

Table of Contents

1 Bienvenido a TeideHPC

- 1.1 Descripción del cluster TeideHPC
- 1.2 Compute nodes
- 1.3 Almacenamiento
- 1.4 Red
- 1.5 Conectividad

I Guías de usuario

2 Cómo empezar en TeideHPC

- 2.1 Flujo de trabajo recomendado en TeideHPC.

I.I Guías de transición

3 Guía de transición a Rocky 8.

- 3.1 Sistema Operativo y nodos de login
- 3.2 Software
- 3.3 Slurm
- 3.4 Repositorio público con ejemplos

- I.II Seguridad y VPN

I.II.I Claves GPG

4 ¿Qué es GPG?

- 4.1 GPG para usuarios linux
 - 4.1.1 Generar una par de claves (pública y privada)
 - 4.1.2 Exportar la clave publica
 - 4.1.3 Importar una clave publica
 - 4.1.4 Encriptar archivos con una clave pública importada
 - 4.1.5 Desencriptar archivos con la clave privada (si es de nuestra propiedad)

5 ¿ Qué es GPG ?

- 5.1 GPG para usuarios windows
 - 5.1.1 Instalación de GPG4win
 - 5.1.2 Generar una par de claves (pública y privada)
 - 5.1.3 Exportar la clave pública
 - 5.1.4 Cifrar ficheros con la clave pública.

- 5.1.5 Descifrar ficheros con la clave privada

6 ¿Qué es GPG?

- 6.1 GPG para usuarios de macOS
 - 6.1.1 Generar una par de claves (pública y privada)
 - 6.1.2 Exportar la clave publica
 - 6.1.3 Importar una clave publica
 - 6.1.4 Encriptar archivos con una clave pública importada
 - 6.1.5 Desencriptar archivos con la clave privada (si es de nuestra propiedad)

7 Publicar una clave GPG en un servidor público

I.II.II OpenVPN

8 OpenVPN para usuarios Linux

- 8.1 Conexión mediante terminal
- 8.2 Conexión mediante interfaz gráfica

9 OpenVPN para usuarios Windows

- 9.1 Descarga e instalación
- 9.2 Importar perfil de conexión
- 9.3 Conectar a la VPN
- 9.4 Log de OpenVPN para usuarios Windows.

10 OpenVPN para macOS

- 10.1 Conectarse a la VPN utilizando Tunnelblick

11 Añadir los servidores DNS de TeideHPC a nuestra conexión

- 11.1 Añadir servidores DNS. Usuarios linux.
 - 11.1.1 Mediante línea de comandos.
 - 11.1.2 Mediante interfaz gráfica (Ubuntu y Debian)
- 11.2 Añadir servidores DNS. Usuarios Windows

I.III Primeros pasos

12 ¿Cómo iniciar sesión en TeideHPC?

- 12.1 Acceso para los usuarios de Linux y macOS
 - 12.1.1 SSH Alias
 - 12.1.2 Acceder con clave pública SSH
- 12.2 Acceso para usuarios de Windows
 - 12.2.1 Acceso remoto SSH con PuTTY
 - 12.2.2 MobaXterm

13 Cuentas de usuario

14 Almacenamiento

- 14.1 Copias de seguridad

15 Repositorio con ejemplos

- II Conceptos HPC

- II.1 Slurm

16 Qué es SLURM

- 16.1 Nodos de login
- 16.2 Particiones
- 16.3 Comandos más comunes en SLURM

II.1.1 Gestión de trabajos

17 Ejecutar trabajos en el cluster

- 17.1 Lanzar una ejecución
 - 17.1.1 Ejecución de un trabajo con sesión interactiva
 - 17.1.2 Ejecutar un trabajo en tiempo real
 - 17.1.3 Ejecutar en SLURM mediante un script

18 Ejecutar un trabajo en la cola

- 18.1 Mi primer trabajo con slurm
 - 18.1.1 Como lanzar un job con sbatch
 - 18.1.2 Mi segundo trabajo con slurm

19 Useful Slurm commands

- 19.1 Encontrar información en la cola de trabajo con squeue
 - 19.1.1 Estado de los trabajos
- 19.2 Detener trabajos con scancel
- 19.3 Información de estado con sstat
 - 19.3.1 Formatear la salida del sstat
- 19.4 Analizar trabajos terminados con sacct
 - 19.4.1 Formatear la salida del sacct
- 19.5 Control de trabajos en cola y en ejecución mediante scontrol
 - 19.5.1 Streaming de salida a un archivo de texto
 - 19.5.2 Canalizar la salida a Grep y encontrar líneas que contengan la palabra "Tiempo"

20 Variables de entorno de salida en SLURM

- 20.1 Script para ver las variables de slurm.

21 Investigar un trabajo fallido

- 21.1 Exceder los límites de recursos
 - 21.1.1 Información de error
- 21.2 Errores de Software

22 Límites de memoria Slurm

- 22.1 Cómo estudiar la eficiencia de un job
 - 22.1.1 Memoria máxima por tipo de nodo
- 22.2 ¿Cómo puedo saber qué tan eficiente es mi trabajo?

II.1.II Temas avanzados

23 Ejecuciones secuenciales

- 23.1 Ejemplo de ejecución secuencial de trabajos de 1 core
- 23.2 Ejemplo ejecución secuencial trabajos de 4 cores

24 Ejecución de arrays

- 24.1 Ejecución de un array de 10 trabajos a un nodo
- 24.2 Ejecución de un array de 100 trabajos con sólo 10 trabajos simultáneos

25 Trabajos dependientes

- 25.1 Uso

26 Cómo Trabajar con la partición /local

- 26.1 Ejemplo de ejecución utilizando la partición /local

27 Paralelismo

- 27.1 Paralelismo de memoria compartida (threading)
- 27.2 Paralelismo de memoria distribuida (MPI)
- 27.3 Paralelismo híbrido: DM + SM
- 27.4 OpenMP/Multithreading vs. MPI
 - 27.4.1 Algunos consejos
 - 27.4.2 Ejemplo básico:
 - 27.4.3 Ejemplo básico:
- 27.5 Más ejemplos

28 GPUs en TeideHPC

- 28.1 Modelos de GPU disponibles
- 28.2 Computación GPU en TeideHPC
 - 28.2.1 Cómo solicitar recursos de GPU
 - 28.2.2 Recursos genéricos
 - 28.2.3 Ejemplos de uso de
 - 28.2.4 Scripts de ejemplo para un trabajo de GPU
 - 28.2.5 Más ejemplos

29 Ejemplos de script

- 29.1 Trabajos de uno y varios núcleos
- 29.2 OMP, MPI y Trabajos Híbridos
- 29.3 Arrays de Jobs
- 29.4 Trabajos de GPU

- II.II Transferencia de datos

- 30 Transferencia de datos masiva

- 30.1 SFTP para usuarios Linux
 - 30.1.1 Comandos más usados
 - 30.2 SFTP para usuarios Windows
 - 30.2.1 psftp
 - 30.2.2 Filezilla
 - 30.2.3 WinSCP
 - 30.2.4 Otros
 - 30.3 SFTP para usuario de MAC
 - 30.4 Enlaces de interés

III Software y herramientas

31 Entorno de módulos (Environment Modules)

- 31.1 Lmod
- 31.2 ¿Cómo se organiza el software en TeideHPC? Nomenclatura plana.
 - 31.2.1 ¿Qué es un
 - 31.2.2 Toolchain
 - 31.2.3 Toolchain

32 Software disponible en los clusters TeideHPC y AnagaGPU

- 32.1 Significado de algunos módulos significativos.
- 32.2 Ejemplo de uso de modules

33 Solicitud de instalación de nuevo software

III.I Guías de uso

34 Conda

- 34.1 Utilizar Conda en TeideHPC
- 34.2 Crear un entorno en /data

35 Python

- 35.1 Utilizar Python en TeideHPC
- 35.2 Instalar paquetes de Python en /data

36 R

- 36.1 Utilizar R en TeideHPC
- 36.2 Instalar paquetes de R en /data
- 36.3 Ejecutar un script de R en Slurm

III.I.I Jupyter

37 Jupyter en el Clúster HPC

- 37.0.1 Jupyter Notebook: la interfaz de notebook clásica
- 37.0.2 JupyterLab: una interfaz de notebook de próxima generación
- 37.0.3 JupyterHub

38 Ejecución remota de Jupyter Notebook con Slurm

- 38.0.1 Crear un entorno conda
- 38.0.2 Ejecutar Jupyter Notebook en un nodo de cómputo mediante
- 38.0.3 Running Jupiter Notebook on a Compute Node via
- 38.1 Notas Importantes
- 38.2 Conceptos básicos de Jupyter Notebook
 - 38.2.1 El dashboard Notebook
 - 38.2.2 Ejemplo básico
 - 38.2.3 Ejecutar de un Notebook

39 JupyterLab: La evolución de Jupyter Notebook

- 39.1 Ejecutar JupyterLab de forma remota con Slurm
 - 39.1.1 Crear un entorno Conda para instalar JupyterLab
 - 39.1.2 Usar JupyterLab como módulo
 - 39.1.3 Cómo iniciar JupyterLab
 - 39.1.4 Ejecutar JupyterLab en un nodo de cómputo mediante
 - 39.1.5 Crear túnel SSH
 - 39.1.6 Ejecutar JupyterLab en un nodo de cómputo mediante
- 39.2 Notas importantes
- 39.3 JupyterLab Extensions Manager
 - 39.3.1 Jupyterlab-slurm
 - 39.3.2 Neptune-notebooks
 - 39.3.3 JupyterLab TensorBoard
 - 39.3.4 Jupyter ML-workspace
 - 39.3.5 JupyterLab Debugger
 - 39.3.6 JupyterLab Git

40 JupyterHub

41 Singularity

- 41.1 Utilizar Singularity en TeideHPC
- 41.2 Contenedores Docker
 - 41.2.1 Descargar un contenedor desde DockerHUB
 - 41.2.2 Ejecución de un contenedor
- 41.3 Ejecutar un contenedor en Slurm
 - 41.3.1 MPI
 - 41.3.2 GPU

- 41.4 Otras opciones

IV Infraestructura en la nube (IaaS)

IV.1 OpenNebula

42 Infraestructura como Servicio (IaaS)

- 42.1 OpenNebula
 - 42.1.1 Acceso a OpenNebula
 - 42.1.2 Interfaz de OpenNebula
 - 42.1.3 Máquinas Virtuales
 - 42.1.4 Plantillas
 - 42.1.5 Vista de usuario
 - 42.1.6 Crear Instantáneas

43 Gestión de Máquinas Virtuales

- V Servicios

- V.1 NextCloud

44 Introducción a NextCloud

- 44.1 ¿Qué es NextCloud?
- 44.2 ¿Para que sirve NextCloud?

45 Características de la cuenta

- 45.1 ¿Cómo se accede?
- 45.2 ¿Qué políticas de seguridad hay?
- 45.3 ¿Qué sucede cuando elimino un archivo?
- 45.4 ¿Qué cuota de espacio tengo disponible?

46 Características de la cuenta

- 46.1 ¿Qué tipos de cuentas hay disponibles?
- 46.2 ¿Cómo configurar más de una cuenta en el cliente de Nextcloud?

47 Interfaz web de NextCloud

- 47.1 Iniciando sesión desde la interfaz web
 - 47.1.1 Navegando por la interfaz de usuario principal

48 Compartiendo archivos en NextCloud

- 48.1 ¿Cómo se crean Espacios Compartidos?
- 48.2 Formas de compartir los archivos
 - 48.2.1 Enlaces públicos compartidos
 - 48.2.2 Uso compartido interno con usuarios y grupos

- 48.2.3 Otros con acceso

49 Control de versiones de NextCloud

50 Aplicación de escritorio en NextCloud

- 50.1 Instalación de la aplicación de escritorio
 - 50.1.1 Instalación en Linux
 - 50.1.2 Instalación en Mac OS X y Windows
- 50.2 Uso de la aplicación de escritorio
 - 50.2.1 Iconos de aplicación utilizados
- 50.3 Configurar una cuenta

51 Migración desde el cliente de escritorio de OwnCloud a NextCloud

- 51.1 Pasos a seguir
 - 51.1.1 1. Detener el cliente owncloud
 - 51.1.2 2. Eliminar el cliente owncloud
 - 51.1.3 3. Instalar el cliente nextcloud
 - 51.1.4 4. Inicie el cliente nextcloud
 - 51.1.5 5. Mueve tú fichero owncloud.cfg
 - 51.1.6 6. Inicie el cliente nextcloud

52 Copias de seguridad en Nextcloud

- 52.1 Razones
 - 52.1.1 No es una solución de backup completa
 - 52.1.2 No se pueden hacer copias de seguridad incrementales
 - 52.1.3 Riesgo de corrupción de datos
 - 52.1.4 Limitaciones de almacenamiento

• V.II OnDemand

53 Introducción a Open OnDemand

- 53.1 Características Clave
- 53.2 Ventajas de Open OnDemand
- 53.3 Casos de Uso
- 53.4 Cómo Comenzar

54 Trabajando con Archivos en OnDemand

- 54.1 Acceso a la página de Archivos en OnDemand
- 54.2 Crear una carpeta
- 54.3 Subir un archivo
- 54.4 Crear un archivo
- 54.5 Editar un archivo
- 54.6 Renombrar un archivo
- 54.7 Copiar/Mover un archivo
- 54.8 Eliminar un archivo

55 Ejecución de trabajos en OnDemand

- 55.1 Acceder a la Pestaña de Trabajos
- 55.2 Enviar un Nuevo Trabajo
- 55.3 Monitorear y Gestionar Trabajos Activos
- 55.4 Consultar Trabajos Completados

56 Clústeres en TeideHPC

- 56.1 Cómo Acceder a la Pestaña "Clusters"

57 Aplicaciones Interactivas en OnDemand

- 57.1 Acceso a las Aplicaciones Interactivas
- 57.2 Aplicaciones Disponibles
 - 57.2.1 Code Server
 - 57.2.2 JupyterLab
 - 57.2.3 RStudio
- 57.3 Consejos Generales
- 57.4 Soporte

58 Preguntas frecuentes.

- 58.1 HPC
- 58.2 Opennebula

1 Bienvenido a TeideHPC



Visite nuestra web para encontrar más información sobre la infraestructura <https://teidehpc.iter.es> o envíenos un mail a teidehpc@iter.es.



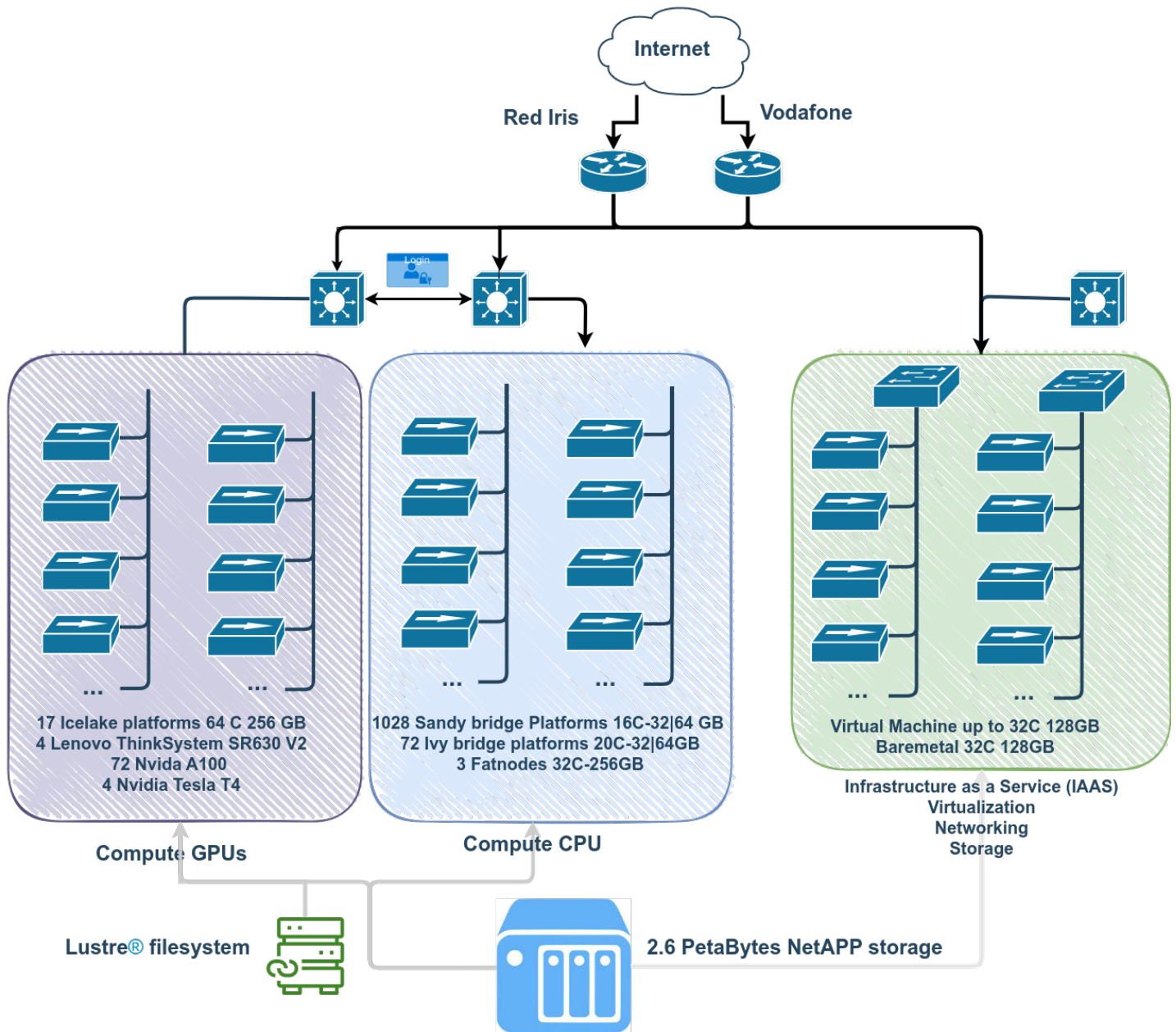
La infraestructura Teide HPC (High Performance Computing) constituye una pieza fundamental del proyecto ALiX para la puesta en marcha de infraestructuras orientadas a la creación de un tejido industrial en torno a las Tecnologías de la Información y la Comunicación (TICs) en Tenerife.

El superordenador Teide, es uno de los más potente de España, ofrece a investigadores, empresas del Parque Tecnológico y Científico de Tenerife, y a la Universidad de La Laguna, un medio de alta capacidad de proceso, para mejorar y ampliar el alcance tanto nacional como internacional de las investigaciones. Además está presente en la lista top500 de los supercomputadores más potentes del mundo ocupando el puesto 138 de la lista de noviembre de 2013.

El superordenador Teide es una infraestructura de computación de altas prestaciones de propósito general. Gestionado en el ITER, el superordenador Teide está alojado en datacenter D-ALiX y provisto de infraestructura eléctrica y de frío de alta disponibilidad, y de conectividad a internet de alta velocidad.

1.1 Descripción del cluster TeideHPC

El siguiente diagrama describe de manera muy sencilla cómo está estructurado el cluster.



1.2 Compute nodes

TeideHPC has these types of computing platforms.

Type	Quantity	Platform	Processors	Cores	Memory	# GPUs
CPU	>500	Sandy bridge	2 X Intel Xeon E5-2670	16	32-64GB	-
CPU	72	Ivy bridge	2 X Intel Xeon E5-2670v2	20	32-64GB	-
CPU	3	Sandy bridge	2 X Intel Xeon E5-4620	32	128-256 GB	-

Type	Quantity	Platform	Processors	Cores	Memory	# GPUs
GPU	16	Icelake	2 X Intel Xeon Gold 6338 32C	64	256 GB	4 Nvidia A100
GPU	1	Icelake	2 X Intel Xeon Gold 6338 32C	64	256 GB	8 Nvidia A100
GPU	4	Icelake	2 X Intel Xeon Gold 6338 32C	64	256 GB	8 Nvidia T4

1.3 Almacenamiento

- Almacenamiento **NetApp** con capacidad de **2.6 Peta Bytes**, configurada en formato clúster contando con todos los elementos redundados para hacer frente a posibles fallos de hardware, con discos de spare según las best practices, siendo éstos globales.
- Almacenamiento paralelo **Lustre** para aplicaciones que requieran un alto número de operaciones de E/S.

1.4 Red

Teide-HPC dispone de una topología de red donde se definen cuatro redes de propósito específico:

- Red dedicada de almacenamiento.
- Red dedicada de gestión.
- Red out of band.
- Red de baja latencia Infiniband EDR a 100Gbps para cómputo.

Como medidas de seguridad TeideHPC dispone de túneles IPSec, conexiones VPN y la posibilidad de establecer VLANs privadas para sus clientes.

1.5 Conectividad

TeideHPC se conecta a internet a través de la red académica y de investigación española, **RedIris**, mediante un enlace de 10 Gb. También dispone de conectividad a través del proyecto Alix mediante un operador de internet nacional.

Para realizar **transferencias de datos**, se dispone de nodos de transferencia que permiten copiar grandes cantidades de datos al espacio de usuario accediendo directamente a la red troncal de datos.

I. Guías de usuario

2 Cómo empezar en TeideHPC

Esta página lo guiará brevemente a la documentación. Si sigue los pasos a continuación, debería tener una buena idea sobre los aspectos más importantes del clúster TeideHPC.

1. Consigue una **cuenta TeideHPC**.

Si tu entidad tiene firmado un convenio con nosotros, obtén una cuenta TeideHPC enviando un correo electrónico a soporte@hpc.iter.es informándonos quién es el responsable del departamento o entidad a la que perteneces.

Si eres un investigador o una empresa interesada en utilizar nuestros servicios, pregúntanos.

2. Envíanos una **clave pública GPG**.

Recibirá sus credenciales de VPN cifradas con su clave pública GPG. Sólo necesitará descifrarlas y **conectarse** a nuestra VPN.

3. **Cambie su contraseña de usuario**

Después de conectarte por primera vez puede cambiar su contraseña. Ésta tiene una serie de políticas de seguridad que deberá cumplir así como una fecha de caducidad.

4. **Antes de comenzar cualquier cálculo serio, eche un vistazo a las **opciones de almacenamiento**.**

¡Tu directorio **home** en TeideHPC tiene un límite flexible de 5GB! **Recomendamos guardar sus datos en el directorio de datos o lustre**

5. **Descubra la sección de comandos útiles y consejos de uso**

Hay bastantes **comandos útiles** y **consejos de uso** que pueden facilitar su vida en el clúster.

6. Antes de enviar un trabajo, recomendamos documentarse sobre los siguientes aspectos":

- [Descripción del cluster](#)
- [Gestor de trabajos Slurm](#)
- [Particiones](#) todas las particiones tienen límites de tiempo diferente.
- [Cómo se carga el software](#)
- [Sesiones interactivas](#)
- [Cómo usar las GPU](#)
- [Ejecuciones multinodo \(MPI\)](#)

7. Familiarizarse con los *trabajos en batch*

Eche un vistazo a la sección [Cómo ejecutar trabajos en batch](#)

8. ¡No malgastes!. Usa los nodos y las GPU eficientemente.

Estudia la [eficiencia de tus trabajos con slurm](#) o indícanos el número de trabajo, hora de inicio y finalización, así como el nombre de los nodos sobre los que se ejecutó y te enviaremos unas gráficas donde podrás ver el rendimiento del mismo.

