

Reglamento del Cluster de Supercómputo en el LANCIS

Mayo de 2017

1 Introducción

La supercomputadora del Laboratorio Nacional de Ciencias de la Sostenibilidad del Instituto de Ecología es un recurso compartido. Si bien tiene la capacidad de ejecutar cientos de tareas al mismo tiempo es necesario observar algunas reglas para su funcionamiento óptimo.

1.1 Infraestructura de hardware y software

La supercomputadora está constituida de decenas de computadoras interconectadas y configuradas para cooperar entre sí como si fueran una sola. Desde el punto de vista del usuario, el instrumento entero se usa desde el nodo central, que es un equipo dedicado a mantener sesiones interactivas y despachar tareas a los nodos de cómputo.

Esto se hace posible a través de componentes de software:

1.1.1 Sistema operativo Se ha configurado la versión 8.5 de Debian, un sistema robusto, con una amplia base de usuarios y muchos casos de éxito en el contexto de cómputo masivo. Se ha configurado para que todos los nodos de la supercomputadora

- tengan una réplica de la base de datos de usuarios
- puedan acceder al sistema principal de almacenamiento

1.1.2 Sistema de cola o fila de ejecución Se ha instalado y configurado HT-Condor, un sistema de cómputo distribuido flexible y poderoso. Este sistema permite a los usuarios compartir la infraestructura brindando un acceso ordenado al conjunto entero de recursos. En esencia implementa una cola en la que las tareas de cómputo deben formarse para su ejecución de acuerdo con el esquema de prioridades configurado. El esquema actual es: first-come-first-serve.

1.1.3 Bibliotecas especializadas Bibliotecas especializadas como Message Passing Interface (MPI) han sido instaladas. Este tipo de software permite la creación y ejecución de programas que mantienen muchos hilos de ejecución en paralelo. La supercomputadora soporta software creado con bibliotecas como MPI pues todos los nodos de cómputo tienen varios procesadores, y estos

programas pueden y deben invocarse desde HT-Condor descrito en la sección anterior.

2 Usuarios y Beneficiarios

La supercomputadora del LANCIS es el resultado del apoyo financiero del proyecto de Laboratorios Nacionales de Conacyt y otros proyectos. Para justificar su uso continuado es reglamentario:

- que cada proyecto de investigación en desarrollo sobre esta infraestructura tenga un resumen en la siguiente carpeta: <https://github.com/sostenibilidad-unam/supercomputadora/tree/master/proyectos>
- incluir en el resumen información de autores, ponencias, artículos sometidos y publicados
- agradecer a la unidad de Supercómputo del LANCIS-Instituto de Ecología en aquellas publicaciones que hayan requerido el uso de la supercomputadora.

3 Ejecución de Software

Para que los usuarios ejecuten sus programas de manera responsable es necesaria cierta familiaridad con el cómputo en paralelo y la configuración de software y hardware de la supercomputadora. Se invita a la comunidad a leer la documentación y quizá a solicitar apoyo en la mesa de ayuda.

Si alguna tarea de cómputo en ejecución está impactando detrimentalmente a los demás usuarios en el uso de la infraestructura el administrador del sistema está facultado para terminarla. Se buscará notificar al usuario interesado, pero la responsabilidad del correcto funcionamiento de las tareas es responsabilidad de quien las programa.

3.1 HT-Condor: Fila de ejecución

Todos los usuarios deben de hacer uso del sistema de colas del cluster, que corre en el nodo central y gobierna el uso de los nodos de cómputo. No se permite el acceso directo a los nodos de cómputo para correr trabajos ahí.

Esta prohibido utilizar el nodo central para correr trabajos sin el uso del administrador de colas.

Los usuarios deben de utilizar el sistema de administración de colas (HT-Condor) para reservar los nodos necesarios para sus trabajos.

La fila de ejecución implementa una política de “atender en el orden en que lleguen” (first-come-first-serve).

3.2 Almacenamiento

La supercomputadora es un recurso de cómputo intensivo. Naturalmente cuenta con espacio para almacenar datos a la entrada, a la salida, y en condiciones intermedias de su procesamiento. Sin embargo ha sido diseñada para optimizar su capacidad de procesamiento, y no para tener una gran capacidad de almacenamiento: el espacio en disco es un recurso limitado.

Por esta razón debe evitarse el uso de la supercomputadora como servidor de almacenamiento a largo plazo. Se requiere que los usuarios sean juiciosos acerca de qué datos mantener en la supercomputadora, i.e. aquellos con los que estén trabajando.

Se prohíbe usar las cuentas de usuario sólo para retener datos.

¡Si el espacio en discos escasea será necesario desarrollar una política de desalojo!

3.3 Ventanas de Mantenimiento

Los usuarios serán notificados por correo electrónico y con anuncios en la mesa de ayuda de cualquier suceso relacionado a la supercomputadora, como ejemplos: actualizaciones de hardware, mantenimientos programados, etc. Se requiere de una comunicación activa para el uso efectivo de este recurso compartido.

4 Mesa de ayuda

Para organizar el soporte técnico en torno al uso de este recurso se brinda una aplicación web en este URL:

<https://github.com/sostenibilidad-unam/supercomputadora>

Cuenta con mecanismo de seguimiento de [issues](#), lo que permite rastrear solicitudes de cambio a la supercomputadora, por ejemplo para corregir desperfectos, instalar nuevo software, etc.

También cuenta con un [wiki](#) para dar soporte a la creación colaborativa de documentación pertinente a la supercomputadora, por ejemplo manuales de uso para las diferentes piezas de software comúnmente ejecutadas.