[eva@cns-master-01 Test_GROMACS_redato]$ sinfo
PARTITION AVAIL TIMELIMIT NODES STATE NODELIST
all* up infinite 5 down* cns-gpu-[01-04],cns-nodo-13
all* up infinite 1 mix cns-nodo-01
all* up infinite 1 alloc cns-nodo-09
all* up infinite 45 idle cns-nodo-[02-08,10-12,14-48]
nodos_fat up 15-00:00:0 1 mix cns-nodo-01
nodos_fat up 15-00:00:0 7 idle cns-nodo-[02-08]
nodos_standard up 15-00:00:0 1 down* cns-nodo-13
nodos_standard up 15-00:00:0 1 alloc cns-nodo-09
nodos_standard up 15-00:00:0 38 idle cns-nodo-[10-12,14-48]
escuela_hpc up 15-00:00:0 6 idle cns-nodo-[43-48]
nodos_gpu up 15-00:00:0 4 down* cns-gpu-[01-04]
[eva@cns-master-01 Test_GROMACS_redato]$
```

Nuevas políticas de uso

A continuación se enlistan las nuevas políticas de uso después de la nueva configuración del clúster. Es muy importante que las siga, ya que en caso contrario se podría prohibir el acceso a la supercomputadora.

- Solo se les otorgara un espacio de almacenamiento en su directorio home de 100GB.
- Se contará con un espacio de almacenamiento en su directorio “/mnt/scratch_huawei” de un máximo de 1TB. En caso de que el usuario rebase su almacenamiento se le solicitara que haga un borrado de su información, si hace caso omiso el departamento de servicios de supercómputo del CNS, se reserva el derecho de realizar el borrado de información.
- Solo se puede guardar información y ejecutar dentro del directorio “/mnt/scratch_huawei”, el gestor de recursos Slurm, está configurado para no encolar los trabajos que se envíen de otro directorio diferente al antes mencionado.
- Debido a los problemas que el clúster presento, se sigue con la misma configuración de cores permitidos para ejecución (128 cores) y la misma cantidad de memoria RAM (768 GB máximo) por cuenta.
- Es necesario que respalden su información periódicamente ya que el CNS no se hace responsable de la integridad de la misma por posibles fallas de hardware y software. Se les recomienda que una vez que termine la ejecución de su trabajo, hagan el respaldo de su información.
- Queda prohibido prestar o transferir su cuenta.
- No se permiten labores de post procesamiento de datos dentro del nodo maestro.
- No se permite ingresar a carpetas diferentes a las permitidas por los usuarios.

Para cualquier duda, comentario, soporte, instalación de aplicaciones, o cualquier situación relacionada con el servicio de supercómputo, mandar correo a soporte.cns@ipicyt.edu.mx para brindarle una atención más rápida y eficiente.