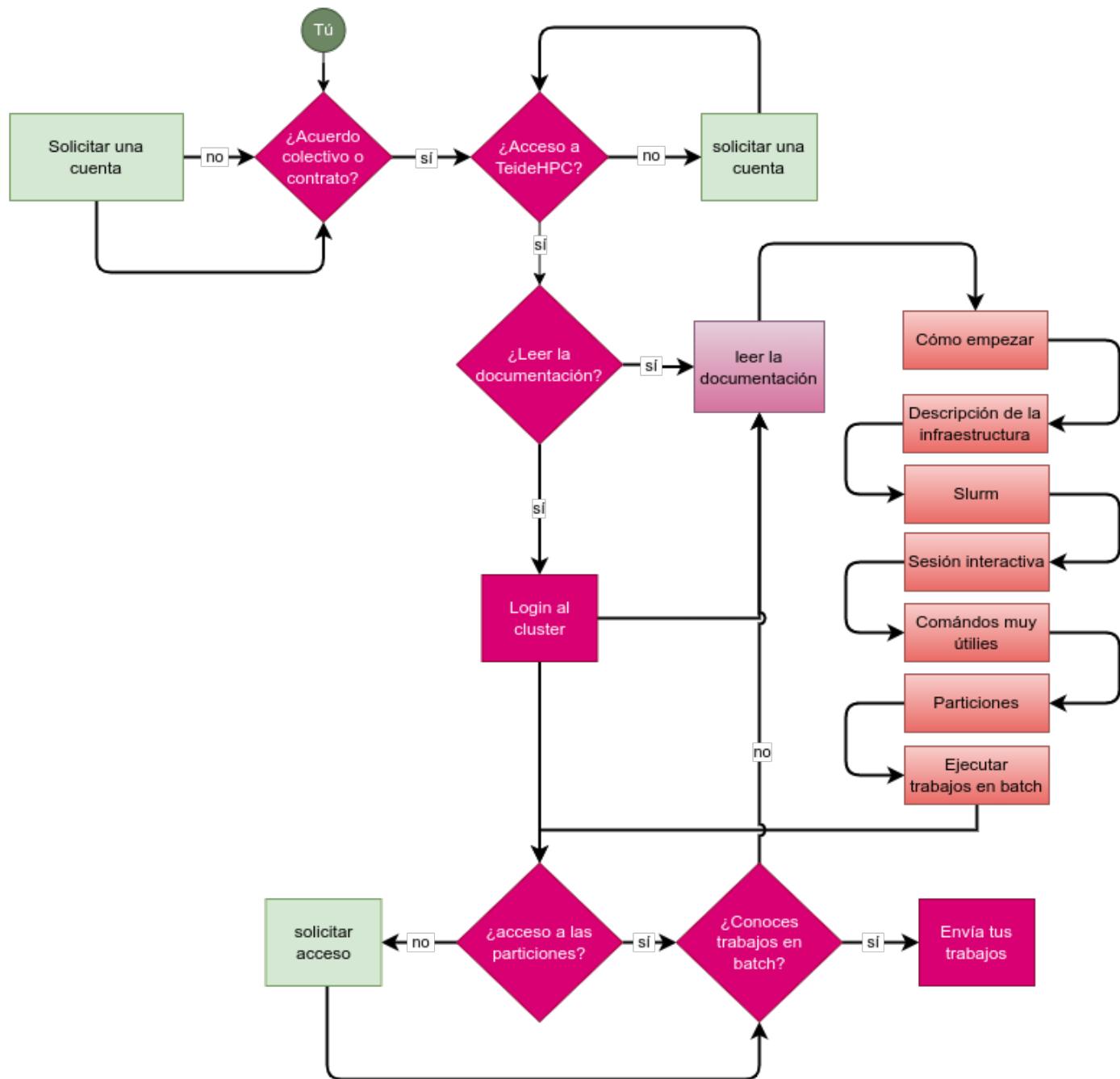
9. ¿Necesita Computación en la nube, IaaS (Infraestructura como servicio), virtualización, almacenamiento, red privada o servidores?

Esríbenos a support@hpc.iter.es

10. Cualquier otra cuestión, ayuda o asesoramiento técnico?

Esríbenos a support@hpc.iter.es

2.1 Flujo de trabajo recomendado en TeideHPC.



I.I Guías de transición

3 Guía de transición a Rocky 8.

Con motivo de la llegada de las GPU al centro de supercomputación TeideHPC se ha introducido una versión nueva del sistema operativo que usan tanto los nodos de cómputo, GPU y nodos de login. Todo el cluster relacionado con Centos 6 y Centos 7 estarán al final de vida en pocos meses.

Por otra parte, se ha creado un nuevo cluster llamado **AnagaGPU** y se han realizado algunos cambios en el cluster TeideHPC, así como en el software, por lo que si ya ha ejecutado trabajos en TeideHPC deberá realizar ajustes en su flujo de trabajo.

Resumiendo, estos son los cambios más significativos a nivel de Sistema Operativo, acceso, software, slurm.

3.1 Sistema Operativo y nodos de login

- Cada cluster (TeideHPC y AnagaGPU) tiene su propia IP de acceso.
- El sistema operativo de ambos cluster y los nuevos nodos es Rocky 8.
- Existen **4 nuevos nodos de login** dispuestos en alta disponibilidad (HA) mediante 2 IPs de acceso.
- La asignación de nodo de login durante el acceso es aleatoria y va en función del número de usuarios.

3.2 Software

- El cambio en el sistema operativo significa que la mayoría del software de los usuarios basado en Centos 6 o CentOS 7 no funcionará y ha de ser recompilado.
- Se deja de usar la herramienta de modules *TCL* (Centos 6) en virtud de **Lmod**.
- El software instalado pasa a estar organizado mediante una **nomenclatura plana**
- Cada tipo de nodos tiene instalado y compilado el software específico para cada arquitectura de nodos. Esto quiere decir:

El software instalado depende de la arquitectura de los nodos

- Básicamente hay **2** arquitecturas: **icelake** (nodos con GPUs) y **sandybrige**(nodos de CPU).
- Mira la **descripción del cluster** en la página principal así cómo la página "**Cómo solicitar recursos de GPU y cómputo**".

Cada cluster tiene su propio software

Para ver el software disponible en cada cluster debe ingresar a través de las Ip de acceso de cada cluster.

Existen módulos que no dependen de la arquitectura

```
----- /share/easybuild/software/common/modules/all -----
EasyBuild/4.7.0          Go/1.18.3             Miniconda3/22.11.1-1   Singularity/3.11.0     slurm/
teide
EasyBuild/4.8.2 (L,D)    Mamba/4.14.0-0        Miniconda3/23.5.2-0   (D)                     Squashfs/4.3
```

3.3 Slurm

- La asignación de nodos pasa de ser **Modo NO compartido a nodos compartidos**.

Esto quiere decir que, al solicitar simplemente 1 nodo de cómputo, no se solicita el nodo completo para el usuario, por lo que se obliga al usuario a realizar una reserva completa de recursos si es lo que se desea.

- Los parámetros por defecto que asigna slurm son:

```
#SBATCH --node=1
#SBATCH --ntask=1
#SBATCH --ntask-per-node=1
#SBATCH --cpu-per-task=1
#SBATCH --mem=2GB
```

- Se ha introducido el uso de una partición nueva para la **solicitud de recursos de GPUs**.
- Recomendamos encarecidamente que para ejecutar aplicaciones se use el comando `srun tu_aplicacion`. **Aquí** puedes ver una explicación sencilla de qué implicaciones puede tener usarlo o no usarlo.
- Puedes estudiar la **eficiencia de tus trabajos** completados con un simple comando.

3.4 Repositorio público con ejemplos

Para facilitar el inicio y acceso a la computación HPC en TeideHPC, hemos creado un repositorio público en github donde iremos publicando ejemplos de uso de aplicaciones.

Te animamos a colaborar en él. https://github.com/hpciter/user_codes

I.II Seguridad y VPN

I.II.I Claves GPG

4 ¿Qué es GPG?

GNU Privacy Guard (GnuPG o GPG) es una herramienta de cifrado y firmas digitales que implementa el **estándar OpenPGP**. GnuPG permite encriptar y firmar tanto datos como comunicaciones (emails por ejemplo). Cuenta con un sistema de gestión de claves versátil, junto con módulos de acceso para todo tipo de directorios de claves públicas. También conocido como GPG, es una herramienta de línea de comandos con funciones para una fácil integración con otras aplicaciones.

GPG cifra los mensajes usando pares de claves individuales asimétricas generadas por los usuarios. Las claves públicas pueden ser compartidas con otros usuarios de muchas maneras, un ejemplo de ello es depositándolas en los servidores de claves.

NOTA: NUNCA COMPARTA SU CLAVE PRIVADA, GUÁRDELA A BUEN RECAUDO Y SÓLO REMITA O PUBLIQUE EN UN SERVIDOR DE CLAVES LA PÚBLICA.

4.1 GPG para usuarios linux

Todas las distribuciones linux incorporan la aplicación por defecto.

4.1.1 Generar una par de claves (pública y privada)

```
gpg --gen-key
```

Se debe contestar a una serie de preguntas en orden para generar la clave pública y privada:

- Key type. RSA-RSA por defecto.
- Expiration time: Tiempo de validez de la clave.
- User data (Real name, email, comment) . Estos datos son usados para identificar la clave en servidores de claves públicas y/o cuando se le envían a otros usuarios
- Password: Contraseña para proteger la clave privada.

4.1.2 Exportar la clave publica

```
gpg --armor --output myname_public_key.asc --export 'Name Surname'
```

4.1.3 Importar una clave publica

```
gpg --import johndoe.asc
```

Listar las claves públicas importadas con el siguiente comando:

```
gpg --list-keys
```

4.1.4 Encriptar archivos con una clave pública importada

```
gpg --encrypt --recipient 'Name Surname' file.txt
```

4.1.5 Desencriptar archivos con la clave privada (si es de nuestra propiedad)

```
gpg --output file.txt --decrypt file.txt.gpg
```

4.1.5.1 Otros links utiles sobre gpg

- <https://www.madboa.com/geek/gpg-quickstart/>

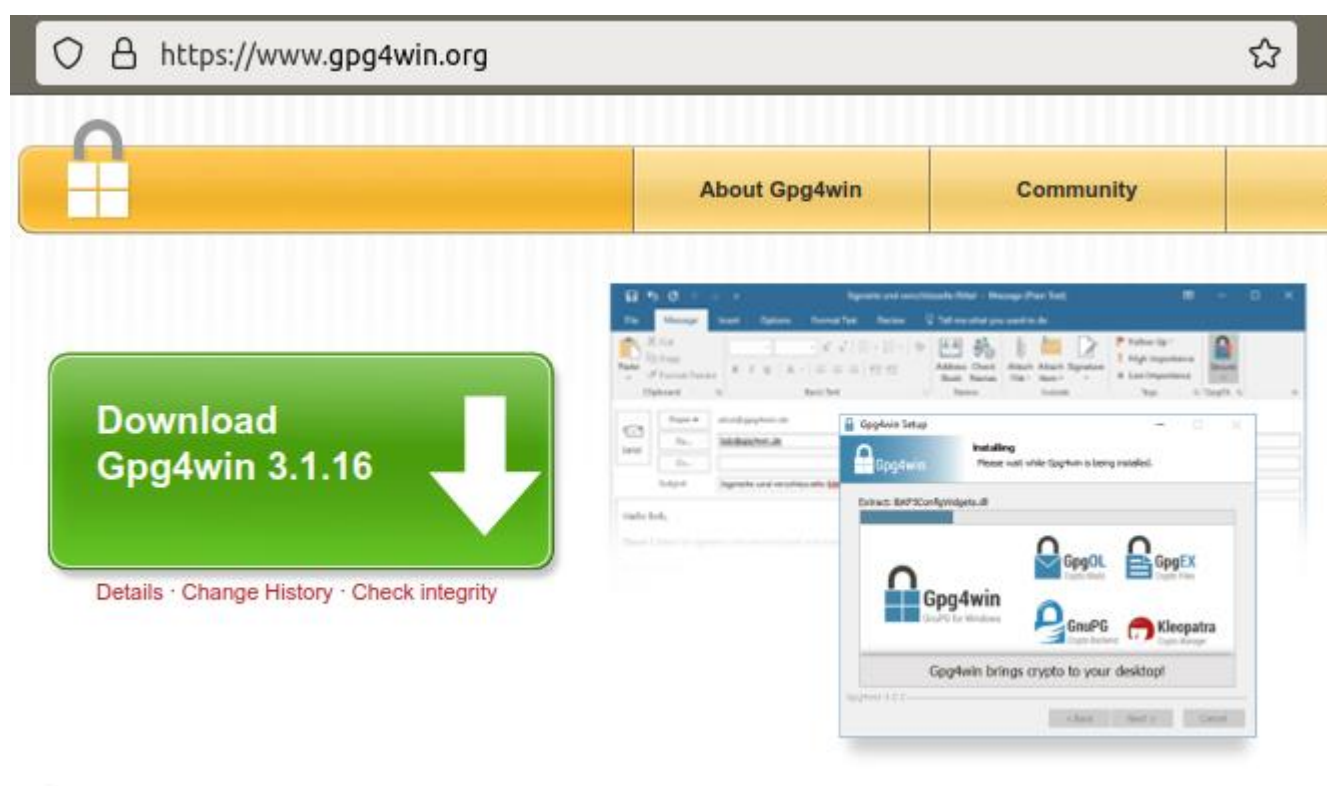
5 ¿ Qué es GPG ?

Ver descripción [aquí](#)

5.1 GPG para usuarios windows

Existen varias aplicaciones bajo el sistema operativo Windows que permiten generar un par de claves gpg. Entre las más conocidas está **Gnu4win** y su módulo [Kleopatra]. La aplicación Gpg4win es un software de encriptación de libre distribución el cual permite el cifrado de ficheros y el envío de documentos a través del correo electrónico utilizando criptografía de clave pública para el cifrado de datos y firmas digitales.

Soporta los estándares de criptografía OpenPGP y S/MIME (X.509).

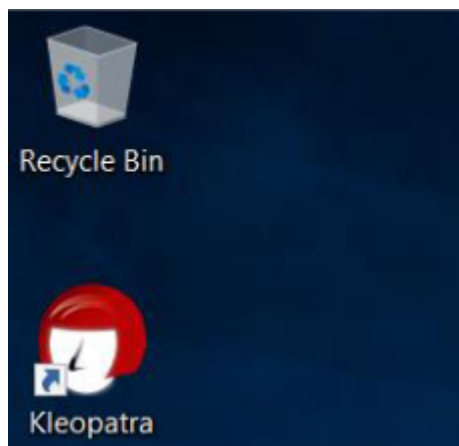
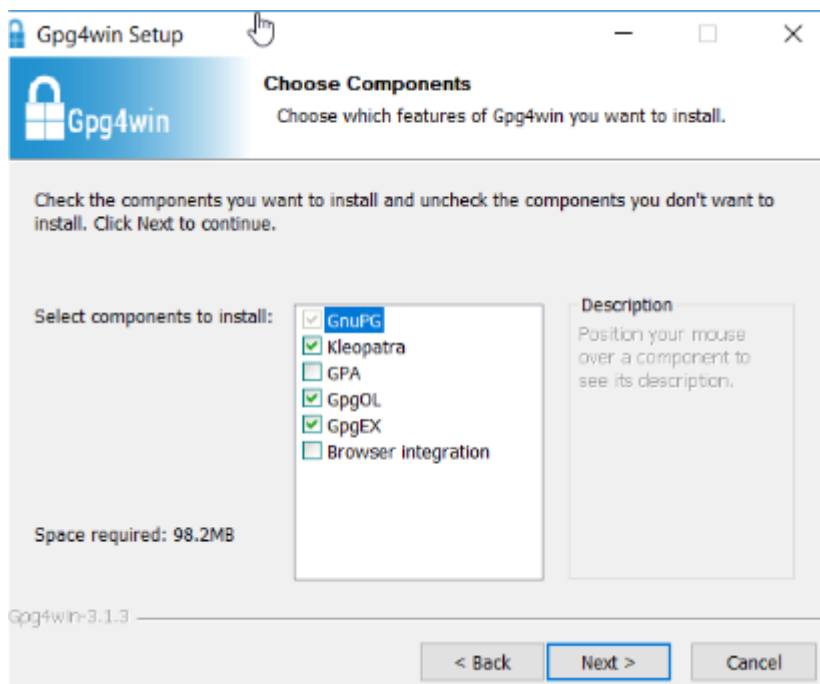


Gpg4win se compone varios módulos, entre ellos:

- GnuPG : la herramienta de cifrado básico
- Kleopatra : administrador de certificados para OpenPGP y X.509
- GPA : un administrador de certificados alternativa (GNU) para OpenPGP y X.509

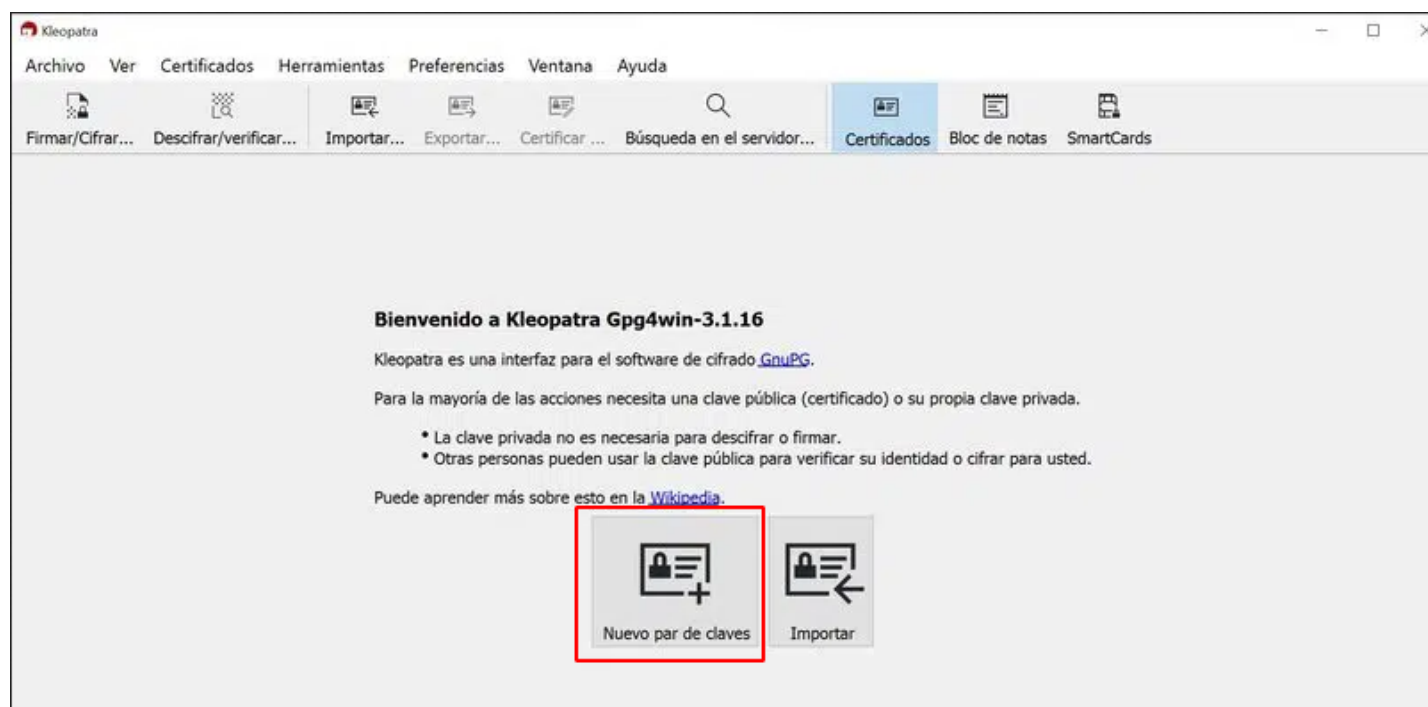
5.1.1 Instalación de GPG4win

En la web del proyecto descargar la versión más reciente del software. Abrir el instalador y autorizar los cambios en caso necesario. Seguir los pasos indicados con el botón Siguiente.



5.1.2 Generar una par de claves (pública y privada)

Una vez instalado Gpg4win vamos a utilizar el módulo **Kleopatra** para generar un nuevo par de claves gpg. Para ello desde la pestaña *Archivo* seleccionamos *Nuevo Certificado*, lo cual abrirá el asistente de generación de certificados.



Pulsa en *Nuevo par de claves GPG*

← Asistente de creación del par de claves

Elegir formato

Por favor, elija qué tipo quiere crear.

- **Crear un par de claves personales OpenPGP**
Los pares de claves OpenPGP están certificados por la confirmación de la huella digital de la clave pública.

- **Crear un par de claves personales X.509 y una petición de certificación**
Los pares de claves X.509 se certifican por una autoridad de certificación (CA). La petición generada necesita enviarse a la CA para finalizar la creación.

Next

Cancel

Introducir los datos requeridos

? ×

← Asistente de creación del par de claves

Introduzca detalles

Por favor, introduzca sus detalles personales debajo. Si desea más control sobre los parámetros, pulse el botón «Configuración avanzada».

Nombre: (opcional)

Correo: (opcional)

Proteger la clave generada con una frase de contraseña.

Julián G.

[Configuración avanzada...](#)

[Crear](#) [Cancel](#)

En la pestaña *Avanzadas* indicar una longitud de clave de 4096 bits y una fecha de caducidad en el desplegable inferior no superior a 2 años.

? ×

Advanced Settings

Technical Details

Key Material

RSA

+ RSA

DSA

+ Elgamal

Certificate Usage

Signing Certification

Encryption Authentication

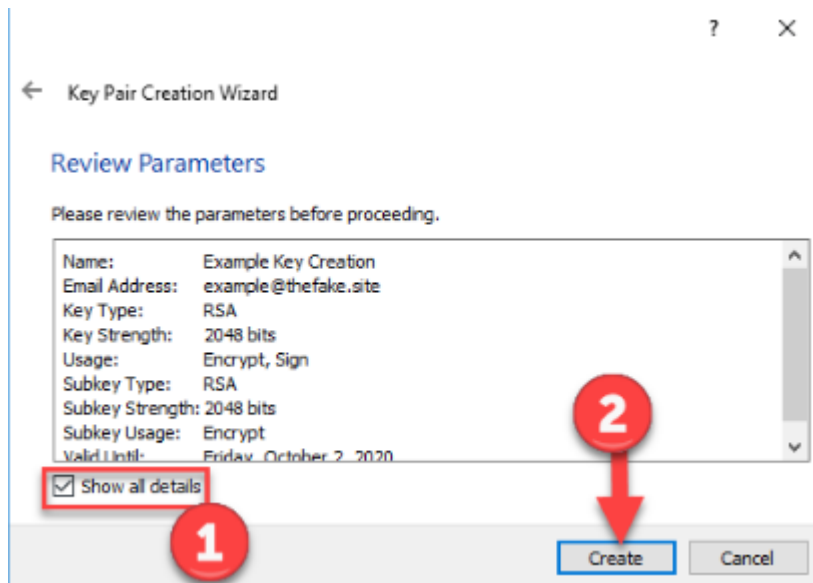
Valid until:

December 2017						
Mon	Tue	Wed	Thu	Fri	Sat	Sun
27	28	29	30	1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31
1	2	3	4	5	6	7

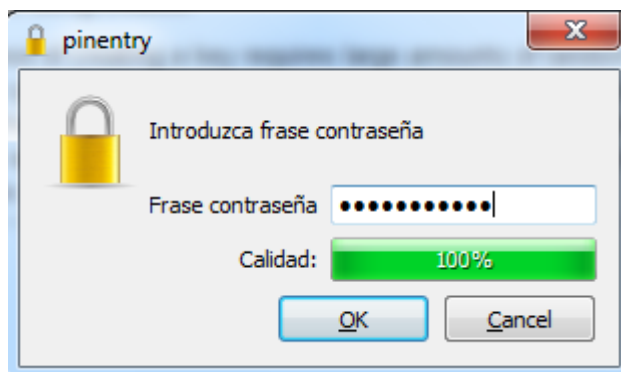
Next Year

Next Month

Confirmar los valores introducidos y generar las claves pulsando en *Crear*.



Introducir una contraseña que recuerde para su clave privada. Esta contraseña será la que deba usar para descifrar los archivos que reciba cifrados con su clave pública.



Al pulsar **OK** verá que su par de claves pública/privadas se han generado correctamente

← Asistente de creación del par de claves

Par de claves creado correctamente

Su nuevo par de claves se ha creado correctamente. Consulte los detalles sobre el resultado y algunos pasos a seguir sugeridos más abajo.

Resultado

Par de claves creado correctamente.
Huella digital: 0A97759CC84026BD56749920AB4FBBA7E0C0E342

Siguientes pasos

Hacer copia de respaldo de su par de claves...

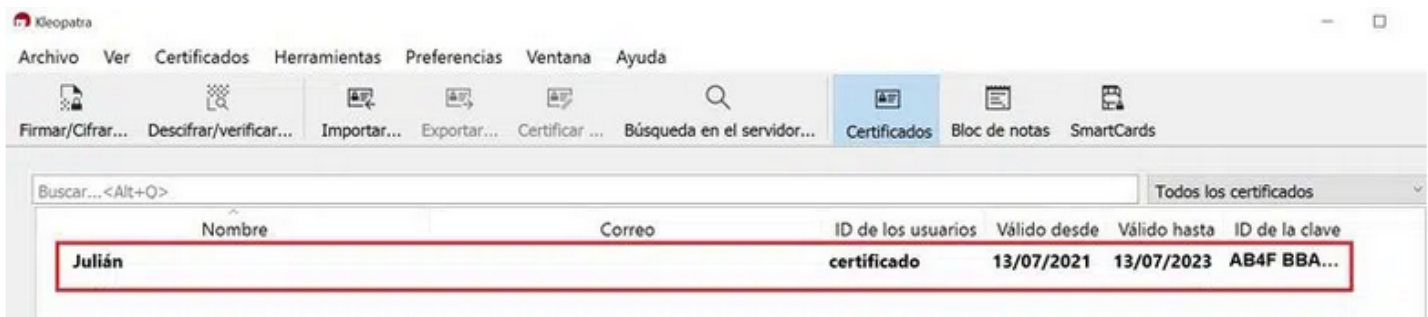
Enviar clave pública por correo...

Enviar clave pública a un servicio de directorio...

Finish

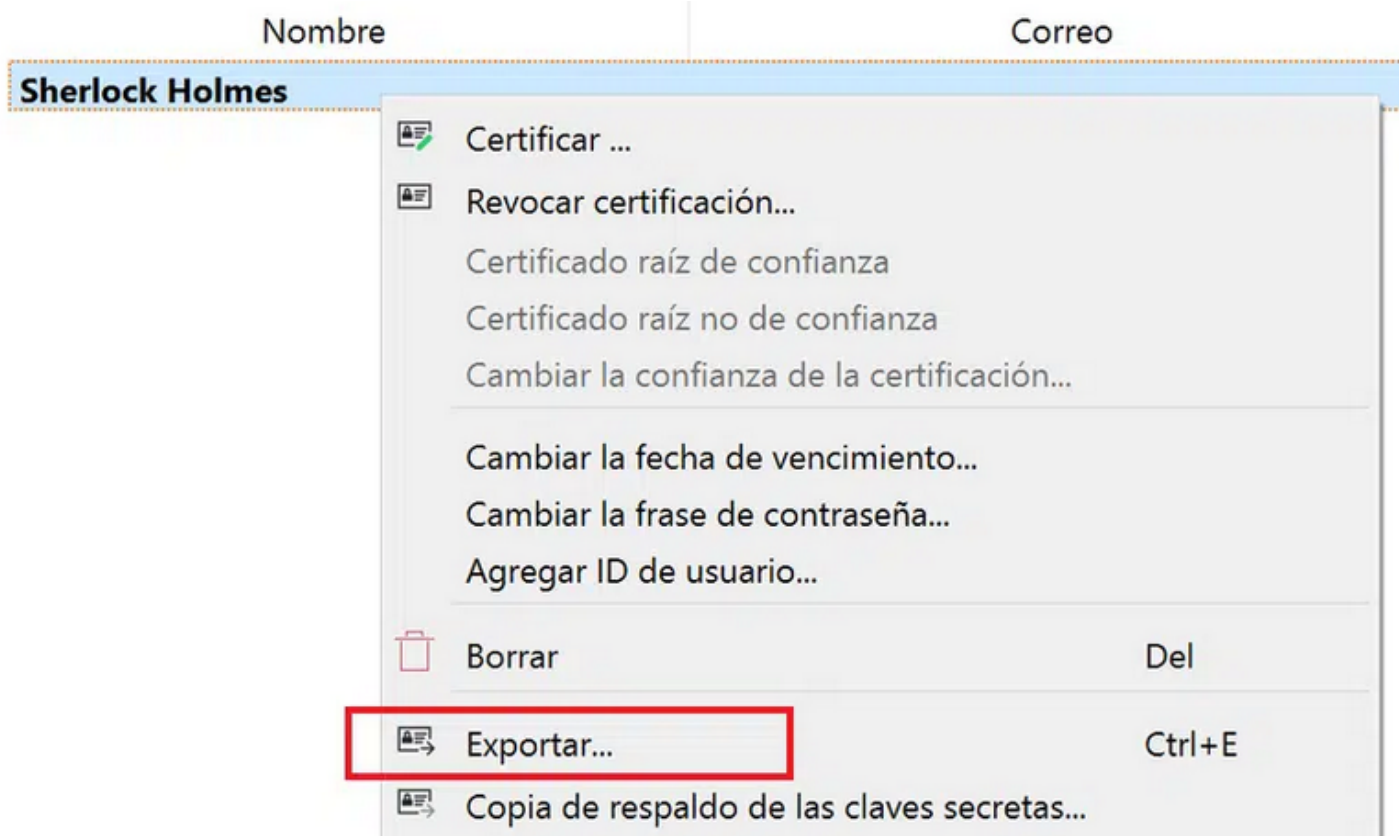
Cancel

Se puede comprobar que se han creado con éxito el par de claves en la pestaña *Certificados*.

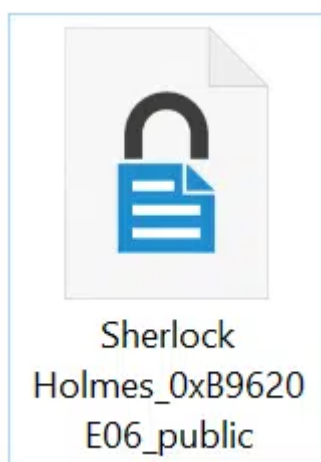


5.1.3 Exportar la clave pública

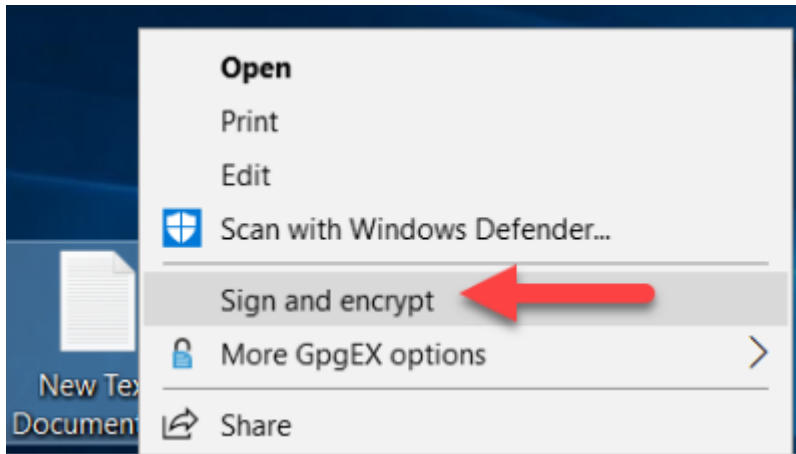
Al exportar la clave pública se genera un fichero de texto que por norma general se almacena con la extensión `.asc`, que es el que podrá distribuir por correo electrónico o bien [publicar en un servidor de claves públicas](#) como *REDIRIS* y donde cualquier persona la puede descargar.



Obtendrá un fichero de este tipo:



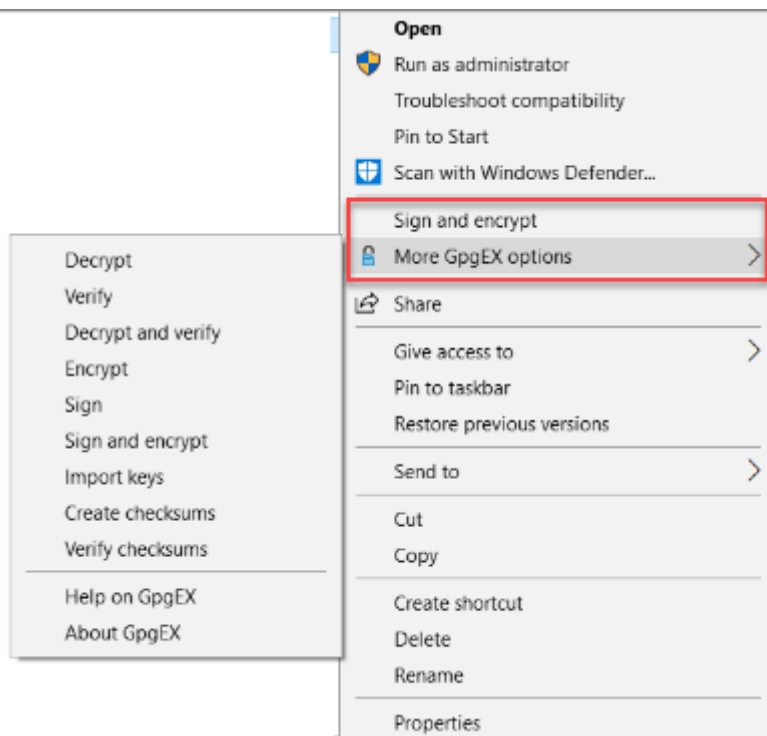
5.1.4 Cifrar ficheros con la clave pública.



5.1.5 Descifrar ficheros con la clave privada

Sólo si un archivo `.gpg` ha sido cifrado con su clave pública y tiene la clave privada, así como su contraseña, podrá descifrar su contenido.

Se puede usar el menú contextual *Descifrar y verificar* para recuperar el fichero original pulsando con el segundo botón. También es posible realizar esto desde Kleopatra.



Una vez descifrado este tendrá el mismo nombre sin la extensión `gpg`.

Todas las operaciones terminadas.



TOP SECRET.docx (1).gpg → TOP SECRET.docx (1):

[Mostrar registro de auditoría](#)

Firma válida por [Dr. Watson](#)

Firma creada en viernes, 16 de julio de 2021 22:01:36

Con certificado:

[Dr. Watson \(9111 B0CB F48D C6DD\)](#)

La firma es válida y la validez del certificado es totalmente confiable.

NOTA IMPORTANTE: Si cambia de ordenador o usa uno diferente asegúrese de tener su par de claves a buen recaudo. Sin ellas no tendrá la posibilidad de descifrar el contenido del archivo. Por ello recomendamos realizar un backup del par de claves GPG

6 ¿Qué es GPG?

GNU Privacy Guard (GnuPG o GPG) es una herramienta de cifrado y firmas digitales que implementa el **estándar OpenPGP**. GnuPG permite encriptar y firmar tanto datos como comunicaciones (emails por ejemplo). Cuenta con un sistema de gestión de claves versátil, junto con módulos de acceso para todo tipo de directorios de claves públicas. También conocido como GPG, es una herramienta de línea de comandos con funciones para una fácil integración con otras aplicaciones.

GPG cifra los mensajes usando pares de claves individuales asimétricas generadas por los usuarios. Las claves públicas pueden ser compartidas con otros usuarios de muchas maneras, un ejemplo de ello es depositándolas en los servidores de claves.

NOTA: NUNCA COMPARTA SU CLAVE PRIVADA, GUÁRDELA A BUEN RECAUDO Y SÓLO REMITA O PUBLIQUE EN UN SERVIDOR DE CLAVES LA PÚBLICA.

6.1 GPG para usuarios de macOS

A la hora de instalar gpg en macOS tenemos dos formas:

- Instalar la **GPG Suite**
- Instalar GPG utilizando el **gestor de paquetes Homebrew**

Recomendamos instalar `gpg` a través de Homebrew y utilizarlo a través de la terminal, como por ejemplo, iTerm.

6.1.1 Generar una par de claves (pública y privada)

```
gpg --gen-key
```

Se debe contestar a una serie de preguntas en orden para generar la clave pública y privada:

- Key type. RSA-RSA por defecto.
- Expiration time: Tiempo de validez de la clave.
- User data (Real name, email, comment) . Estos datos son usados para identificar la clave en servidores de claves públicas y/o cuando se le envían a otros usuarios
- Password: Contraseña para proteger la clave privada.

6.1.2 Exportar la clave publica

```
gpg --armor --output myname_public_key.asc --export 'Name Surname'
```

6.1.3 Importar una clave publica

```
gpg --import johndoe.asc
```

Listar las claves públicas importadas con el siguiente comando:

```
gpg --list
```

6.1.4 Encriptar archivos con una clave pública importada

```
gpg --encrypt --recipient 'Name Surname' file.txt
```

6.1.5 Desencriptar archivos con la clave privada (si es de nuestra propiedad)

```
gpg --output file.txt --decrypt file.txt.gpg
```

6.1.5.1 Otros links utiles sobre gpg

- <https://www.madboa.com/geek/gpg-quickstart/>

7 Publicar una clave GPG en un servidor público

Como se comentaba en apartados anteriores es posible publicar la **clave pública** en servidores de claves GPG que facilitan el acceso a ellas.

Entre los más conocidos está el servidor de la **Rediris**.

Este servidor permite tanto consultar claves públicas gpg como publicar las nuestras, así como revocarlas. Estas claves estarán vigentes durante el tiempo que haya dispuesto al generarlas.


◀ [Servicios](#) ▶ [Seguridad](#) ▶ [Servidor de claves públicas PGP](#)

Servidor de claves públicas PGP



[Consultar](#) · [Enviar clave](#) · [Borrar clave](#) · [Estadísticas de uso](#) · [Estadísticas de claves](#) · [Grafo de sincronización](#) · [Documentación](#)

Para facilitar el acceso a las claves PGP, RedIRIS dispone de un servidor de claves de uso público, situado en el equipo **pgp.rediris.es**, es posible emplear el formulario que aparece más abajo para consultar y/o actualizar la base de datos de claves PGP disponibles en este servidor.

Formulario web original en el que nos basamos: [Brian A. LaMacchia](#).
Software del servidor de claves: [Marc Horowitz](#) , ([Página Personal](#))

[[Consultar](#)] [[Enviar clave](#)] [[Eliminar clave](#)] [[Estadísticas](#)] [[Estadísticas por dominios](#)]
[[Documentación](#)]

- [Consultar una clave en el servidor](#)

La base de datos está indexada por palabras. Por ejemplo, si el identificador de una clave es Pepito Perez <pepitop@rediris.es>, esa clave estará indexada bajo las palabras 'Pepito', 'Perez', 'pepitop', 'rediris' y 'es'. A efectos de búsqueda es equivalente a Pepito Rediris <perez@pepito.es>, a menos que solicite la búsqueda *exacta*.

También puede especificar un *KeyID* hexadecimal, como 0x1234ABCD. El uso de un *KeyID* implica una búsqueda exacta.

Cadena de búsqueda:

- Listar sólo claves. Mostrar huellas digitales.
 Incluir también sus certificados. Sólo búsquedas exactas.

- [Enviar una clave al servidor](#)

AVISO IMPORTANTE: Tenga en cuenta que una vez que envíe una clave a un servidor, ésta será distribuida por todo el mundo en un plazo relativamente corto. Por tanto, **no realice esta acción a la ligera.**

La **ÚNICA** forma de anular una clave pública del servidor, es mediante un certificado de revocación de clave, lo que requiere acceso a su clave privada.

Pulse [aquí](#) si quiere generarlo ahora. Esto le permitirá revocar su clave pública más adelante incluso si pierde la privada u olvida su password.

Introduzca clave PGP en ASCII aquí:

- [Eliminar una clave del servidor](#)

I.II.II OpenVPN

8 OpenVPN para usuarios Linux

Con el fin de proporcionar un acceso seguro a la infraestructura de TeideHPC todas las comunicaciones serán mediante una red privada virtual (VPN).

Los usuarios recibirán en su momento un email con los ficheros de configuración y las credenciales de acceso cifrados con su clave pública GPG.

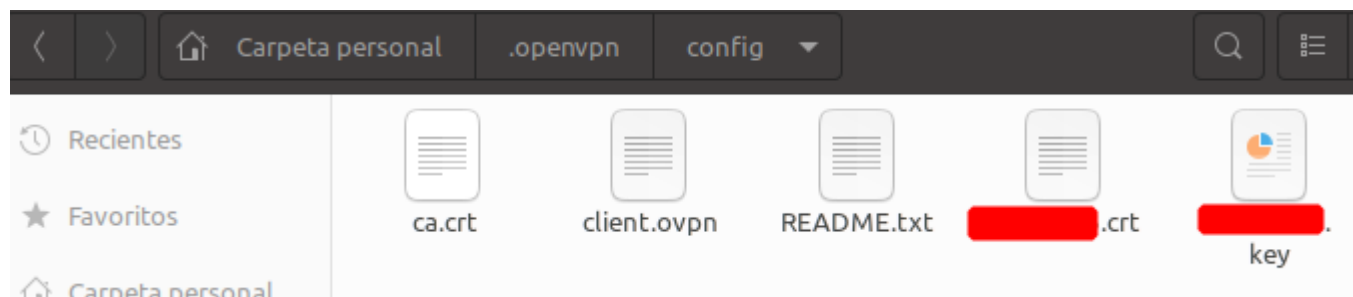
El contenido del mismo es el siguiente:

- Cuatro ficheros que contienen los certificados del cliente.
- Un fichero client.ovpn que contiene los datos de conexión
- Un fichero README.txt con las credenciales de acceso.

Recordamos que las credenciales VPN tienen carácter personal e intransferible y que solo se permite una única conexión simultánea por VPN.

La mayoría de distribuciones actuales de linux traen instalado un cliente de openVPN por defecto. Éste es posible usarlo mediante terminal o bien configurarlo y usarlo mediante interfaz gráfica.

Cómo primer paso **descifre** y posteriormente descomprima el fichero gpg que se le adjuntó y colóquelo en una carpeta dentro de su `/home`.



8.1 Conexión mediante terminal

Para conectarse a la VPN sólo tiene que ir a la ruta `~$home/.openvpn/su_configuración/` y ejecutar:

```
sudo openvpn client.ovpn
```

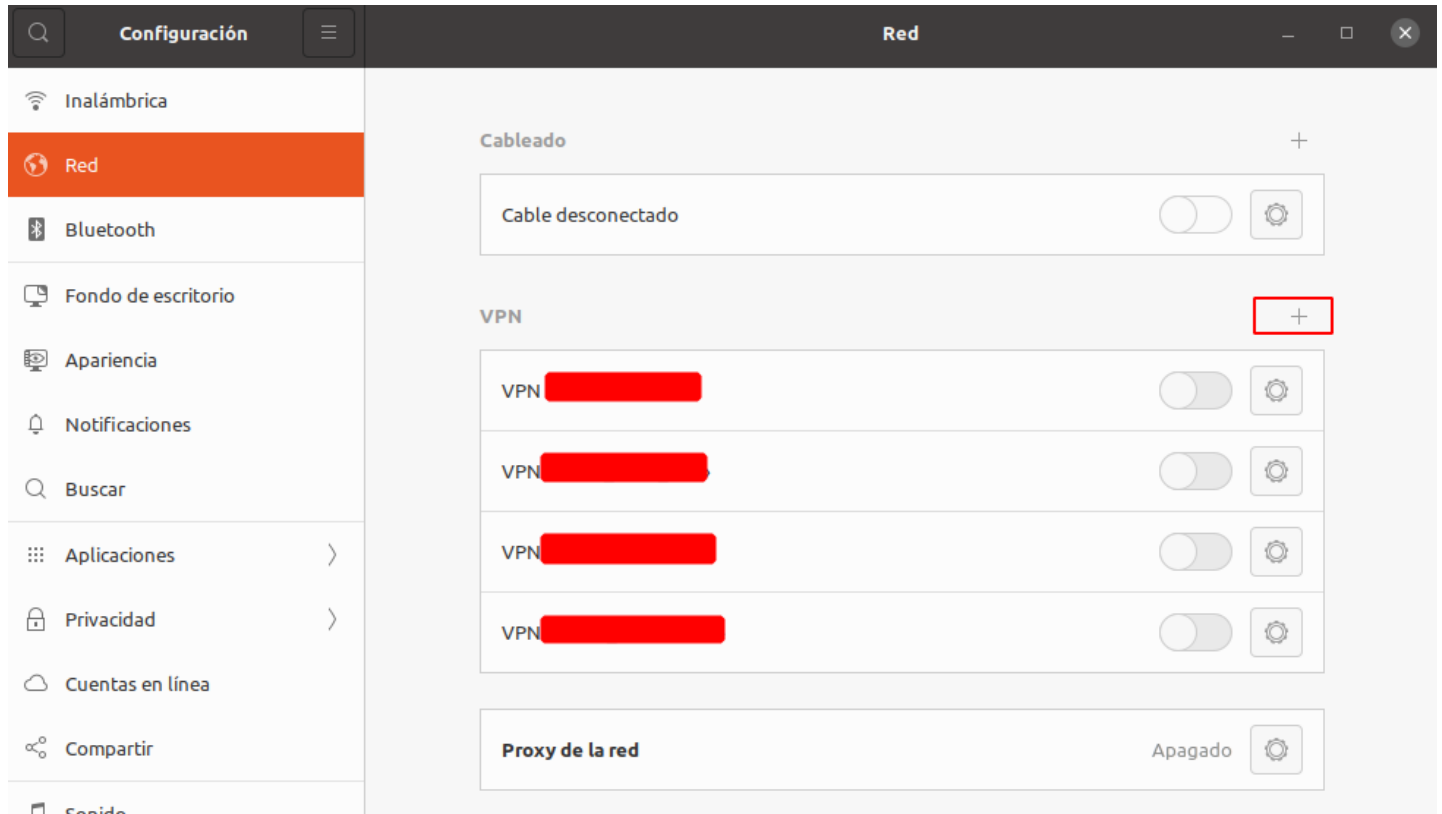
8.2 Conexión mediante interfaz gráfica

Mostramos a continuación una pequeña guía para la distribución Ubuntu 20.04.

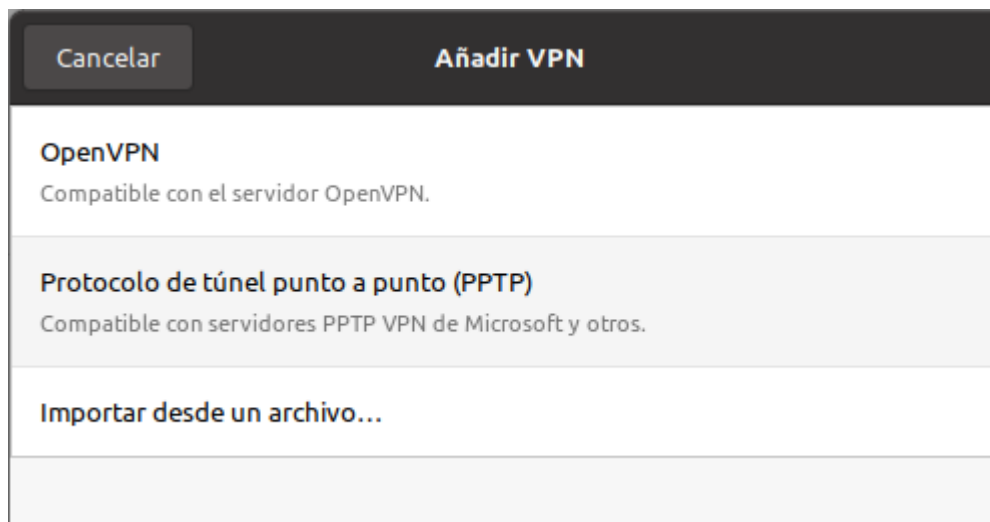
Abra el gestor de redes mediante el acceso en la barra de escritorio o bien teclear en el menú de aplicaciones *Configuración*.



Abrimos la configuración de *Red* y agregamos una nueva conexión VPN:



Importamos la configuración desde un archivo



Introducimos nuestro usuario y contraseña

Cancelar **Añadir VPN** Añadir

Identidad IPv4 IPv6

Nombre

General

Pasarela

Autenticación

Tipo

Nombre de usuario

Contraseña

certificado CA

certificado Usuario

Usuario clave privada

contraseña de clave Usuario

Mostrar la contraseña

⚙️ Avanzado...

Guardamos la configuración y ya podemos probar a conectarnos a la VPN de TeideHPC.

9 OpenVPN para usuarios Windows

Con el fin de proporcionar un acceso seguro a la infraestructura de TeideHPC todas las comunicaciones serán mediante una red privada virtual (VPN).

Los usuarios recibirán en su momento un email con los ficheros de configuración y las credenciales de acceso cifrados con su clave pública GPG.

El contenido del mismo es el siguiente:

- 4 ficheros que contienen los certificados del cliente.
- 1 fichero client.ovpn que contiene los datos de conexión
- 1 fichero README.txt con las credenciales de acceso.



Recordamos que las claves y ficheros de acceso a la VPN y el usuario y contraseña son personales e intransferibles y sólo permite una conexión simultánea.

9.1 Descarga e instalación

Aunque Windows 10 y Windows 11 dispone de un cliente VPN propio que podría usarse, recomendamos OpenVPN en su versión 3.

Descargar el cliente OpenVPN en el siguiente [link](#).

[Download OpenVPN Connect v3](#)

sha256 signature: c001c26f0141d66409f91aca069b402f73913556ecc39156d0522922216809d7

For Windows 7, 8, 10, and 11.

A 32 bits version is also available:

[Download OpenVPN Connect v3 for 32 bits](#)

sha256 signature: c4cedee5eebc3c7d329b1efd071ac544feac50e4e4e94d9755f171cbd8b6d4d9

Previous generation OpenVPN Connect V2 is available here:

[Download OpenVPN Connect v2.7.1](#)

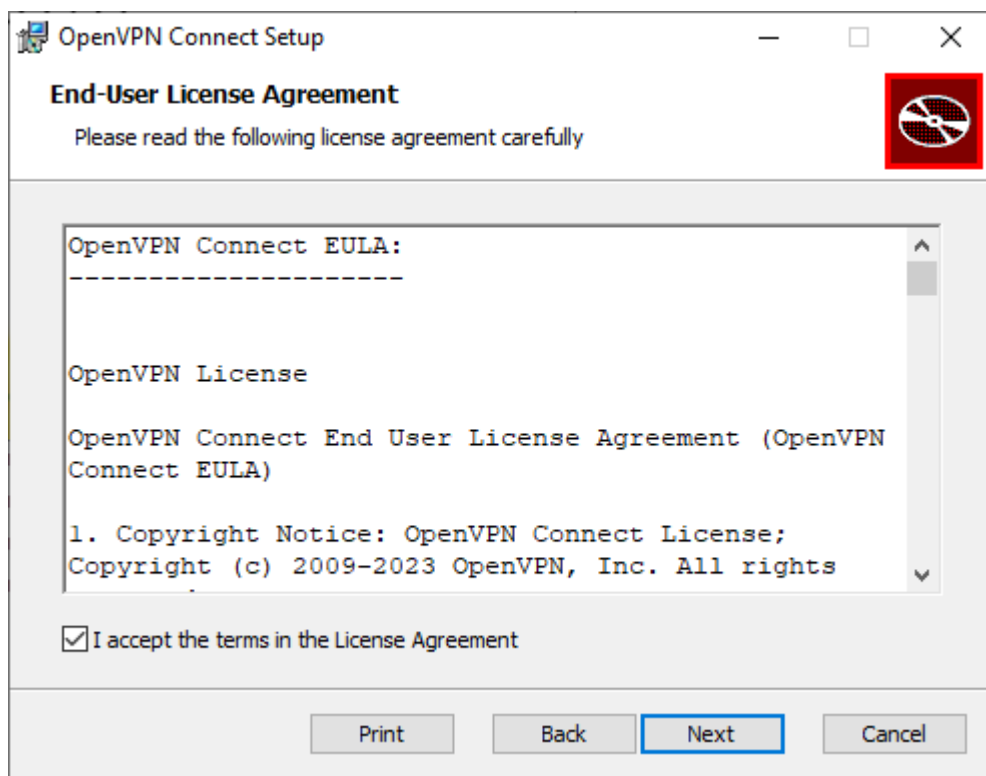
sha256 signature: f65dd0ea784dd63632be64f89b1f83d51c199fd7319888883780cb9e975c325a

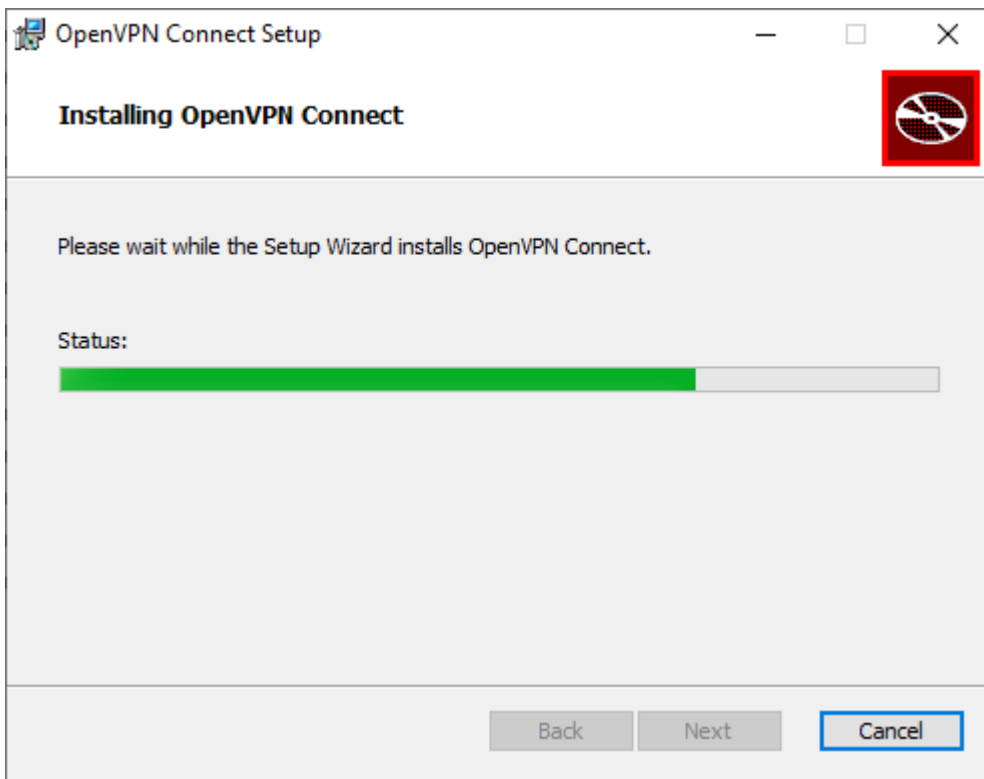
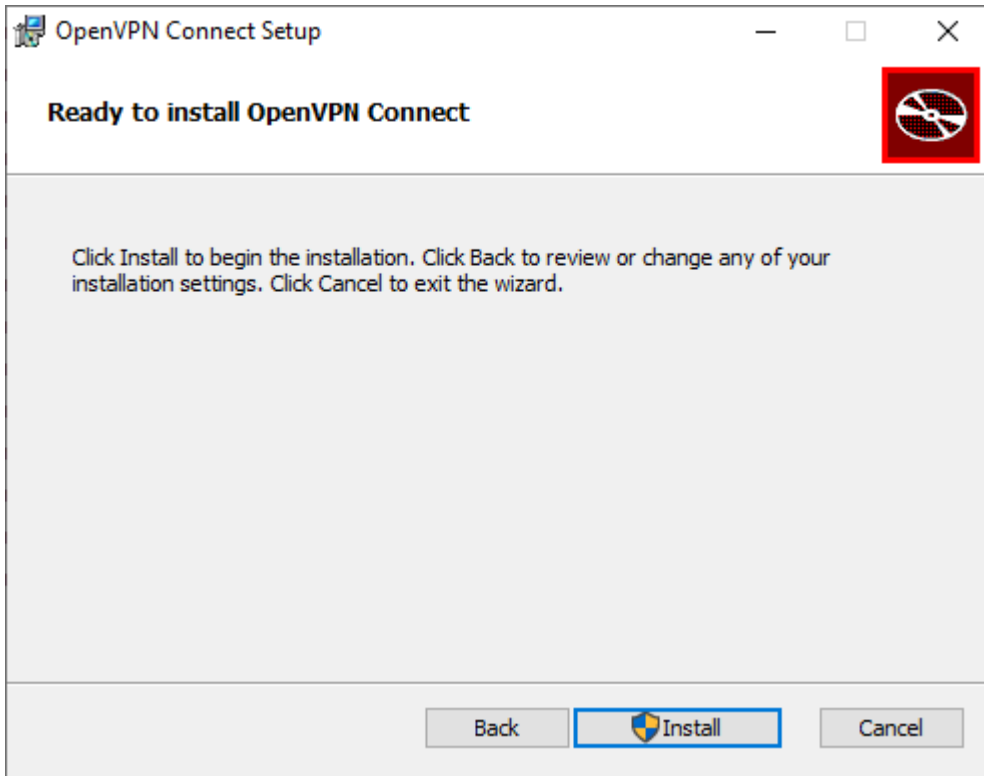
For Windows 7, 8, and 10.

Ejecutar el archivo descargado y le saldrá el asistente de instalación

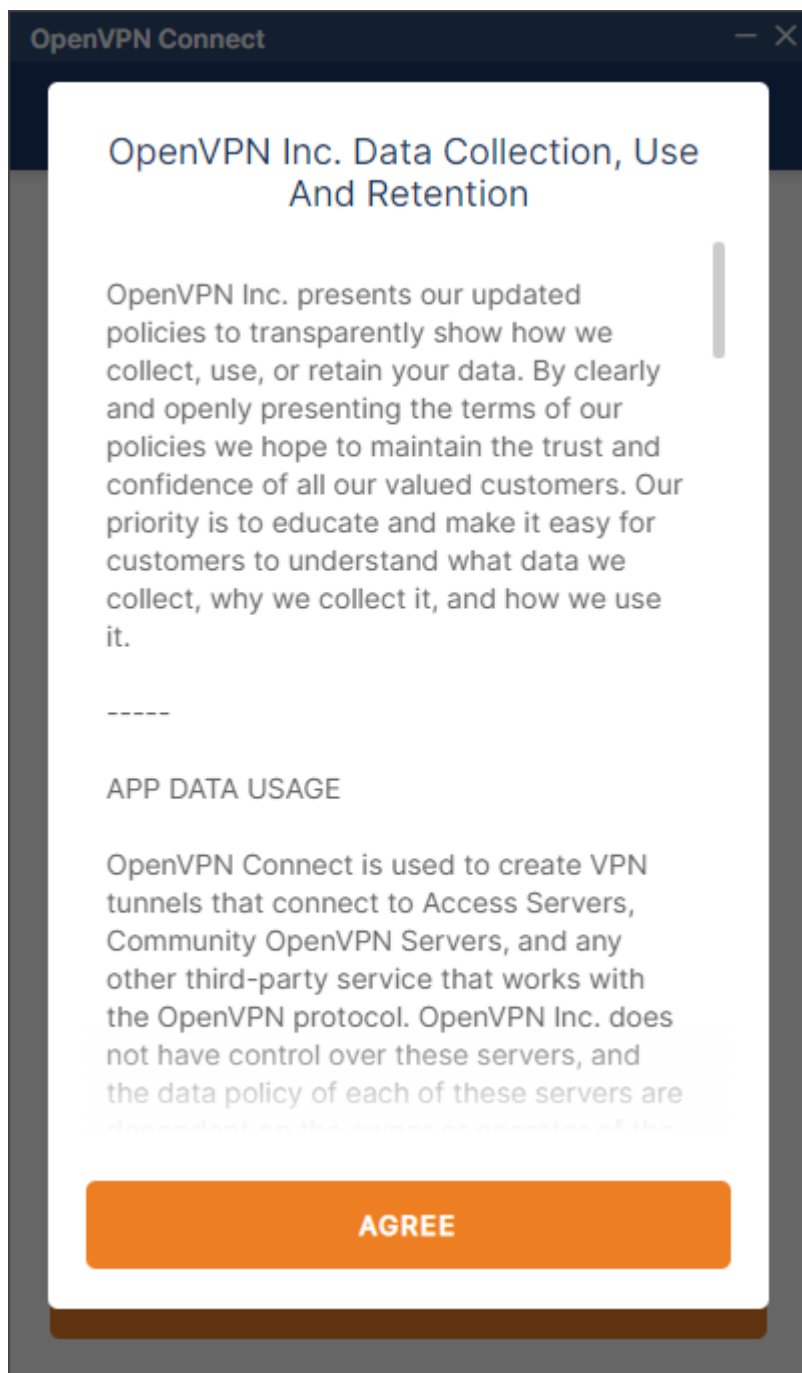


Acepte los términos de uso y siga las instrucciones de instalación:

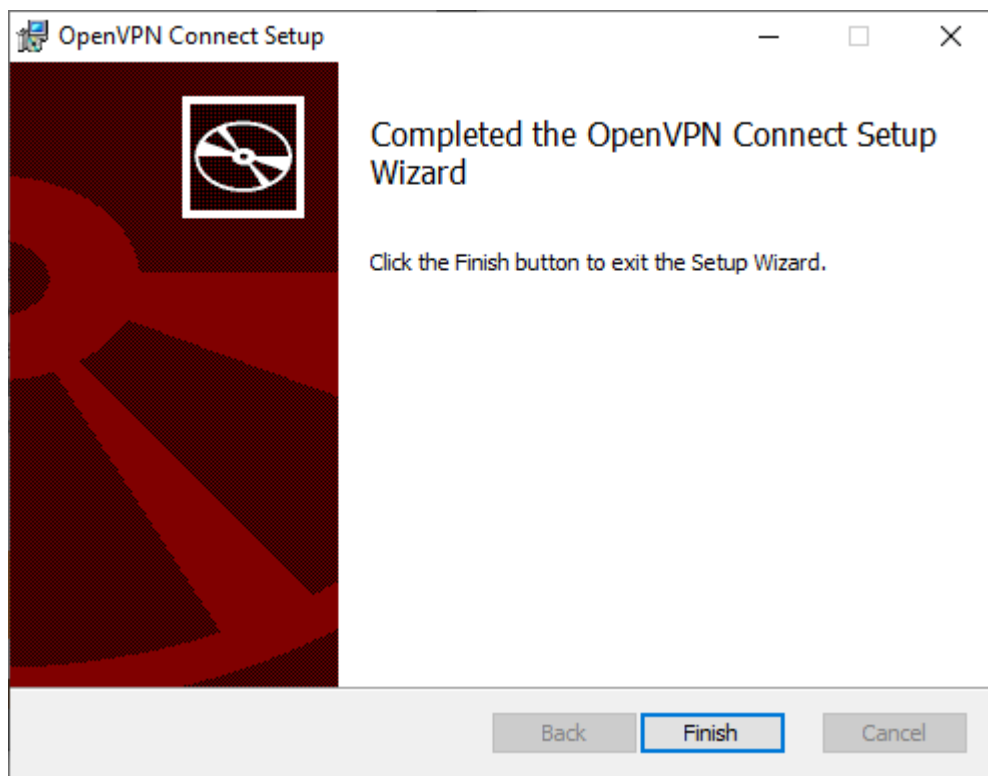




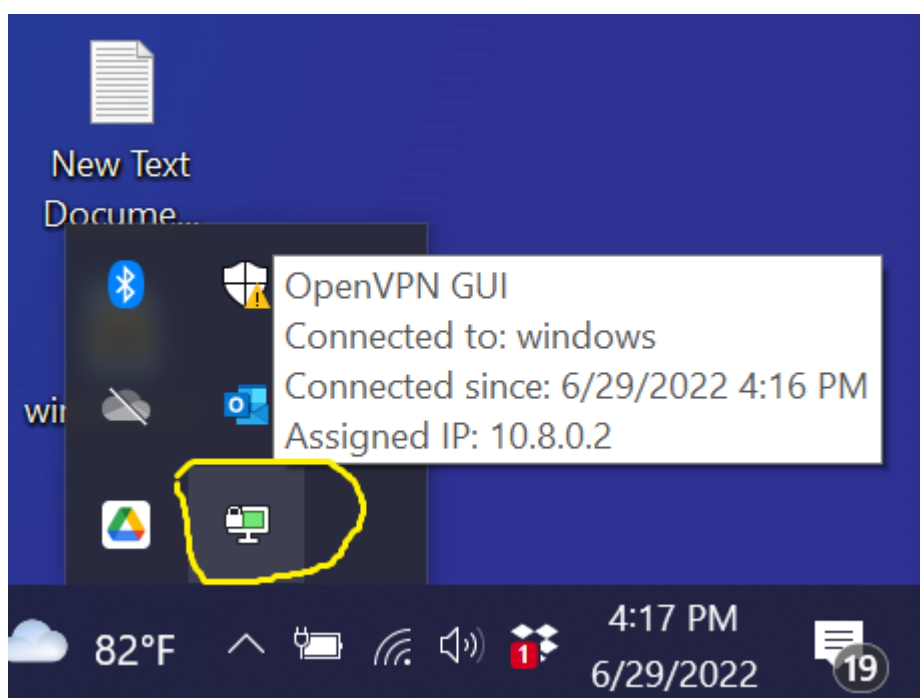
Acepte la licencia



y finalice la instalación



Con esto ya tendrá instalado el cliente de OpenVPN en su ordenador personal.



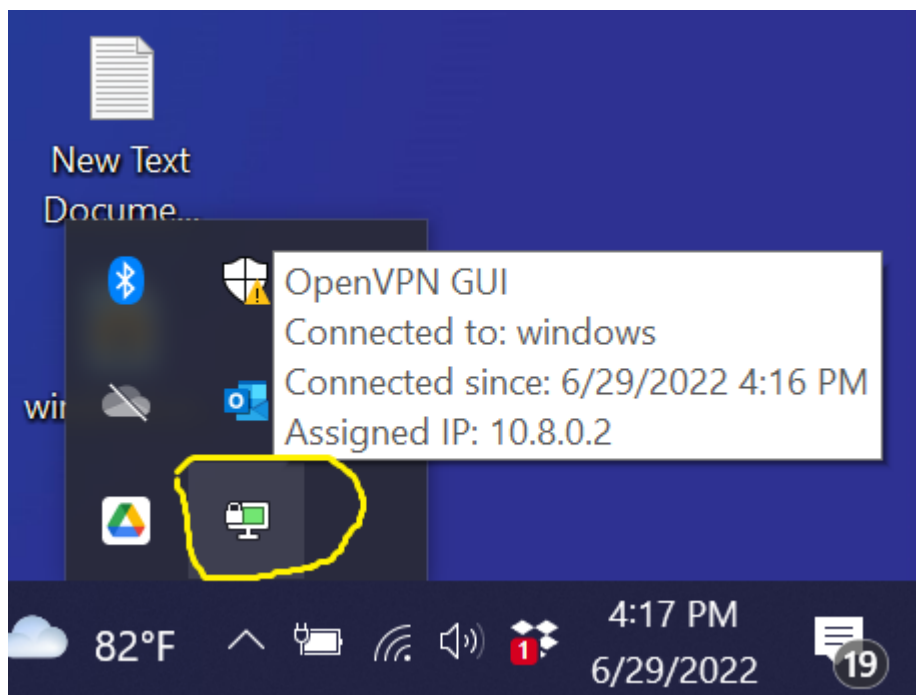
9.2 Importar perfil de conexión

Después de recibir las credenciales cifradas con su clave pública gpg y **descifrarlas** deberá todos los archivos recibidos en el mismo directorio antes de importar la configuración, por ejemplo:

```
C:\Users\your_user\VPN\  
  
client_teide-YYYY  
├── ca.crt  
├── client.ovpn  
├── README.txt  
├── ta.key  
├── teide-YYYY.crt  
└── teide-YYYY.key
```

9.3 Conectar a la VPN

Busque el icono del cliente OpenVPN en el área de notificaciones de windows.




Una vez abierto podrá configurar la conexión manualmente o lo más sencillo es importar el archivo *client.ovpn* arrastrando, haciendo doble click sobre él o bien, pulsando el icono *BROWSE*

OpenVPN Connect

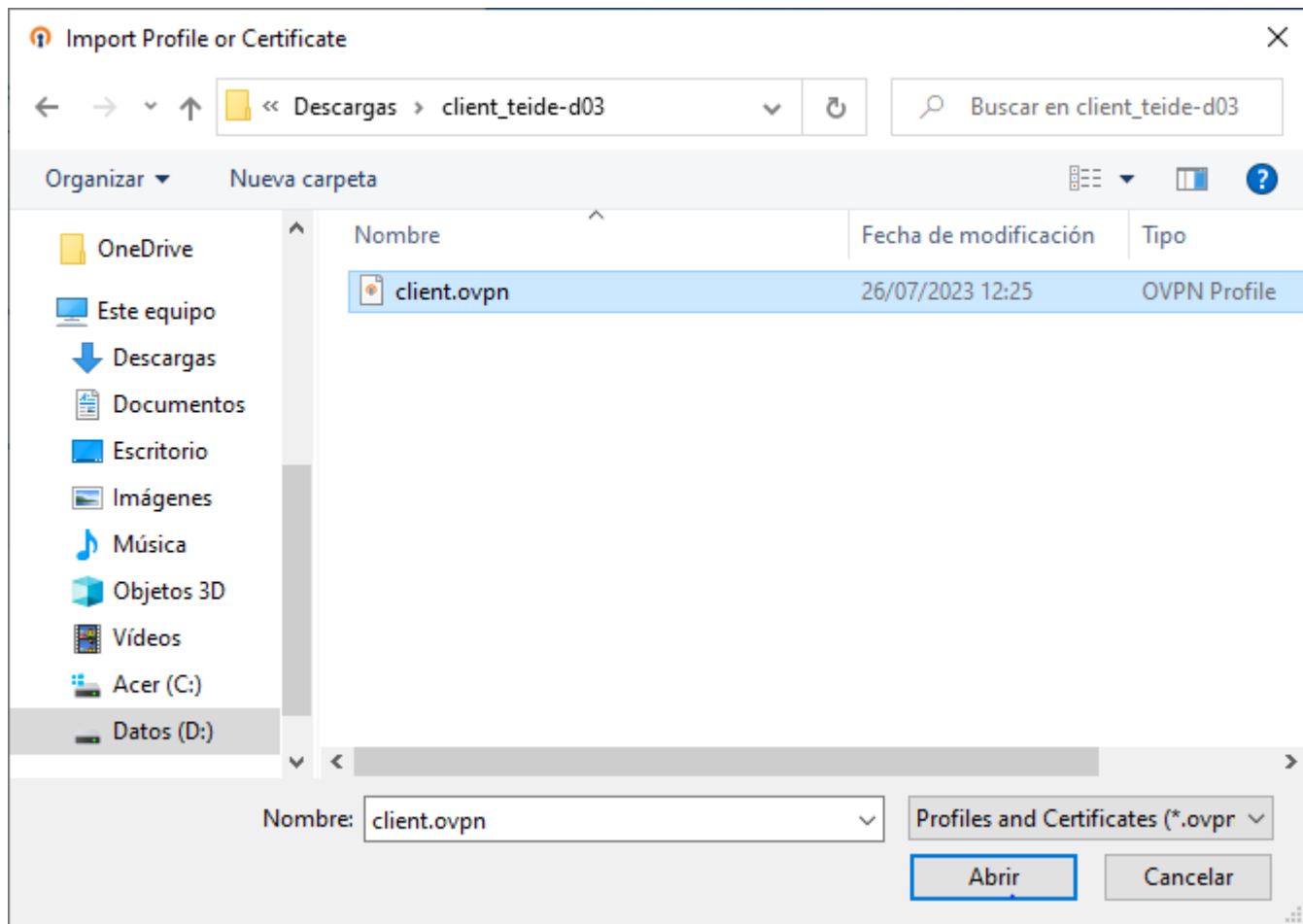
Import Profile

VIA URL **UPLOAD FILE**

Drag and drop to upload .OVPN profile.
You can import **only one profile** at a time.



BROWSE



Una vez introducidas las credenciales, el su sistema operativo queda en disposición de acceder a la infraestructura de TeideHPC.

OpenVPN Connect


Imported Profile

Profile Name
VPN TeideHPC

Server Hostname (locked)

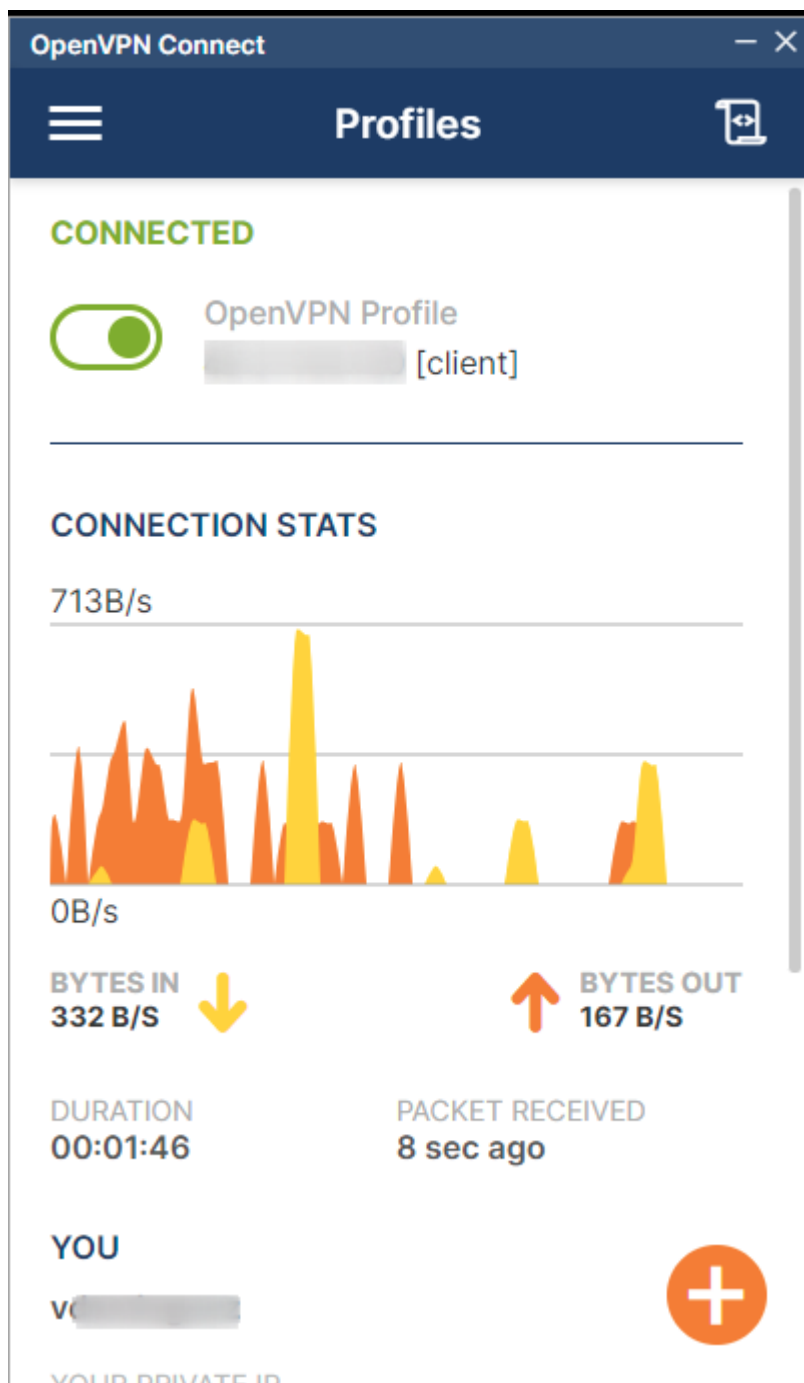
Username
youruser

Save password

Password
..... 

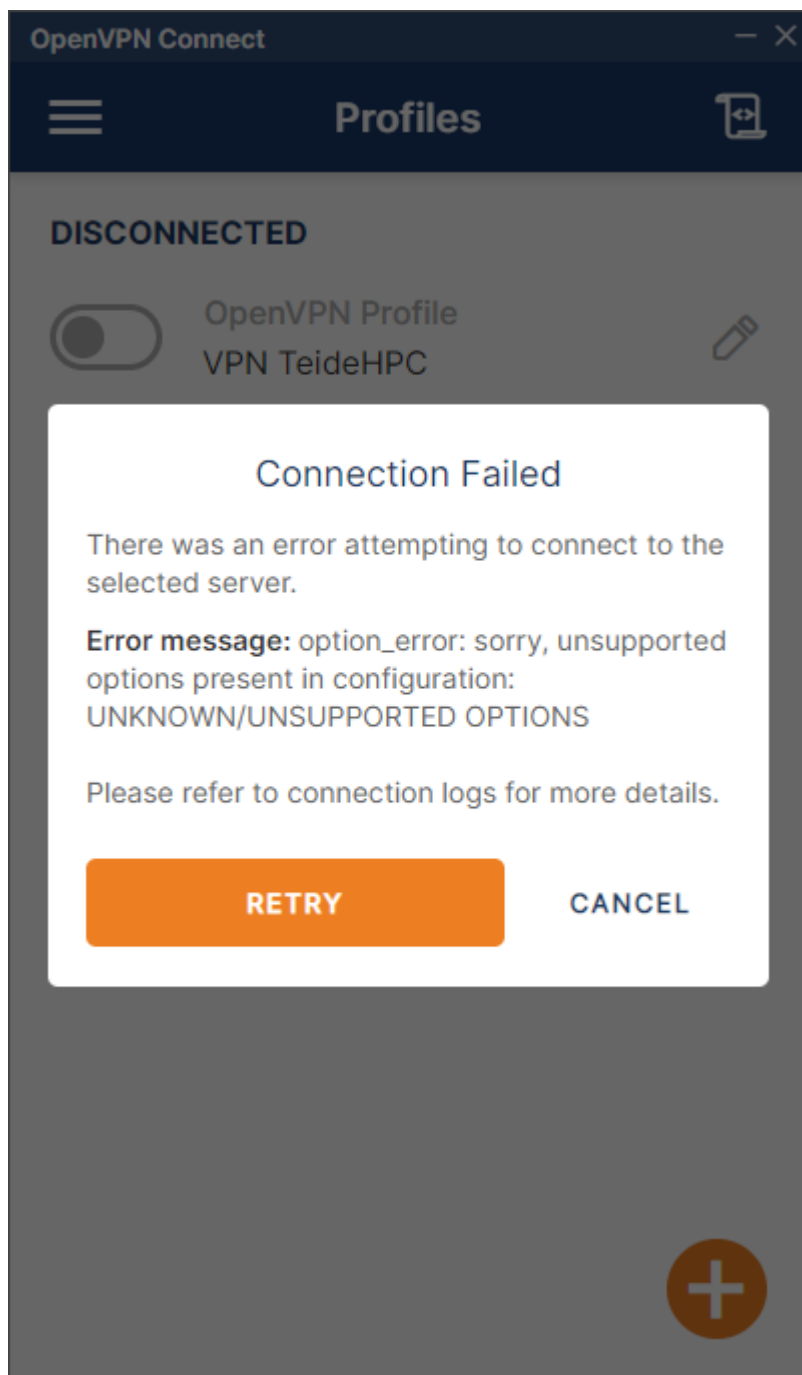
PROFILES **CONNECT**

Si su conexión es satisfactoria verá el siguiente estado en la conexión creada.

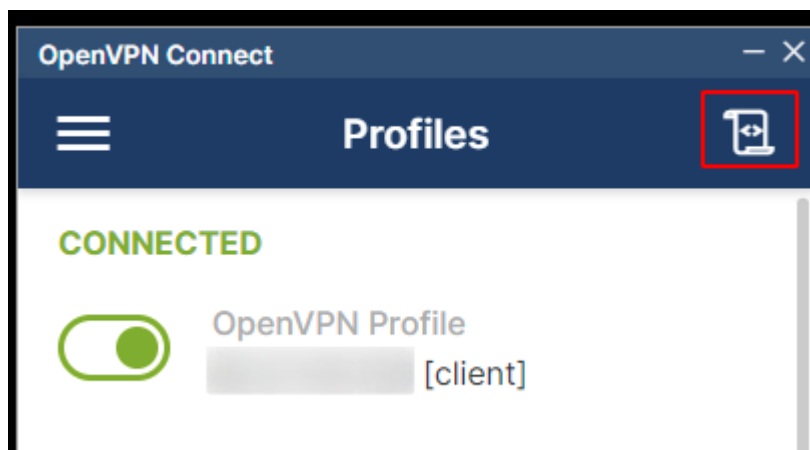


9.4 Log de OpenVPN para usuarios Windows.

En caso de que exista algún problema con la conexión a TeideHPC es importante consultar el log de conexión.



En caso de que la conexión no sea satisfactoria, éste siempre contendrá la información del error.




En caso de que no pueda solucionar el problema por si mismo, siempre le pediremos el log completo, el cual puede exportarlo haciendo click en el sobre.



10 OpenVPN para macOS

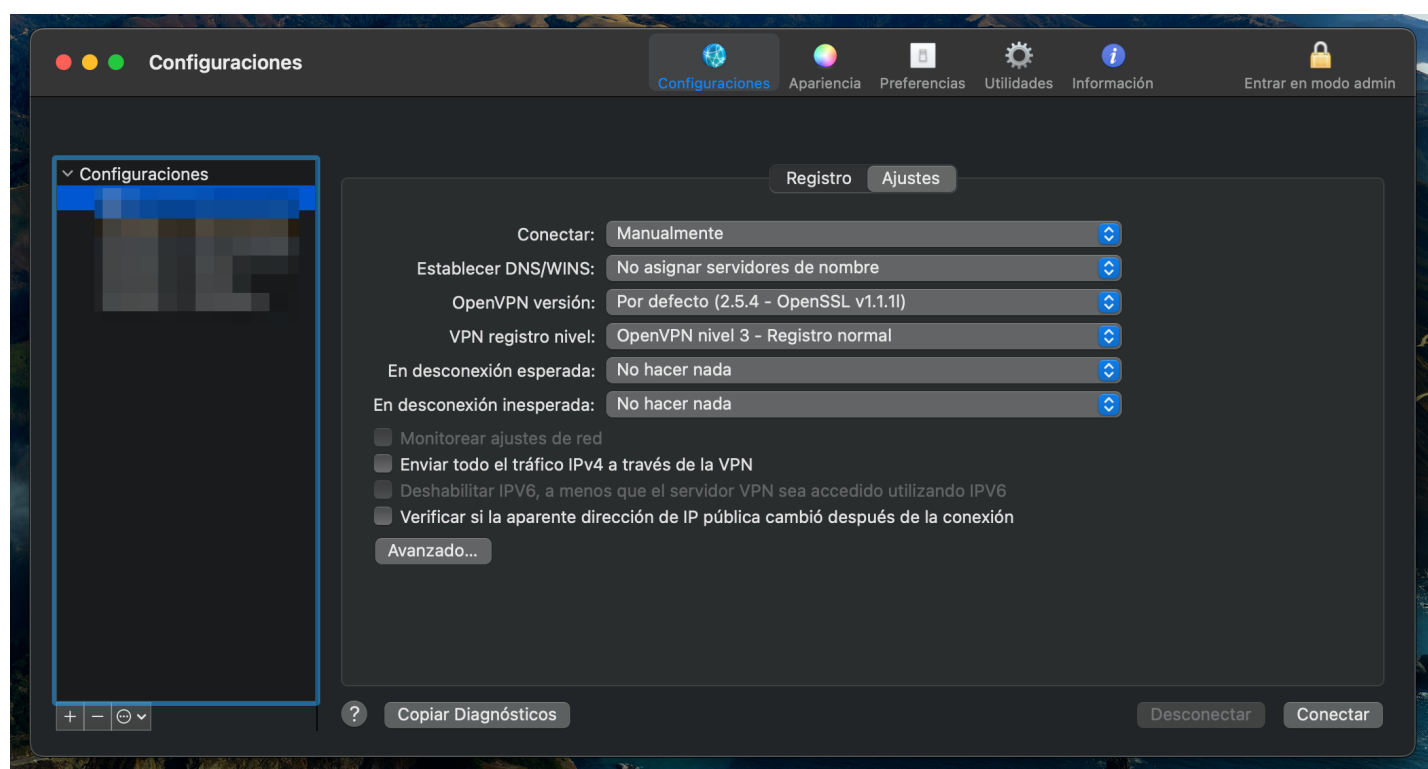
Para utilizar OpenVPN en macOS no disponemos de un cliente oficial, así que tenemos que utilizar una aplicación de terceros, **Tunnelblick**, que es la opción más recomendada para abrir una conexión de OpenVPN en macOS.

El presente manual se ha hecho utilizando macOS Ventura y Tunnelblick en su **versión estable v3.8.7a**.

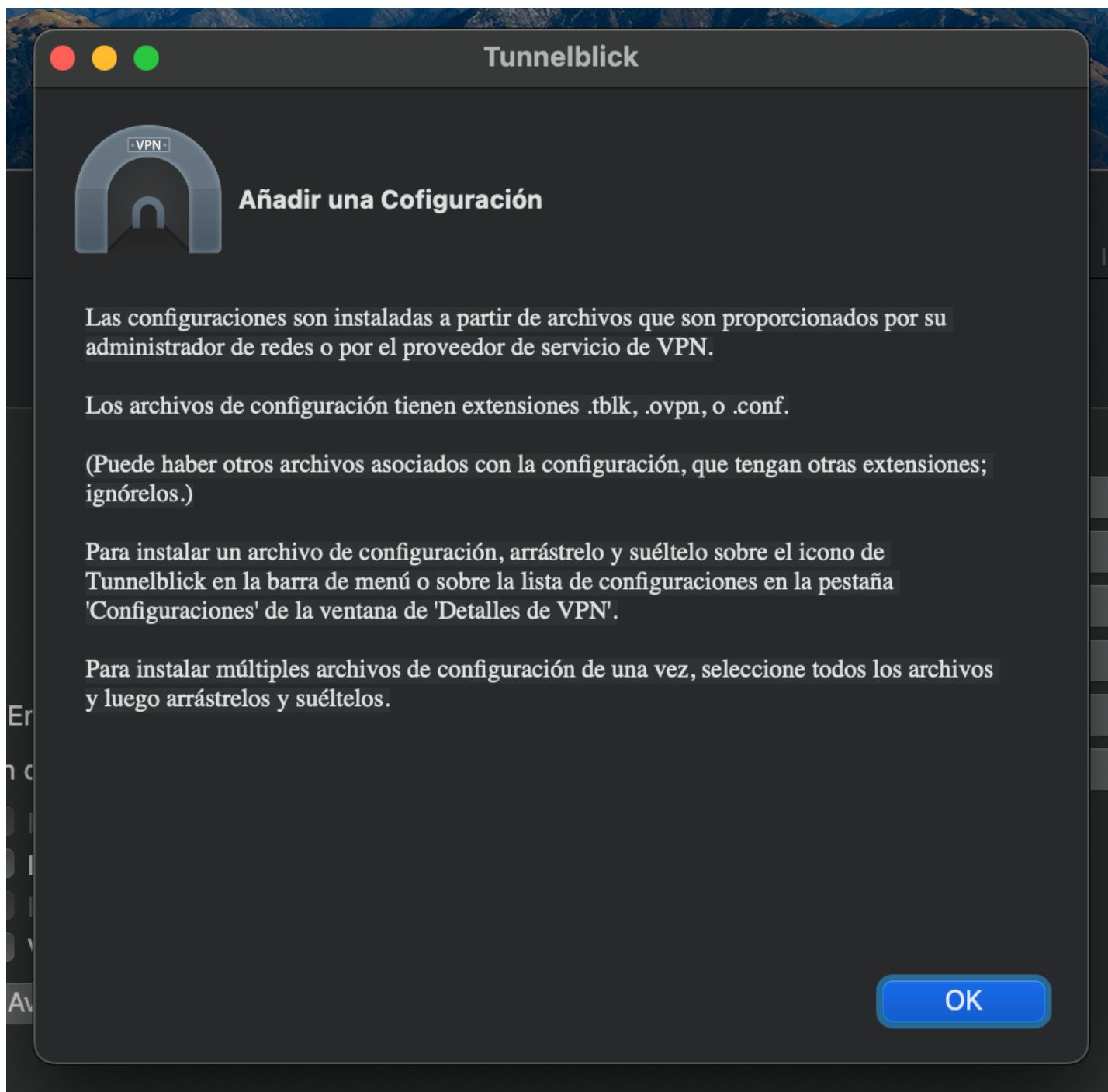
 **En este manual no se explica su instalación.**

10.1 Conectarse a la VPN utilizando Tunnelblick

Una vez instalado el software lo abrimos para añadir una nueva conexión VPN:

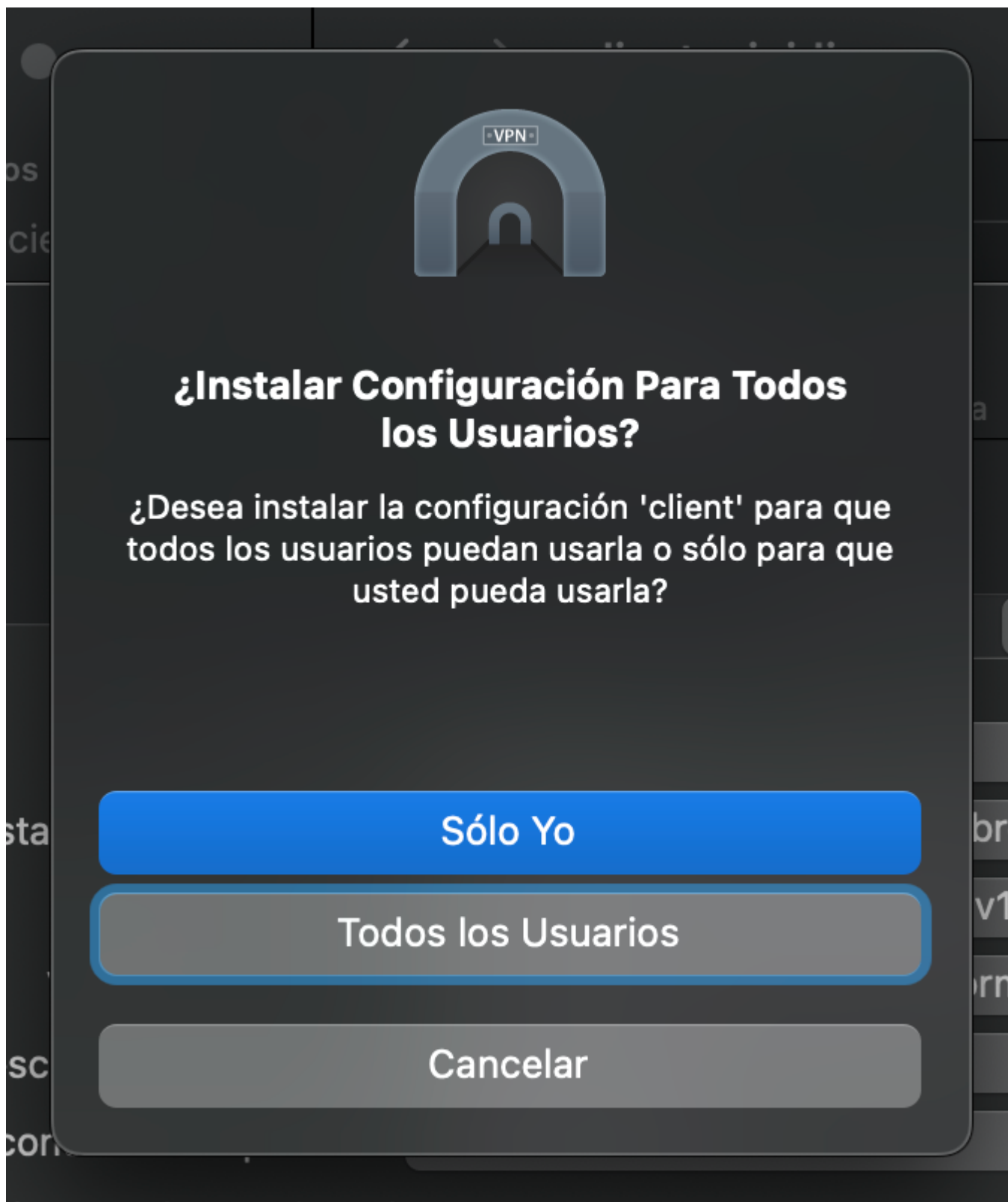


Para añadir una nueva conexión, podemos darle al símbolo "+" que hay en la esquina inferior izquierda:



Tal y como vemos, para crear una nueva conexión tenemos que arrastrar el fichero de configuración al apartado de *Configuraciones*. En nuestro caso sería el fichero `<nombre_usuario>.ovpn`. Lo hacemos, arrastramos el fichero de configuración para crear una nueva configuración VPN.

Dependiendo del ordenador en el que estemos trabajando, nos puede interesar crear la conexión únicamente para nuestro usuario o disponible para todos. En este caso, será solo para nuestro usuario, por tanto, "Solo Yo".



Se nos pedirá que introduzcamos la contraseña de usuario (del ordenador) para permitir que haga cambios:



Tunnelblick

Tunnelblick necesita:

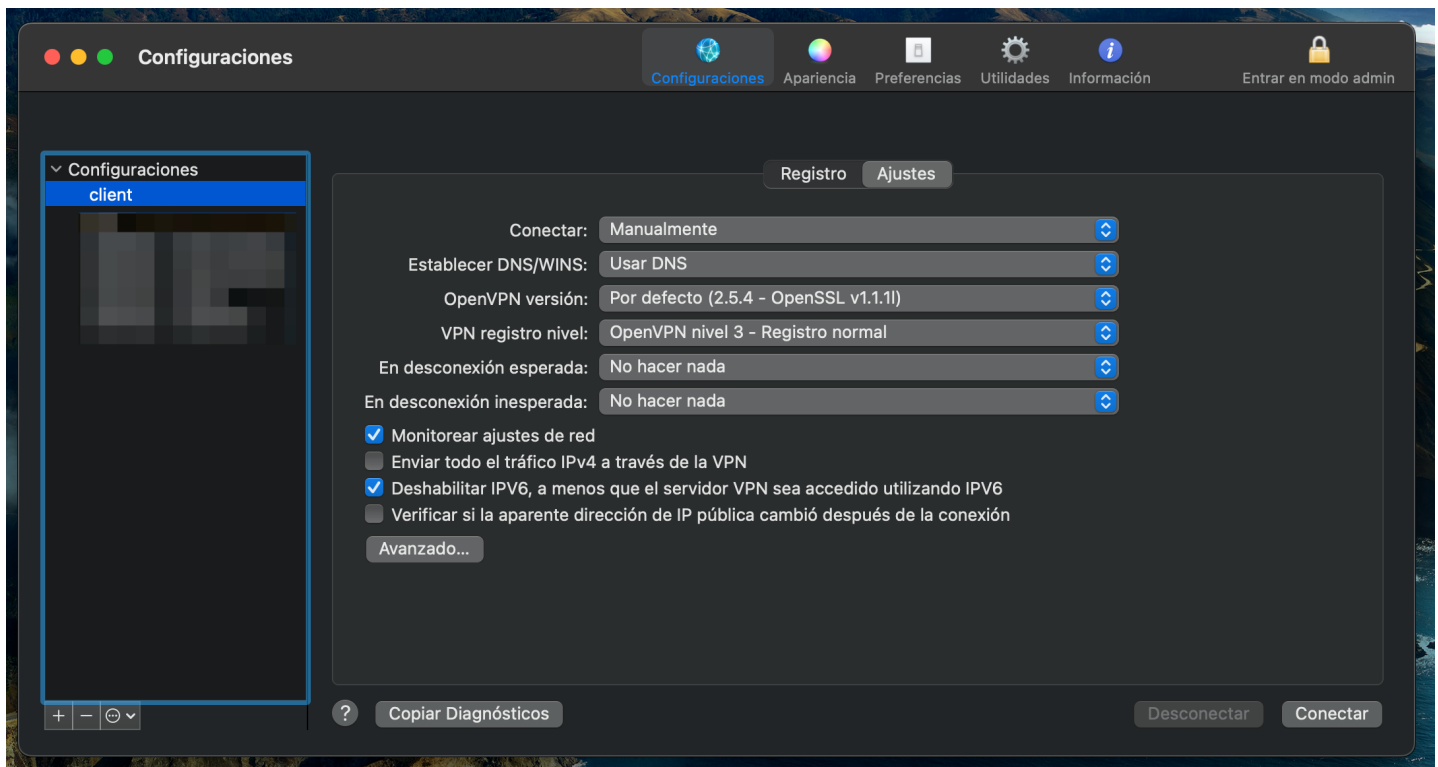
- Instalar una configuración

Introduce la contraseña para permitir esta operación.

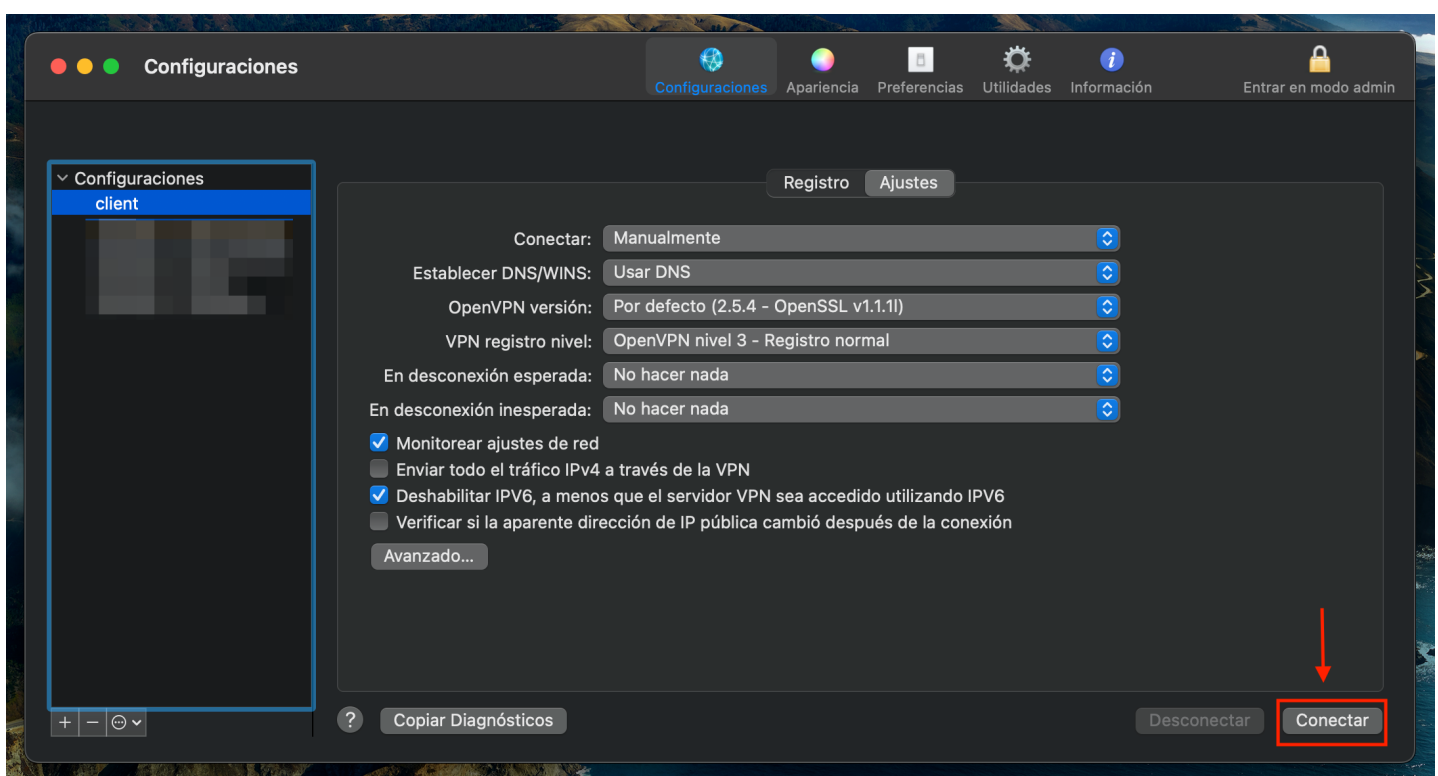
Cancelar

Aceptar

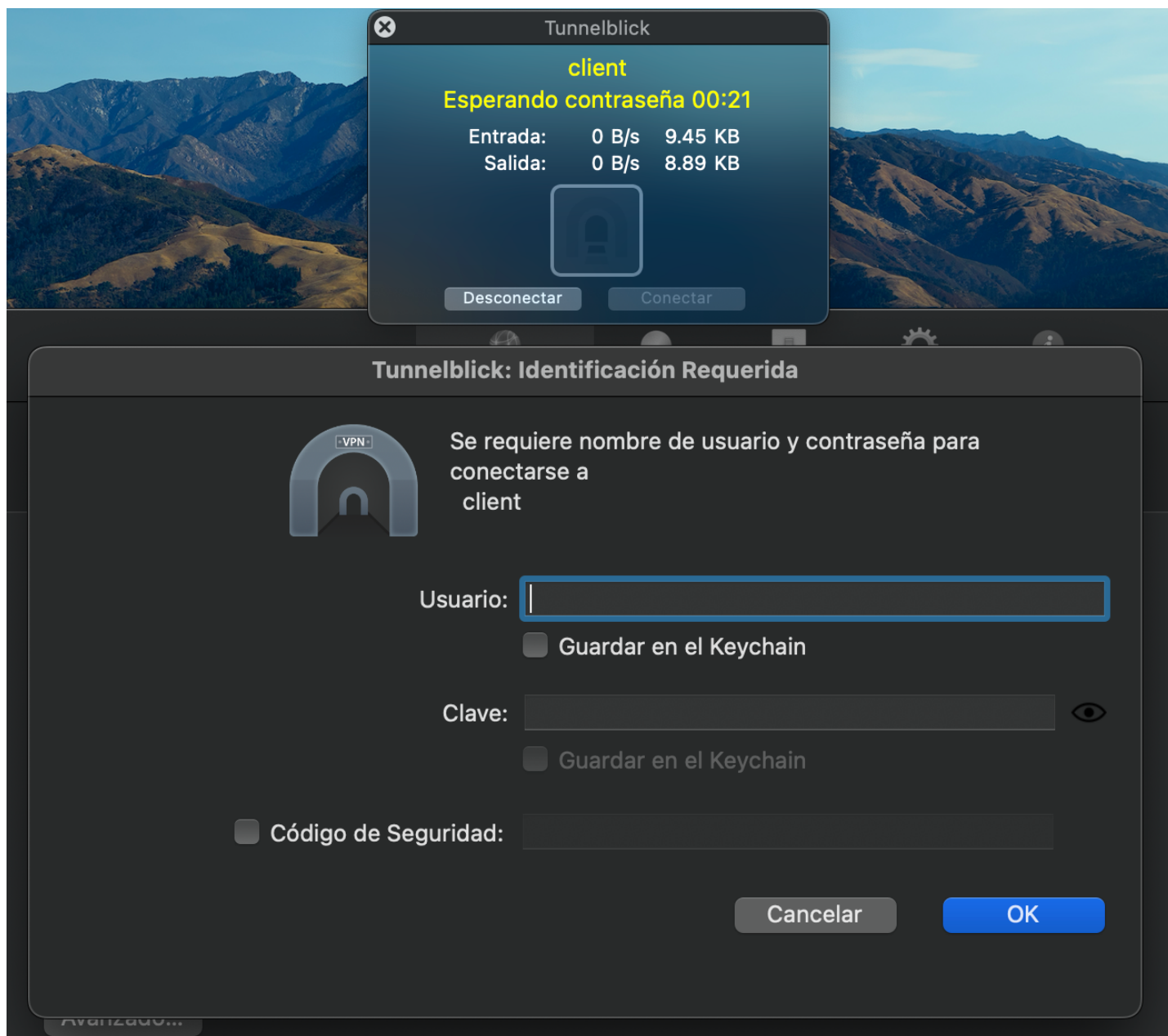
Una vez creada, aparecerá en el recuadro de la izquierda



En principio no se va a modificar la configuración de la conexión, por tanto, podemos darle a **"Conectar"**:



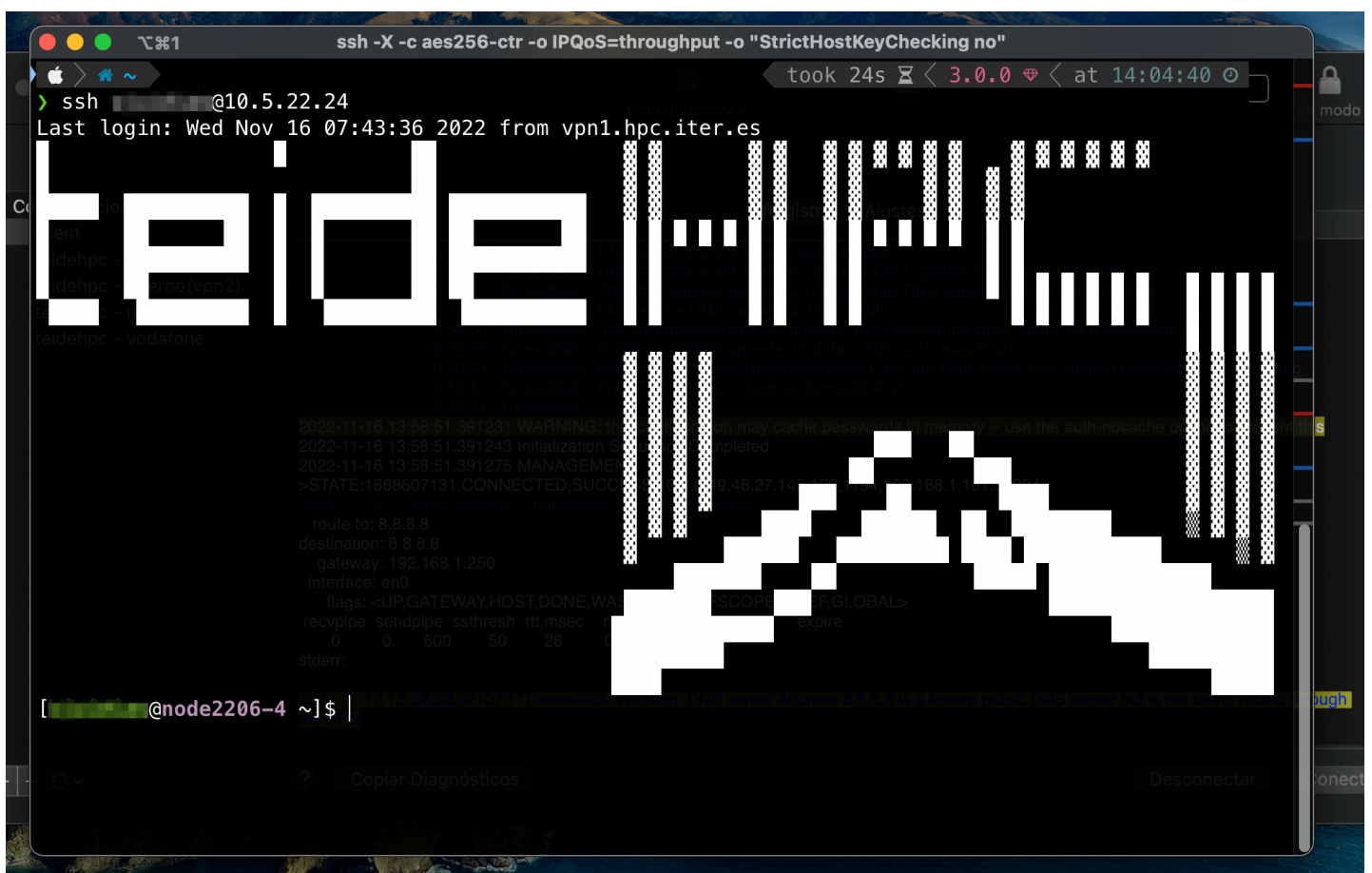
Al hacerlo, nos pedirá que introduzcamos las credenciales de la VPN, es decir, el usuario y contraseña:



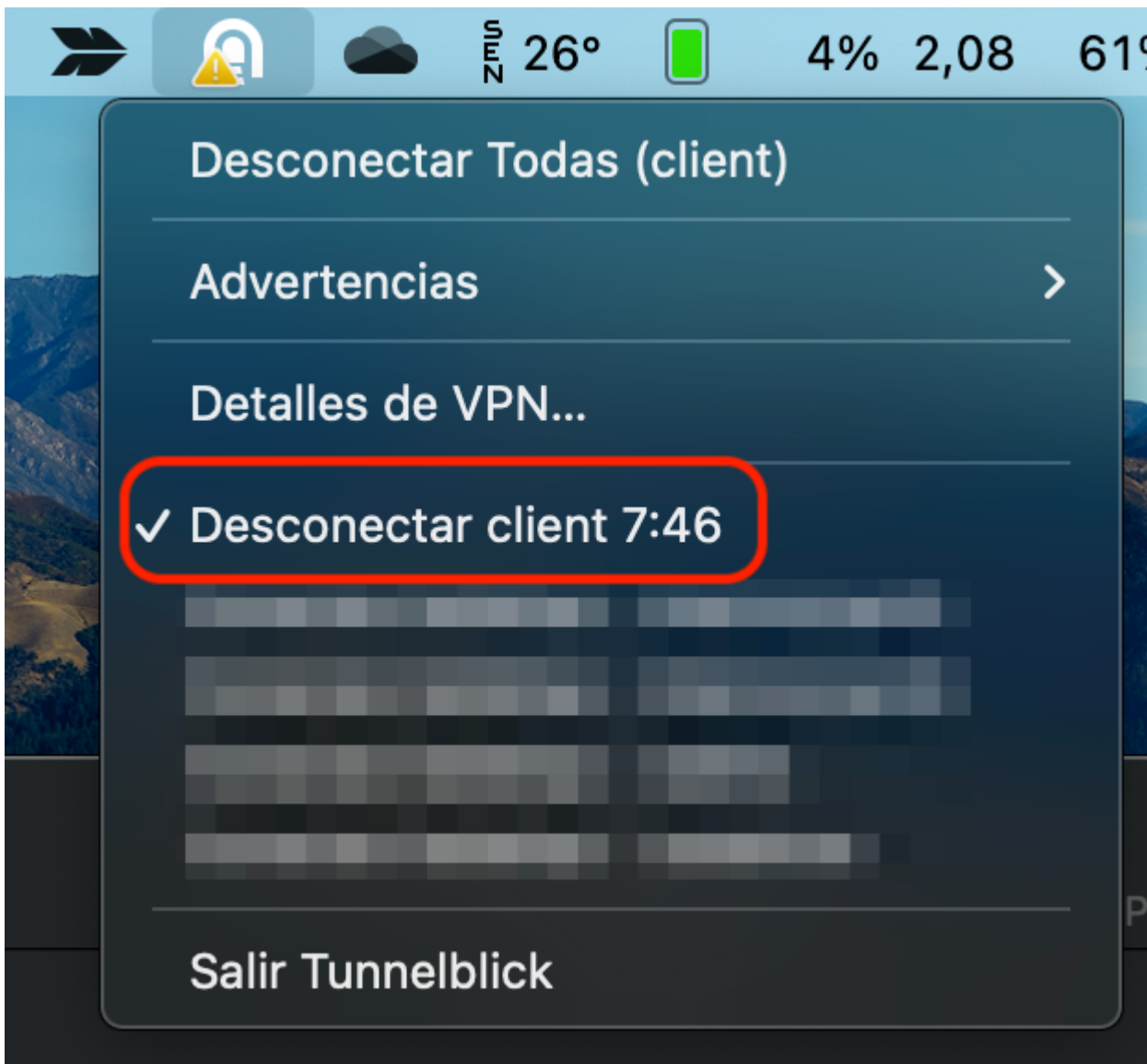
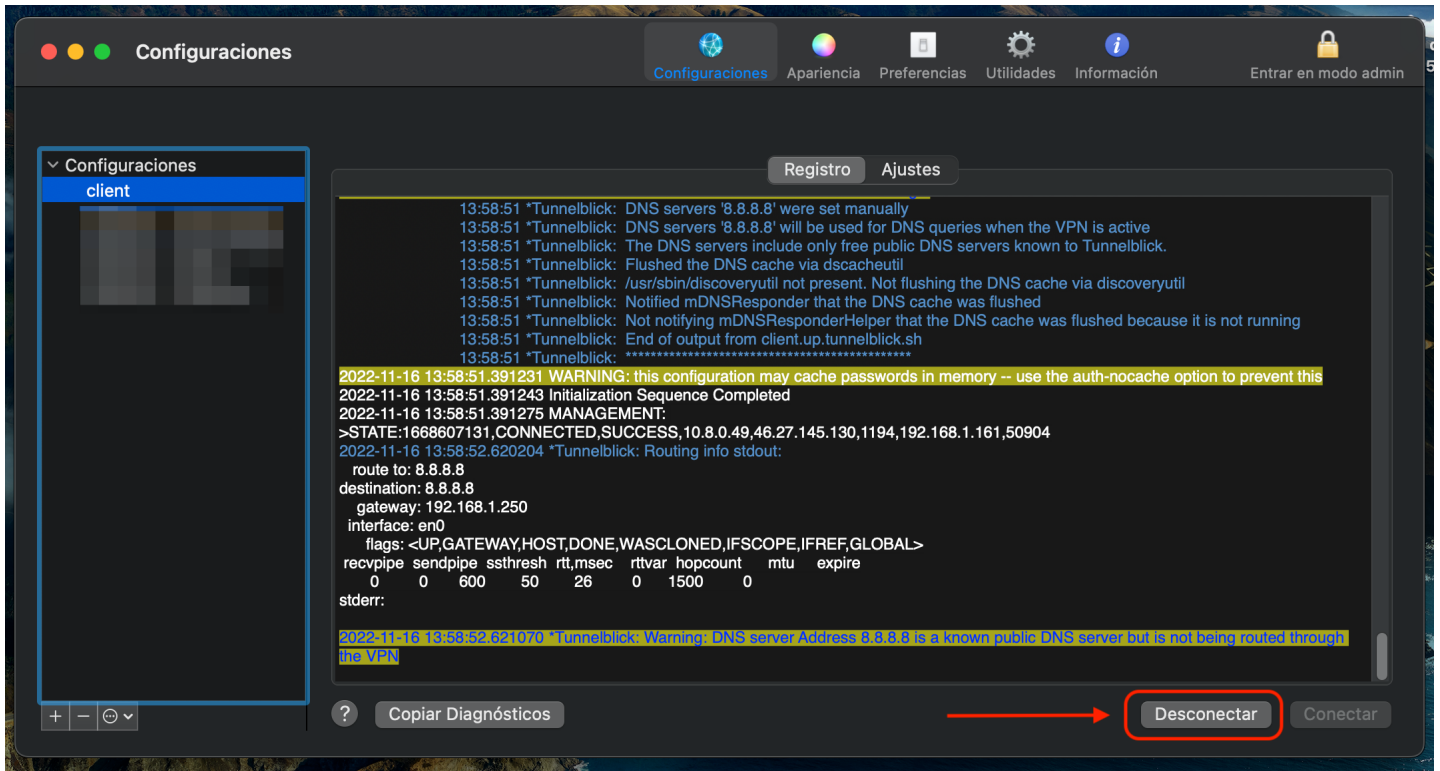
Si lo hemos introducido correctamente y todo ha ido bien, nos habremos conectado correctamente:



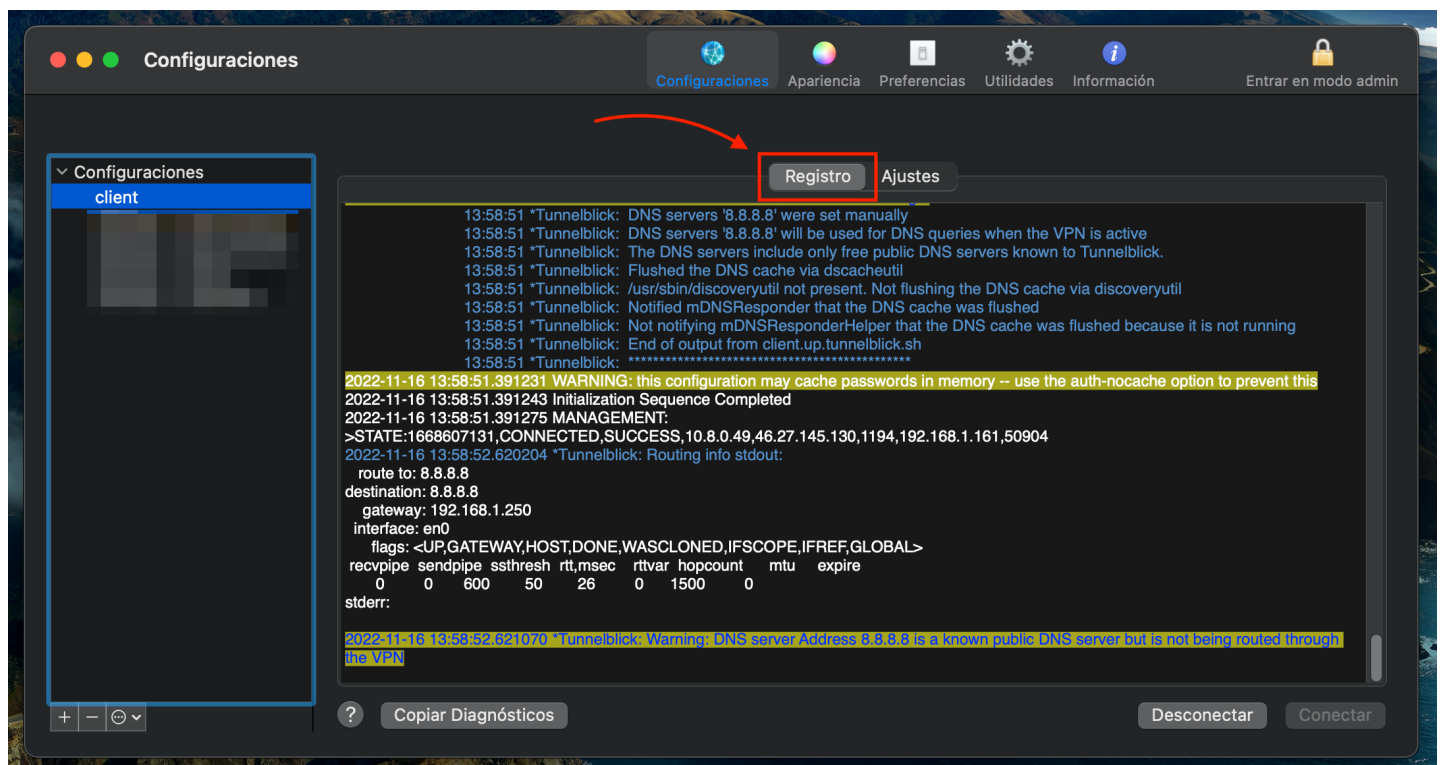
Una vez conectado a la VPN, podremos conectarnos vía SSH al nodo de login de TeideHPC utilizando la terminal:



Para desconectarnos, podemos hacerlo desde el propio programa o desde el icono del programa que encontramos en la barra superior:



En caso de tener problemas con la conexión, tenemos disponible la pestaña "**Registro**" para ver los mensajes de la conexión y así poder depurar un posible problema.



11 Añadir los servidores DNS de TeideHPC a nuestra conexión

Para facilitar la vida a los usuarios y no tengan que recordar todas las **IPs de acceso al cluster** de cómputo tenemos en funcionamiento un servidor DNS que permite resolver esas direcciones sin la necesidad de recordad la IP.

Nombre	IP Address
dns1	10.5.22.37
dns2	10.5.22.38

Para configurarla siga los siguientes pasos:

11.1 Añadir servidores DNS. Usuarios linux.

Los usuarios linux pueden añadir los DNS de TeideHPC en la **configuración de su conexión VPN**

11.1.1 Mediante línea de comandos.

Lista tus conexiones:

```
nmcli connection show
```

```

NAME                UUID                                TYPE      DEVICE
XXXXXXXX            00000000-7777-4f70-a5ca-a5be5c9551b5  wifi      --
....
teide hpc           12345678-1234-1234-1234-622e596d061e  vpn       --
....

```

Con el siguiente comando añade nuestros servidores DNS

```
nmcli c modify "<vpn-settings-name>" ipv4.dns '10.5.22.37 10.5.22.38'
```

Añade el dominio de búsqueda

```
nmcli c modify "<vpn-settings-name>" ipv4.dns-search 'hpc.iter.es'
```

11.1.2 Mediante interfaz gráfica (Ubuntu y Debian)

En una terminal escribe el siguiente comando:

```
nm-connection-editor
```

Selecciona tu conexión VPN y edítala. En la pestaña *IPv4 Settings* podrás añadir los DNS y el dominio de búsqueda como se muestra a continuación.

Editing hpc new rediris

Connection name: hpc new rediris

General | VPN | Proxy | **IPv4 Settings** | IPv6 Settings

Method: Automatic (VPN) addresses only

Additional static addresses

Address	Netmask	Gateway
---------	---------	---------

Add

Delete

DNS servers: 10.5.22.37, 10.5.22.38

Search domains: hpc.iter.es

Routes...

Export... | Cancel | Save

11.2 Añadir servidores DNS. Usuarios Windows

Antes de añadir tu tu fichero de configuración *client.ovpn* al cliente de OpenVPN añade las siguientes líneas:

```
dhcp-option DNS 8.8.8.8
dhcp-option DNS 10.5.22.37
dhcp-option DNS 10.5.22.38
push "dhcp-option hpc.iter.es"
```

I.III Primeros pasos

12 ¿Cómo iniciar sesión en TeideHPC?

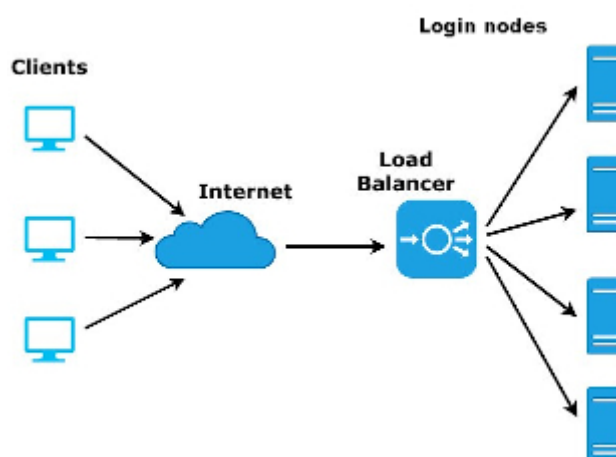
Con la llegada de la nueva infraestructura a TeideHPC, se han dispuesto varios nodos de login que permitan tener la redundancia y alta disponibilidad necesaria para mantener el cluster operativo el 100% del tiempo.

Para el acceso a la infraestructura el TeideHPC dispone de los siguientes *2 puntos de acceso*, uno para cada cluster.

TeideHPC	AnagaGPU
10.5.22.100	10.5.22.101

 **Recuerde que debe estar conectado a la VPN para acceder**

Para asegurar que el número de usuarios no se concentre en un sólo nodo de login, el cual eventualmente acaba saturándose debido a preferencias de acceso y ejecuciones no intencionadas de los usuarios, se ha dispuesto un **balanceador de carga para cada cluster el cual se encarga de distribuir a los usuarios de forma equitativa en cada uno de los 2 nodos de login que tiene cada cluster.**



 **Todos los nodos de login comparten todo su \$HOME**

Es decir, **independientemente a qué IP de acceso se conecten y de qué nodo le sea asignado, sus datos serán accesibles de manera indistinta.**

 **Una vez accedes, es posible ir de un nodo a otro sin ningún tipo de restricción**

```
ssh login-cpu01
ssh login-cpu02
ssh login-gpu01
ssh login-gpu02
```

El software disponible en cada cluster NO es el mismo

No todos los nodos son iguales, cada cluster tiene un tipo de nodos según la arquitectura de estos por lo que el software está compilado en su mayoría atendiendo a la arquitectura de los nodos.

En la [página principal](#) de esta documentación puede ver ver qué arquitecturas existen en TeideHPC.

12.1 Acceso para los usuarios de Linux y macOS

A través de una terminal:

```
ssh miusuario@IPservidor o
ssh miusuario@login.hpc.iter.es (*)
```

Usar DNS en lugar de IPs

Recuerda que puedes acceder a usando nuestro DNS siempre que los [hayas añadido a tu configuración](#).

12.1.1 SSH Alias

Para que sea más sencillo trabajar y no estar recordando la IP del servidor, podemos utilizar un alias para guardar la conexión. Para ello, editamos el fichero `~/.ssh/config` y añadimos lo siguiente (si no existe, lo creamos):

```
~/.ssh/config
```

```
Host teidelogin
  Hostname 10.5.22.100
  User miusuario

Host anagalogin
  Hostname 10.5.22.101
  User miusuario
```

Ahora para conectarnos por ssh a los nodos de login podemos hacerlo de la siguiente manera:

```
ssh teidelogin
ssh anagalogin
```

12.1.2 Acceder con clave pública SSH

Podemos acceder a los nodos de login sin contraseña utilizando una clave pública SSH. Si no tenemos ninguna, podemos hacerlo de la siguiente manera:

```
ssh-keygen -b 4096
```

Una vez se ejecute el comando, nos pedirá dos cosas:

- Una localización donde guardar la clave y un nombre para el archivo.
- Una contraseña para encriptar la clave y que tendremos que utilizar cada vez que usemos la clave pública.

Podemos dejar vacío ambos campos. En el caso de la localización, por defecto, el par de claves se guardarán en el directorio `~/.ssh`:

- Para la clave privada: `~/.ssh/id_rsa`
- Para la clave pública: `~/.ssh/id_rsa.pub`

Warning

La clave privada la debe guardar usted y no compartirla con nadie. Es la que se usará para realizar la autenticación con el servidor. Si la pierde, no podrá conectarse utilizando la clave pública ssh.

En cuanto a la contraseña, eso ya es decisión de cada uno.

Para copiar la clave pública SSH al nodo de login, debe hacer lo siguiente:

```
ssh-copy-id -i ~/.ssh/id_rsa.pub mi-usuario@ip-nodo-login
```

Nos pedirá la contraseña de nuestro usuario para proceder y ya podremos conectarnos a los nodos de login sin necesidad de contraseña, siempre y cuando lo hagamos desde el ordenador donde está la clave privada.

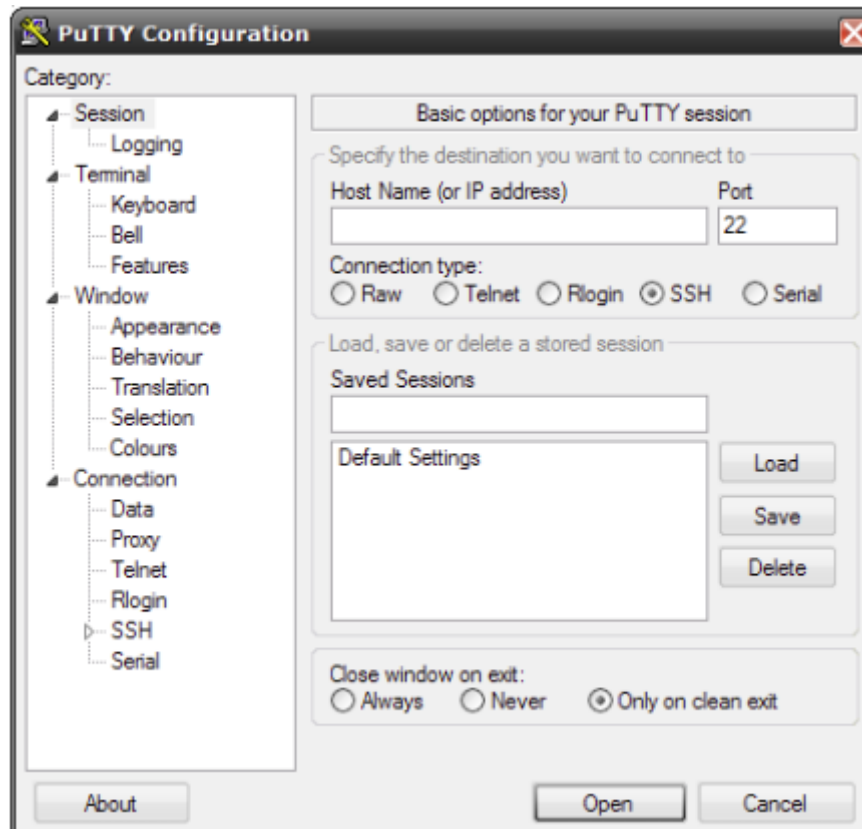
12.2 Acceso para usuarios de Windows

Los usuarios de windows disponen de varias alternativas para conectarse vía SSH a los nodos de login. Entre ellas están **PuTTY** y **MobaXterm**:

12.2.1 Acceso remoto SSH con PuTTY

PuTTY es un cliente de red que soporta los protocolos SSH, Telnet y Rlogin y sirve principalmente para iniciar una sesión remota con otra maquina o servidor. Es de licencia libre y a pesar de su sencillez es muy funcional y configurable.

Una vez descargado e instalado el software habrá que seguir los siguientes pasos para establecer la conexión con los nodos de login de TeideHPC



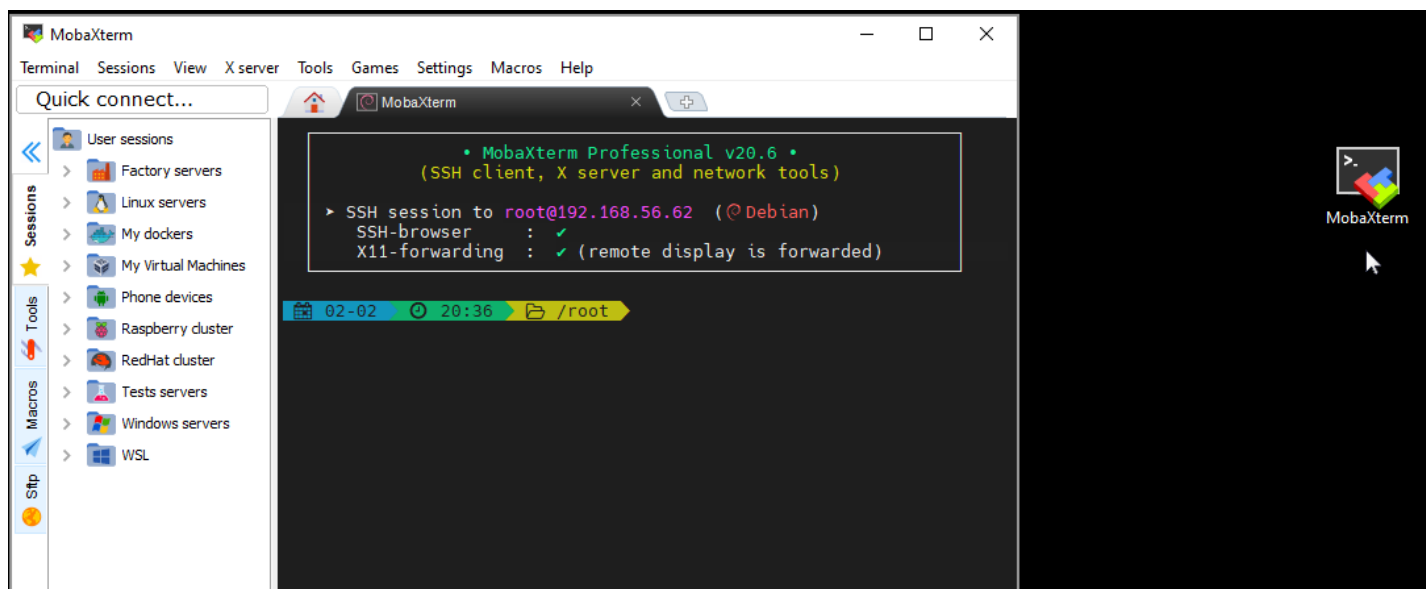
1. En el menú de configuración seleccione la categoría Session.
2. Introduzca el nombre de su dominio o IP en el campo Host Name y seleccione el protocolo SSH.
3. Introduzca un nombre para esta conexión en el campo Saved Sessions.
4. Vuelva al menú de configuración y seleccione la categoría SSH.
5. Asegúrese de que está marcada la opción 2 en Preferred SSH protocol version.
6. Seleccione nuevamente la categoría Session.
7. Para guardar la configuración pulse Save y Open para conectar.

Consejos:

- For slow connections you can enable compression. You can find a checkbox in the Connection > SSH menu.
- SSH version 2 must be set as the preferred protocol version in Connection > SSH menu.

12.2.2 MobaXterm

MobaXterm es un toolbox para trabajar de manera remota. En una sola aplicación proporciona un montón de funciones (SSH, X11, RDP, VNC, FTP, sFTP, MOSH, comandos unix) que están diseñadas para programadores, webmasters, administradores de TI y prácticamente todos los usuarios que necesitan manejar sus trabajos remotos de una manera más simple.



La versión gratuita tiene ciertas limitaciones como el número de sesiones simultáneas, pero puede usarse libremente sin fines comerciales.

Para configurar un cliente SSH en MobaXterm, sigue estos pasos:

1. Iniciar MobaXterm: Abre MobaXterm en tu computadora.
2. Sesión Nueva: Haz clic en el botón "Session" en la parte superior izquierda de la pantalla de inicio de MobaXterm.
3. Seleccionar SSH: En la ventana de "Session settings", selecciona la opción "SSH" en la lista de tipos de sesión.
4. Configurar Parámetros de SSH:

- Remote host: Introduce la dirección IP o el nombre de host del servidor al que deseas conectarte.
- Specify username: Puedes especificar un nombre de usuario si no quieres ingresar el nombre de usuario cada vez que te conectes al servidor.
- Port: Cambia el puerto si es necesario (el puerto por defecto para SSH es 22).
- Advanced SSH settings: Si necesitas configuraciones avanzadas como el uso de una clave SSH específica, haz clic en "Advanced SSH settings" y realiza las configuraciones necesarias.

1. Guardar la sesión (opcional): Puedes guardar la configuración de la sesión para futuras conexiones. Dale un nombre a la sesión y guarda la configuración.
2. Conectar: Haz clic en "OK" o "Open" para iniciar la conexión SSH. Si es la primera vez que te conectas a este servidor, es posible que se te pida que verifiques la clave del host y que ingreses tu contraseña.
3. Autenticación: Ingresa tu contraseña cuando se te solicite. Si configuraste una clave SSH y la contraseña de la clave, ingrésala.
4. Sesión Iniciada: Una vez que se complete la autenticación, deberías estar conectado a tu servidor a través de SSH en la terminal de MobaXterm.

En este [enlace](#) puede encontrar una guía más detallada:

13 Cuentas de usuario

Las cuentas de usuarios son proporcionados por el equipo de administración de TeideHPC previa solicitud. Las cuentas de usuario tienen carácter **personal e intransferible**, lo que significa que está **prohibido compartir las credenciales VPN, así como las contraseñas con terceras personas**.

Recordamos que solo **se permite una conexión VPN simultánea**, es decir, no podrá estar conectado a la VPN al mismo tiempo desde dos dispositivos.

Para cualquier consulta, dificultad o necesidad de asistencia adicional, no dudes en ponerte en contacto con nosotros a través del correo electrónico support@hpc.iter.es. Estamos aquí para ayudarte.

14 Almacenamiento

Salvo casos particulares, las cuentas tendrán asociadas el siguiente esquema de almacenamiento:

Almacenamiento	Capacidad	Descripción
<code>/home</code>	5 GB	Recomendable no almacenar datos o código en esta partición. Cuando se alcance su máxima capacidad, tendrá que mover los archivos a <code>/data</code>
<code>/data</code>	X TB	Partición con una " cuota soft " sobre el almacenamiento contratado. Para grandes volúmenes será necesario solicitarlo mediante el email de soporte (support@hpc.iter.es). Es accesible en <code>/home/<usuario>/data</code>
<code>/scratch</code>	--	Almacenamiento sólo disponible bajo solicitud expresa
<code>/local/<jobid></code>	300GB	En el HDD local del nodo que le asigne el planificador a la ejecución. El contenido de esta partición se elimina una vez finalizada la ejecución. Al lanzar una ejecución al gestor de colas se crea la carpeta en el path indicado. El jobid es accesible a través de la variable de entorno <code>\$SLURM_JOBID</code>
<code>/lustre</code>	X TB	Almacenamiento de alto rendimiento basado en LUSTRE con mayores prestaciones en cuanto a ancho de banda. Debe utilizarse como almacenamiento de datos en utilización por tareas de cómputo, una vez disponibles los resultados tienen que transferirse a home o data, ya que la permanencia de datos en esta partición no está asegurada y puede someterse a operaciones de limpieza periódicamente. Es accesible <code>/home/<usuario>/lustre</code>

Cualquier cambio en este esquema de almacenamiento será notificado.



Tu directorio *home* tiene un límite *soft* de 5GB.

!!!Recomendamos guardar tus entornos virtuales, datos de aplicación y resultados en el directorio *data* or *lustre* (si necesitas almacenamiento de alto rendimiento!!!

14.1 Copias de seguridad

Los datos de los usuarios están respaldados por el sistema de copias de seguridad del sistema de almacenamiento. El esquema de las copias de seguridad es el siguiente:

- Para `/home`, 6 copias horarias, 2 copias diarias y 2 semanales.
- Para `/data`, 6 copias horarias, 2 copias diarias y 2 semanales.

Para `/local` no hay copias de seguridad, porque es un almacenamiento situado en los discos duros de los nodos de cómputo, enfocado a mejorar el rendimiento de los trabajos. Una vez éstos terminen, el contenido de esta partición se borra por completo. La idea es que, una vez se inicie la ejecución del trabajo, el usuario copie los datos con los que va a trabajar a esta partición del nodo y, antes de que finalice por completo el trabajo, es decir, en el mismo script de ejecución, mueva los resultados obtenidos a su espacio en **data**.

Para `/lustre` tampoco hay ningún sistema de copias de seguridad por la misma razón. Es un almacenamiento dedicado a optimizar las ejecuciones de los usuarios, pero esta vez, utilizando la red infiniband que interconecta todos los nodos y que cuenta con velocidades muy superiores al resto de redes. La idea es que el usuario copie los datos de las ejecuciones aquí, y cuando estas terminen, mueva los resultados de vuelta a su espacio en **data**. En este caso, nosotros no borramos nada, pero si por algún motivo ocurre algún problema y los datos se ven afectados o el usuario borra datos por equivocación, no hay copias de seguridad para recuperar dichos datos.

15 Repositorio con ejemplos



El conocimiento no tiene ningún poder si no puede ser compartido. (José A. Pallavicini)

Según las solicitudes de nuestros usuarios experimentamos con diferentes aplicaciones y librerías e intentamos documentar ejemplos simples para ellos.

Visita nuestro repositorio en github https://github.com/hpciter/user_codes



GitHub

II. Conceptos HPC

II.1 Slurm

16 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.

16.1 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.

16.2 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 rectricciones 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>todos los usuarios</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 las `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.

16.3 Comandos más comunes en SLURM

Para familiarizarse con el planificador a continuación se detallan los comandos más comunes de slurm. Para obtener mas 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 seccion [Comandos útiles en slurm](#) o en la [página oficial de slurm](#).

II.I.I Gestión de trabajos

17 Ejecutar trabajos en el cluster

17.1 Lanzar una ejecución

Para enviar una trabajo a Slurm se puede hacer de 3 formas diferentes: usando una sesión interactiva, lanzando la aplicación en tiempo real o por medio de un script de ejecución.

En cualquier caso, siempre habrá que informar a Slurm qué recursos se están solicitando.

17.1.1 Ejecución de un trabajo con sesión interactiva

Mediante el comando `salloc` Slurm le dará al usuario un nodo de cómputo donde, de manera interactiva, puede trabajar y ejecutar software cualquier software que necesite. Es una opción muy útil para probar software nuevo o nuevos datos con los que trabajar, sin necesidad de lanzar trabajos a la cola, con el riesgo de que fallen nada más empezar.

- Solicitamos un nodo:

```
salloc -N 1
```

- Solicitamos un nodo de una partición concreta, se le asigna un nombre al trabajo y una duración:

```
salloc -p <partition> -J <job_name> -t <HH:MM:SS> -N 1
```

- Una vez concedido el nodo y el trabajo, recibirás el siguiente mensaje:

```
salloc: Granted job allocation 968978
```

- Y cuando el trabajo comienza:

```
srun: Job 968978 step creation temporarily disabled, retrying
srun: Step created for job 968978
The authenticity of host 'node1404-3.hpc.iter.es (10.0.14.15)' can't be established.
RSA key fingerprint is 6b:6f:2e:33:2c:92:11:d6:d6:9d:aa:55:9d:c1:a9:bb.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'node1404-3.hpc.iter.es,10.0.14.15' (RSA) to the list of known hosts.
xxxxxx@node1404-3.hpc.iter.es's password:
Welcome to node1404-3. Deployment version 0.5-hpc. Deployed Wed May 13 19:58:32 WEST 2022.
```

Info

Una vez tenemos un nodo mediante `salloc`, podemos, desde otra terminal, acceder a dicho nodo por ssh para tener varias sesiones abiertas sobre el mismo nodo y trabajar en varias cosas a la vez.

- Para salir de la sesión interactiva:

```
[xxxxxx@node0101-1 ~]$ exit
```

Una vez salgamos del nodo desde la terminal donde hicimos `salloc`, Slurm liberará el nodo y dará por completado el trabajo. Obviamente, una vez salgamos, se cerrarán de forma automática todas aquellas sesiones extras que hayamos abierto mediante SSH.

17.1.2 Ejecutar un trabajo en tiempo real

Con el comando `srun` es posible lanzar un trabajo a la cola directamente.

```
srun -p <partition> -J <job_name> -t <days-HH:MM:SS> <aplicacion>
```

Más opciones disponibles:

Nombre	Dirección IP
-p <partition>	partición en la que se ejecutarán los trabajos
-N <nodes>	nodos que se solicitan para la ejecución
-n=<num_tasks>	número de tareas
--tasks-per-node=<number>	número de tareas a ejecutar en cada nodo (en combinación con -N)
-J <job_name>	nombre que se le asignará al trabajo
-t <days-HH:MM:SS>	tiempo estimado que durará la ejecución
-d=<type:job_id[:job_id]>	el trabajo depende de otros para ejecutarse (opcional)
-o </path/to/file.out>	fichero para la salida estándar de las ejecuciones
-e </path/to/file.out>	fichero para la salida de errores de las ejecuciones
-D <directory>	directorio en donde se realizará la ejecución
--mail-user=<email>	email al que slurm notificará eventos del trabajo
--mail-type=<eventos>	lista de eventos de notificación

17.1.3 Ejecutar en SLURM mediante un script

El comando `sbatch` envía un trabajo a la cola para ser ejecutado por uno o varios nodos, según los recursos que se hayan especificados.

```
sbatch [-p <partition>] [-J <job_name>] [-t <days-HH:MM:SS>] mi_script.sh
```

La estructura más básica de un script es la siguiente.

```
# !/bin/bash

#SBATCH -J <job_name>
#SBATCH -p <partition>
#SBATCH -N <nodes>
#SBATCH --tasks=<number>
#SBATCH --cpus-per-task=<number>
#SBATCH --constraints=<node architecture> # sandy, ilk (icelake)... architecture
#SBATCH -t <days-HH:MM:SS>
#SBATCH -o <file.out>
#SBATCH -D .
#SBATCH --mail-user=<cuenta_de_correo>
#SBATCH --mail-type=BEGIN,END,FAIL,TIME_LIMIT_50,TIME_LIMIT_80,TIME_LIMIT_90
#####

module purge
module load <modules>

srun <aplicacion>
```

18 Ejecutar un trabajo en la cola

Normalmente, un trabajo es creado mediante un script de lanzamiento (shell script) donde las primeras líneas del archivo deben contener las **directivas SBATCH** así como la línea de comienzo `#!/bin/bash` que indica el interprete de shell a usar.

18.1 Mi primer trabajo con slurm

En el siguiente ejemplo es un sencillo script que le ayude a familiarizarse con ellos.

Simplemente realiza la solicitud de recursos (particiones, nº de nodos, memoria, tiempo máximo de ejecución, directorios de trabajo, ficheros de salida,...), comienza con una ejecución de comandos linux básicos (steps) y espera 120 segundos antes de finalizar el trabajo.

SBATCH_SCRIPT_EXAMPLE.SH

```
#!/bin/bash

#SBATCH --job=mi_primer_test      # Nombre del trabajo
#SBATCH --partition=batch        # Partición (express/sbatch/long/fatnodes)
#SBATCH --nodes=1               # N° de nodos
#SBATCH --mem=30000M            # Memoria solicitada por nodo (30000M or 60000M)
#SBATCH --tasks-per-node=1      # N° tareas por nodo
#SBATCH --constrains=sandy      # sandy, ilk (icelake)... architecture
#SBATCH --time=02:00           # Límite de tiempo

#SBATCH --output=file_%j.log     # Log de salida estandar
#SBATCH --error=file_%j.err      # Log de salida errores
#SBATCH -D .                     # Directorio de trabajo

#SBATCH --mail-user=EMAIL        # Donde será enviado el mail
#SBATCH --mail-type=END,FAIL     # Eventos email
#####

# UN COMENTARIO
echo "Comienza mi script"
pwd
hostname
date
sleep 120
echo "Finaliza mi script"
```

El mismo script se puede escribir con la *opciones corta*:

SBATCH_SCRIPT_EXAMPLE_2.SH

```
#!/bin/bash

#SBATCH -J mi_primer_test           # Nombre del trabajo
#SBATCH -p batch                    # Partición (express/sbatch/long/fatnodes)
#SBATCH -N 1                        # N° de nodos
#SBATCH --mem=30000M                # Memoria solicitada por nodo (30000M or 60000M)
#SBATCH --tasks-per-node=1         # N° tareas por nodo
#SBATCH --constrains=sandy         # sandy, ilk (icelake)... architecture
#SBATCH -t 02:00                   # Límite de tiempo

#SBATCH -o file%j.log               # Log de salida
#SBATCH -e file_%j.err             # Log de salida errores
#SBATCH -D .                        # Directorio de trabajo

#SBATCH --mail-user=EMAIL           # Donde será enviado el mail
#SBATCH --mail-type=END,FAIL       # Eventos email
#####

# UN COMENTARIO
echo "Comienza mi script"
pwd
hostname
date
sleep 120
echo "Finaliza mi script"
```

Guarde el fichero creado anteriormente con un nombre apropiado en un directorio de trabajo creado para ello, por ejemplo `mi_primer_test.sh` o `mi_primer_test.sbatch`.

Info

Hay determinadas opciones que tienen valores ya establecidos por defecto, como es la partición, donde la partición por defecto es la `batch`.

Para ver todas las opciones disponibles sobre las **directivas sbatch** puede visitar la documentación oficial de slurm o ejecutando los siguientes comandos:

```
man sbatch
sbatch --help
sbatch --usage
```

18.1.1 Como lanzar un job con sbatch

Para lanzar el trabajo solo tiene que ejecutar el siguiente comando.

```
sbatch mi_primer_test.sh
or
sbatch mi_primer_test.sbatch
```

18.1.2 Mi segundo trabajo con slurm

En este segundo ejemplo simplemente **cargaremos una aplicación (module)** para poder usarla. Concretamente un módulo de python y ejecutaremos un pequeño script de python.

hello_world.py

```
from time import sleep

print ("Hello World")
sleep(120)
print ("Bye world)
```

Es recomendable crear un entorno virtual para ejecutar python (virtualenv, venv, pyenv, conda environment, pipenv). Por ejemplo para usar `venv`, en el directorio de trabajo ejecutar:

```
module load GCCcore/11.2.0 Python/3.8.6
python3 -m venv /path/workdir/venv
pip install ....
```

- Crear el script de lanzamiento:

mi_python_test.sh

```
#!/bin/bash

#SBATCH --job=mi_python_test          # Nombre del trabajo
#SBATCH --nodes=1                    # N° de nodos
#SBATCH --constrains=sandy           # sandy, ilk (icelake)... architecture
#SBATCH --time=02:00                 # Límite de tiempo

#SBATCH --output=file_%j.log         # Log de salida estandar
#SBATCH --error=file_%j.err          # Log de salida errores
#SBATCH --chdir=.                    # Directorio de trabajo

#SBATCH --mail-user=email            # Donde será enviado el mail
#SBATCH --mail-type=END,FAIL         # Eventos email
#####

# antes de cargar el módulo vemos si ha ejecutable de python3
echo "python3 antes:"; python3 --version

# Carga modulos
module purge
module load GCCcore/11.2.0 Python/3.8.6

echo "python3 despues"; python3 --version
which python3

echo "COMIENZA STEP1"
# Activo el entorno de python
source /path/workdir/venv/bin/activate

# Lanzamos el script de python
python3 hello_world.py

echo "FINALIZA MI SCRIPT"
```


Recuerde que su directorio de trabajo es `.`, que es el mismo directorio donde está script de ejecución y. Si `hello_world.py` está en otro directorio, debe especificar la ruta completa.

Otra opción es usar la **variable de entorno de slurm*** `SLURM_SUBMIT_DIR`. Para lanzar el trabajo:

```
SBATCH mi_python_test.sh
or
SBATCH mi_python_test.sbatch
```

19 Useful Slurm commands

Slurm proporciona una variedad de herramientas que permiten al usuario administrar y comprender sus trabajos. Este tutorial presentará estas herramientas y proporcionará detalles sobre cómo usarlas.

19.1 Encontrar información en la cola de trabajo con squeue

El comando `squeue` es una herramienta que usamos para obtener información sobre los trabajos que están en la cola. De manera predeterminada, el comando `squeue` imprimirá el ID del trabajo, la partición, el nombre del trabajo, el usuario del trabajo, el estado del trabajo, el tiempo que lleva en ejecución, la cantidad de nodos y la lista de nodos asignados:

```
squeue
JOBID PARTITION NAME USER ST TIME NODES NODELIST(REASON)
111111 batch my_job myuser R 1:21:59 1 node0101-1
```

Podemos generar información no abreviada con la marca `--long`. Esta bandera imprimirá la información predeterminada no abreviada con la adición de un campo de límite de tiempo:

```
squeue -l
squeue --long
```

El comando `squeue` también brinda a los usuarios un medio para calcular la hora de inicio estimada de un trabajo agregando el indicador `--start` a nuestro comando. Esto agregará la hora de inicio estimada de Slurm para cada trabajo en nuestra información de salida.

```
squeue --user=username --start
```

Nota

La hora de inicio proporcionada por este comando puede ser inexacta. Esto se debe a que la hora calculada se basa en los trabajos en cola o en ejecución en el sistema. Si un trabajo con una prioridad más alta se pone en cola después de ejecutar el comando, su trabajo puede ser demorado.

Al verificar el estado de un trabajo, es posible que desee llamar repetidamente al comando `squeue` para buscar actualizaciones. Podemos lograr esto agregando el indicador `--iterate` a nuestro comando `squeue`. Esto ejecutará `squeue` cada `n` segundos, lo que permite una actualización frecuente y continua de la información de la cola sin necesidad de llamar repetidamente a `squeue`:

```
squeue --start --iterate=n_seconds
```

Presione `ctrl-c` para detener el bucle del comando y regresar a la terminal.

19.1.1 Estado de los trabajos

Una vez que se ha enviado un trabajo a una cola de trabajos, la ejecución seguirá estos estados:

- **PENDING** o **PD**: El trabajo ha entrado en la cola pero aún no están disponibles los recursos solicitados para que empiece a trabajar, es decir, que no haya nodos libres.
- **RUNNING** o **R**: El trabajo se está ejecutando en la cola con los recursos que se han solicitado.
- **COMPLETED** o **CD**: El trabajo se ha ejecutado correctamente, o al menos, lo que se ha especificado en el script de lanzamiento.
- **COMPLETING** o **CG**: El trabajo está en proceso de terminar en un estado correcto.
- **SUSPEND** o **S**: la ejecución del trabajo se ha suspendido y los recursos utilizados se han liberado para otros trabajos.
- **CANCELLED** o **CA**: el trabajo ha sido cancelado o por el usuario o por los administradores de sistemas.
- **FAILED** o **F**: Ha fallado la ejecución del trabajo.
- **NODE_FAIL** o **NF**: Ocurrió un error con el nodo y el trabajo no pudo lanzarse. Por defecto, Slurm relanza el trabajo de nuevo.

19.1.1.1 Razones de que un trabajo esté en PENDING

Cuando un trabajo se encuentra en el estado de **PENDING**, se añade la razón de porqué está pendiente para ejecutar y ésta puede ser:

- **(Resources)**: el trabajo está esperando hasta que estén disponibles los recursos solicitados.
- **(Dependency)**: el trabajo es dependiente de otro y, por tanto, no entrará a ejecutar hasta que se cumpla la condición establecida para la dependencia.
- **(DependencyNeverSatisfied)**: El trabajo está esperando por una dependencia que no ha sido cumplida. El trabajo se quedara en este estado para siempre, por tanto, hay que cancelar el trabajo.
- **(AssocGrpCpuLimit)**: El trabajo no se puede ejecutar porque se ha consumido la cuota asignada de CPU.
- **(AssocGrpJobsLimit)**: El trabajo no se puede ejecutar porque se ha alcanzado el límite de trabajos simultáneos que tiene permitido ejecutar el usuario o la cuenta.
- **(ReqNodeNotAvail)**: El nodo especificado no está disponible. Puede ser que esté en uso, que esté reservado o puede que esté marcado como "fuera de servicio".

Info

Para obtener más información sobre squeue, visite la página de Slurm en [squeue](#)

19.2 Detener trabajos con scancel

En ocasiones, es posible que deba detener un trabajo por completo mientras se está ejecutando. La mejor manera de lograr esto es con el comando `scancel`. Este comando permite cancelar los trabajos que está ejecutando en los recursos de Research Computing utilizando la ID del trabajo. El comando se ve así:

```
scancel your_job-id
```

Para cancelar varios trabajos, puede usar una lista de ID de trabajos separados por comas:

```
scancel your_job-id1, your_job-id2, your_jobid3
```

Info

Para obtener más información, visite el manual de Slurm en [scancel](#)

19.3 Información de estado con sstat

El comando `sstat` permite a los usuarios obtener fácilmente información sobre el estado de sus trabajos actualmente en ejecución. Esto incluye información sobre el uso de la CPU, información de la tarea, información del nodo, tamaño del conjunto residente (RSS) y memoria virtual (VM). Podemos invocar el comando `sstat` como tal:

```
sstat --jobs=your_job-id
```

19.3.1 Formatear la salida del sstat

De forma predeterminada, `sstat` extraerá mucha más información de la que se necesitaría en la salida predeterminada de los comandos. Para remediar esto, podemos usar el indicador `--format` para elegir lo que queremos en nuestra salida. El indicador de formato toma una lista de variables separadas por comas que especifican los datos de salida:

```
sstat --jobs=your_job-id --format=var_1,var_2, ... , var_N
```

Algunas de estas variables se enumeran en la siguiente tabla:

Variable	Descripción
avecpu	Promedio de tiempo de CPU de todas las tareas en el trabajo.
averss	Tamaño medio del conjunto residente de todas las tareas.
avevmsize	Memoria virtual promedio de todas las tareas en un trabajo.
jobid	ID del trabajo.
maxrss	Número máximo de bytes leídos por todas las tareas del trabajo.
maxvsize	Número máximo de bytes escritos por todas las tareas en el trabajo.

Variable	Descripción
ntasks	Número de tareas en un trabajo.

Por ejemplo, imprima la identificación de trabajo promedio de un trabajo, el tiempo de CPU, el máximo de rss y la cantidad de tareas. Podemos hacer esto escribiendo el comando:

```
sstat --jobs=your_job-id --format=jobid,cputime,maxrss,ntasks
```

Info

Se puede encontrar una lista completa de variables que especifican datos manejados por sstat con el indicador `--helpformat` o visitando la página de slurm en [sstat](#).

19.4 Analizar trabajos terminados con sacct

El comando `sacct` permite a los usuarios obtener información sobre el estado de los trabajos terminados. Este comando es muy similar a `sstat`, pero se usa en trabajos que se ejecutaron previamente en el sistema en lugar de los trabajos que se están ejecutando actualmente. Podemos usar la identificación de un trabajo.

- Para todos los trabajos ejecutados:

```
sacct
```

- Para un único trabajo, identificado por su ID:

```
sacct --jobs=your_job_id
```

De forma predeterminada, `sacct` solo extraerá los trabajos que han ejecutado en **el día en curso**. Podemos usar el indicador `--starttime` para decirle al comando que mire más allá de su caché de trabajos a corto plazo.

```
sacct --jobs=your_job-id --starttime=YYYY-MM-DD
```

Para ver una versión no abreviada de la salida de `sacct`, use el indicador `--long`:

```
sacct --jobs=your_job-id --starttime=YYYY-MM-DD --long
```

19.4.1 Formatear la salida del sacct

Al igual que `sstat`, es posible que la salida estándar no proporcione la información que queremos. Para remediar esto, podemos usar el indicador `--format` para elegir lo que queremos en nuestra salida. De manera similar, el indicador de formato es manejado por una lista de variables separadas por comas que especifican los datos de salida:

```
sacct --user=your_rc-username --format=var_1,var_2, ... ,var_N
```

A continuación se proporciona una lista de algunas variables:

Variable	Description
account	Cuenta con la que se ejecutó el trabajo
avecpu	Tiempo promedio de CPU de todas las tareas en el trabajo.
averss	Average resident set size of all tasks in the job.
cputime	Tiempo transcurrido de CPU usado por un job o paso
elapsed	Tiempo transcurrido de los trabajos con formato DD-HH:MM:SS
exitcode	El código de salida devuelto por el script de trabajo o salloc.
jobid	ID del trabajo.
jobname	Nombre del trabajo.
maxdiskread	Número máximo de bytes leídos por todas las tareas.
maxdiskwrite	Número máximo de bytes leídos por todas las tareas.
maxrss	El código de salida devuelto por el script de trabajo o salloc.
ncpus	Cantidad de CPU asignadas.
nnodes	Número de nodos usados.
ntasks	Número de tareas en un job.
priority	Prioridad Slurm.
qos	Calidad de servicio.
reqcpu	Número de CPUs solicitados
reqmem	Cantidad de memoria necesaria para un trabajo
user	Nombre de usuario de la persona que ejecutó el trabajo.

Como ejemplo, suponga que desea buscar información sobre los trabajos que se ejecutaron el 12 de marzo de 2018. Desea mostrar información sobre el nombre del trabajo, la cantidad de nodos utilizados en el trabajo, la cantidad de CPU, el maxrss y el tiempo transcurrido. Su comando se vería así:

```
sacct --starttime=2018-03-12 --format=jobname,nnodes,ncpus,maxrss,elapsed
```

Como otro ejemplo, suponga que desea obtener información sobre los trabajos que se ejecutaron el 21 de febrero de 2018. Le gustaría obtener información sobre la identificación del trabajo, el nombre del trabajo, la calidad del servicio, la cantidad de nodos utilizados, la cantidad de CPU utilizadas, el RSS máximo y la CPU. tiempo, tiempo promedio de CPU y tiempo transcurrido. Su comando se vería así:

```
sacct --starttime=2018-02-21 --format=jobid,jobname,qos,nnodes,ncpu,maxrss,cputime,avecpu,elapsed
```

Info

Se puede encontrar una lista completa de variables que especifican datos manejados por sacct con el indicador `--helpformat` o visitando la página de slurm en [sacct](#).

19.5 Control de trabajos en cola y en ejecución mediante scontrol

El comando `scontrol` proporciona a los usuarios un mayor control de sus trabajos ejecutados a través de Slurm. Esto incluye acciones como suspender un trabajo, detener la ejecución de un trabajo o extraer información detallada sobre el estado de los trabajos.

Para suspender un trabajo que se está ejecutando actualmente en el sistema, podemos usar `scontrol` con el comando `suspend`. Esto detendrá un trabajo en ejecución en su paso actual que se puede reanudar en un momento posterior. Podemos suspender un trabajo escribiendo el comando:

```
scontrol suspend job_id
```

Para reanudar un trabajo en pausa, usamos `scontrol` con el comando `resume`:

```
scontrol resume job_id
```

Slurm también proporciona una utilidad para retener trabajos que están en cola en el sistema. Retener un trabajo colocará el trabajo en la prioridad más baja, efectivamente "reteniendo" el trabajo para que no se ejecute. Un trabajo solo se puede retener si está esperando que el sistema se ejecute. Usamos el comando de espera para poner un trabajo en estado de espera:

```
scontrol hold job_id
```

Luego podemos liberar un trabajo retenido usando el comando de `release`:

```
scontrol release job_id
```

scontrol también puede proporcionar información sobre trabajos mediante el comando `show job`. La información proporcionada por este comando es bastante extensa y detallada, así que asegúrese de borrar la ventana de su terminal, recopilar cierta información del comando o canalizar la salida a un archivo de texto separado:

```
scontrol show job job_id
```

19.5.1 Streaming de salida a un archivo de texto

```
scontrol show job job_id > outputfile.txt
```

19.5.2 Canalizar la salida a Grep y encontrar líneas que contengan la palabra "Tiempo"

```
scontrol show job job_id | grep Time
```


20 Variables de entorno de salida en SLURM

Dentro de los script de ejecución se pueden invocar ciertas variables con las que podemos conocer cierta información de la ejecución dentro un script.

Documentación para variables de entorno de slurm

Las más usadas son las siguiente:

Variable	
SLURM_ARRAY_JOB_ID	Job array's master job ID number.
SLURM_JOB_ID	The ID of the job allocation.
SLURM_JOBID	New version for the ID of the job allocation
SLURM_JOB_DEPENDENCY	Set to value of the --dependency option
SLURM_JOB_NAME	Name of the job.
SLURM_JOB_NODELIST	List of nodes allocated to the job
SLURM_JOB_NUM_NODES	Total number of nodes in the job's resource allocation
SLURM_JOB_PARTITION	Name of the partition in which the job is running
SLURM_NODELIST	List of nodes allocated to the job
SLURM_SUBMIT_DIR	The directory from which sbatch was invoked

20.1 Script para ver las variables de slurm.

Ejecutando el siguiente script usted puede ver las variables nombradas en el apartado anterior:

```
#!/bin/bash
#SBATCH -J GNUParallel -o %x-%J.out
#SBATCH --time=00:10:00
#SBATCH --mem-per-cpu=2G
#SBATCH -n 16
#SBATCH -c 4
# SBATCH --ntasks-per-node=8
#SBATCH --constrains=<node architecture> # sandy, ilk (icelake)... architecture

date
echo "SUBMITTED ON: $SLURM_SUBMIT_HOST IP: $SLURM_LAUNCH_NODE_IPADDR DIR: $SLURM_SUBMIT_DIR NODES
ALLOCATED: $SLURM_JOB_NODELIST"
echo "RUNNING ON: $(hostname) $SLURMD_NODENAME"
echo "JOB_ID: $SLURM_JOB_ID ARRAY_JOB_ID: $SLURM_ARRAY_JOB_ID"
echo "STEP: $SLURM_STEP_ID NODEID: $SLURM_NODEID LOCALID: $SLURM_LOCALID PROCID: $SLURM_PROCID"
echo "ARRAY_TASK_COUNT: $SLURM_ARRAY_TASK_COUNT ARRAY_TASK_ID: $SLURM_ARRAY_TASK_ID ARRAY_TASK_MAX:
$SLURM_ARRAY_TASK_MAX ARRAY_TASK_MIN: $SLURM_ARRAY_TASK_MIN ARRAY_STEPSIZE: $SLURM_ARRAY_TASK_STEP"
echo "Args: $*"
sleep 10
```

21 Investigar un trabajo fallido

No siempre los trabajos se ejecutan correctamente. Hay una lista de motivos por los que los trabajos o las aplicaciones se detienen o fallan. Las causas más comunes son:

- Exceder los límites de recursos solicitados, como la memoria o el tiempo límite, por ejemplo.
- Errores específicos del software.
- Errores en el procesamiento de los datos.

Nota

Es importante recopilar mensajes de error/salida ya sea escribiendo dicha información en la ubicación predeterminada o especificando ubicaciones específicas mediante las opciones `--error/--output`, por ejemplo, aunque esto dependerá de cada software.

Los mensajes de error y de salida son el punto de partida para investigar un error en el trabajo.

21.1 Exceder los límites de recursos

Cada partición define límites máximos y predeterminados para el tiempo de ejecución y el uso de la memoria. Dentro de la especificación del trabajo, los límites actuales se pueden definir dentro de los rangos. Para una mejor programación, se deben estimar los requisitos del trabajo y adaptar los límites a las necesidades. Cuanto más bajos sean los límites mejor, así SLURM podrá encontrar un lugar. Además, cuanto menos gastos generales de recursos se especifican, menos recursos se desperdician, como por ejemplo, con la memoria.

Info

Si un trabajo excede el tiempo de ejecución o el límite de memoria, SLURM matará el trabajo.

21.1.1 Información de error

En ambos casos, el archivo de errores proporciona información adecuada:

- Tiempo límite:

```
(...)  
slurmstepd: error: *** JOB 41239 ON fnode01 CANCELLED AT 2016-11-30T11:22:57 DUE TO TIME LIMIT ***  
(...)
```

- Límite de memoria:

```
(...)
slurmstepd: error: Job 41176 exceeded memory limit (3940736 > 2068480), being killed
slurmstepd: error: Exceeded job memory limit
slurmstepd: error: *** JOB 41176 ON fnode01 CANCELLED AT 2016-11-30T10:21:37 ***
(...)
```

21.2 Errores de Software

Slurm captura el código de salida de un trabajo y lo guarda como parte del registro del trabajo. **Para trabajos lanzados con sbatch, se captura el código de salida del script. Para trabajos lanzados con srun, el código de salida será el valor de retorno del comando ejecutado. Cualquier código de salida distinto de cero se considera un error de trabajo y da como resultado un estado de trabajo de FAILED.**

Si una señal es responsable de la terminación de un trabajo/paso, el número de la señal también se capturará y se mostrará después del código de salida (separado por dos puntos).

Según el orden de ejecución de los comandos en el script sbatch, es posible que un comando específico falle, pero el script por lotes devolverá cero, lo que indica que se realizó correctamente. Por ejemplo:

fail.R

```
var<-sq(1,1000000000)
```

job_sbatch.sh

```
#!/bin/bash

#SBATCH --job=my_r_test           # Nombre del trabajo
#SBATCH --nodes=1                # N° de nodos
#SBATCH --constrains=sandy       # sandy, ilk (icelake)... architecture
#SBATCH --time=02:00            # Límite de tiempo

#SBATCH --output=file_%j.log     # Log de salida estandar
#SBATCH --error=file_%j.err      # Log de salida errores
#SBATCH --chdir=.                # Directorio de trabajo

#SBATCH --mail-user=email        # Donde será enviado el mail
#SBATCH --mail-type=END,FAIL     # Eventos email
#####

Rscript fail.R
echo "Script finished"
```

El código de salida y el estado indican incorrectamente que el trabajo finalizó correctamente:

```
sbatch job_sbatch.sh
Submitted batch job 41585
```

```
sacct -j 41585
```

JobID	JobName	Partition	Account	AllocCPUS	State	ExitCode
41585	Simple E +	all	id	1	COMPLETED	0:0
41585.batch	batch		id	1	COMPLETED	0:0

Solo el archivo de salida específico de R muestra el error:

fail.Rout

```
(...)
> var<-sq(1,1000000000)
Error: could not find function "sq"
Execution halted
```

Puede evitar este problema "saliendo" con un código de salida adecuado tan pronto como falle el comando:

job_sbatch.sh

```
#!/bin/bash

#SBATCH --job=my_r_test           # Nombre del trabajo
#SBATCH --nodes=1                # N° de nodos
#SBATCH --constrains=sandy       # sandy, ilk (icelake)... architecture
#SBATCH --time=02:00             # Límite de tiempo

#SBATCH --output=file_%j.log     # Log de salida estandar
#SBATCH --error=file_%j.err      # Log de salida errores
#SBATCH --chdir=.                # Directorio de trabajo

#SBATCH --mail-user=email        # Donde será enviado el mail
#SBATCH --mail-type=END,FAIL     # Eventos email
#####

Rscript fail.R || exit 91
echo "Script finished"
```

Ahora, el código de salida y el estado coinciden con el verdadero resultado:

```
sbatch job_sbatch.sh
Submitted batch job 41925
```

```
sacct -j 41925
```

JobID	JobName	Partition	Account	AllocCPUS	State	ExitCode
41925	Simple E +	all	id	1	FAILED	91:0
41925.batch	batch		id	1	FAILED	91:0

Compruebe siempre los archivos de salida específicos de la aplicación en busca de mensajes de error.

22 Límites de memoria Slurm

Slurm impone un límite de memoria en cada trabajo. De forma predeterminada, es deliberadamente relativamente pequeño: 2 GB por nodo. Si su trabajo usa más que eso, recibirá un error que indica que su trabajo ha excedido la memoria solicitada (*Exceeded job memory limit*).

Para establecer un límite mayor, agregue a su envío de trabajo:

```
#SBATCH --mem X
```

donde X es la cantidad máxima de memoria que usará su trabajo por nodo, en MB.

Cuanto mayor sea su conjunto de datos de trabajo, mayor deberá ser, pero cuanto menor sea el número, más fácil será para slurm encontrar un lugar para ejecutar su trabajo.

Para determinar un valor apropiado, consulte [siguiente sección](#).

22.1 Cómo estudiar la eficiencia de un job

Muchas veces nos preguntamos cómo puedo saber qué tan eficiente es mi trabajo y para ello tenemos el comando **seff**.

```
seff JOBID
```

donde el JOBID es el id del trabajo en el que estás interesado. Por ejemplo:

```
$ seff 1553

Job ID: 1553
Cluster: teide
User/Group: viddata/viddata
State: COMPLETED (exit code 0)
Nodes: 4
Cores per node: 16
CPU Utilized: 2-11:55:56
CPU Efficiency: 89.73% of 2-18:47:28 core-walltime
Job Wall-clock time: 01:02:37
Memory Utilized: 26.08 GB (estimated maximum)
Memory Efficiency: 23.85% of 109.38 GB (27.34 GB/node)
```



Solicite memoria un poco mayor a lo que seff dice.

Debe configurar la memoria que solicita en algo un poco más grande que lo que dice seff.

i seff informa del máximo de memoria usada en jobs paralelos.

Tenga en cuenta que **para trabajos paralelos que abarcan varios nodos, esta es la memoria máxima utilizada en cualquier nodo**; si no está configurando una distribución uniforme de tareas por nodo (por ejemplo, con `--ntasks-per-node`), el mismo trabajo podría tener valores muy diferentes cuando se ejecuta en diferentes momentos.

i Espera a que el trabajo termine satisfactoriamente.

También tenga en cuenta que el número registrado por slurm para el uso de la memoria será inexacto si el trabajo finalizó de forma no satisfactoria. **Para obtener una medición precisa, debe tener un trabajo que se complete correctamente, ya que slurm registrará el pico de memoria real.**

22.1.1 Memoria máxima por tipo de nodo

Tipo de nodo	Solicitud máxima de memoria en slurm	
Sandy bridge 16 Cores - 32 GB	30000 MB	
Sandy bridge 16 Cores - 64 GB	62000 MB	
Fat Nodes 32 Cores - 256 GB	254000 MB	bajo solicitud
Icelake Nodos GPU 64 Cores - 256 GB	254000 MB	bajo solicitud

22.2 ¿Cómo puedo saber qué tan eficiente es mi trabajo?

Puede ver la eficiencia de su trabajo comparando MaxRSS, MaxVMSize, Elapsed, CPUtime, NCPUS con el comando `sacct`.

```
sacct -j 999997 --
format=User,JobID,Jobname%25,partition,elapsed,MaxRss,MaxVMSize,nnodes,ncpus,nodelist%20
```

User	JobID	JobName	Partition	Elapsed	MaxRSS	MaxVMSize	NNodes
NCPUS	NodeList						
eolicase	999996	WRF_2023080618	batch	13:32:47			4
64	node1511-[1-4]						
	999996.batch	batch		13:32:47	216700K	815628K	1
16	node1511-1						
	999996.0	geogrid.exe		00:00:57	236240K	2762564K	4
8	node1511-[1-4]						
	999996.1	metgrid.exe		00:01:54	15395M	591912K	4
4	node1511-[1-4]						
	999996.2	real.exe		00:00:26	338148K	2928876K	4
16	node1511-[1-4]						
	999996.3	wrf.exe		13:26:36	598864K	3768748K	4
64	node1511-[1-4]						

En este trabajo, ve que el usuario usó 64 núcleos y su trabajo se ejecutó durante 13,5 horas. Sin embargo, su tiempo de CPU es de 13,5 horas, que está cerca de las 64×13 horas. Si su código escala efectivamente según esta fórmula $CPU\ Time = NCPUS * Elapsed$ su aplicación escala perfectamente. Si no lo es, el resultado divergirá de la fórmula y la mejor manera de probar esto y determinar cómo escala la aplicación es hacer algunas pruebas de escala.


Hay dos formas de hacer esto: **Strong scaling** (escalado fuerte) and **Weak scaling** (escalado débil).

22.2.0.1 Strong scaling (Escalado fuerte)

El escalado fuerte es donde deja el tamaño del problema igual pero aumenta la cantidad de núcleos. Si su código se escala bien, debería tomar menos tiempo proporcional a la cantidad de núcleos que usa.

22.2.0.2 Weak scaling (escalado débil)

La cantidad de trabajo por núcleo sigue siendo la misma, pero aumenta la cantidad de núcleos, por lo que el tamaño del trabajo se escala proporcionalmente a la cantidad de núcleos. Por lo tanto, si su código se escala en este caso, el tiempo de ejecución debería seguir siendo el mismo.

 **Por lo general, la mayoría de los códigos tienen un punto en el que la escala se rompe debido a ineficiencias en el código.**

Por lo tanto, más allá de ese punto, no hay ningún beneficio en aumentar la cantidad de núcleos que arroja al problema. Ese es el punto que quieres buscar. Esto se ve más fácilmente trazando el registro de la cantidad de núcleos frente al registro del tiempo de ejecución.

II.I.II Temas avanzados

23 Ejecuciones secuenciales

La unidad de reserva mínima para ejecutar en TeideHPC es el nodo, por lo que, independientemente del trabajo a ejecutar, a la hora de realizar el accounting del mismo, se considerará que se ha hecho uso de nodos completos. Es decir, **el cómputo de horas de uso de la infraestructura viene dado por el número de horas que cada nodo está ejecutando el trabajo en cuestión, se use o no todos los recursos disponibles del mismo.**

Para que el usuario no se vea perjudicado por este hecho, habrá de estructurar sus ejecuciones de manera que pueda agruparlas en igual cantidad de número de cores al de los nodos (16).

Las ejecuciones secuenciales suelen requerir muchas ejecuciones de trabajos a pocos cores, por lo que el usuario deberá estructurar los datos de entrada para las ejecuciones en carpetas o nombres de ficheros identificados mediante un número de trabajo, de manera que éste se pueda manipular de manera sencilla dentro del script de lanzamiento.

Una vez se tengan los datos de entradas organizados de esta manera se puede proceder a la ejecución de la misma, lanzando en un script de submit tantas tareas como quepan en un nodo.

23.1 Ejemplo de ejecución secuencial de trabajos de 1 core

El siguiente script de ejemplo lanza 16 trabajos de 1 core simultáneamente.

```
# !/bin/bash

#SBATCH -J <job_name>
#SBATCH -p <partition>
#SBATCH -N 1
#SBATCH --constrains=<node architecture> # sandy, ilk (icelake)... architecture
#SBATCH -t <days-HH:MM:SS>
#SBATCH -o <file.out>
#SBATCH -D .
#####

module purge
module load <modules>

for i in {1..16}; do
    srun --ntasks=1 --exclusive --output=slurm-%J.out serial-program --input /path/to/input.$i &
done
# wait until all background processes are ended
wait
```

El comando `srun` lanzará los trabajos individuales en diferentes cores. El símbolo `&` al final de la línea significa que el trabajo se ejecute en segundo plano, de manera que no haya que esperar para lanzar la siguiente tarea. El comando `wait` del final evita que el trabajo termine hasta que no hayan terminado todos los procesos anteriores.

23.2 Ejemplo ejecución secuencial trabajos de 4 cores

El siguiente script de ejemplo lanza 4 trabajos de 4 cores simultáneamente lo que hace un total de 16 cores utilizados:

```
# !/bin/bash

#SBATCH -J <job_name>
#SBATCH -p <partition>
#SBATCH -N 1
#SBATCH --constrains=<node architecture> # sandy, ilk (icelake)... architecture
#SBATCH -t <days-HH:MM:SS>
#SBATCH -o <file.out>
#SBATCH -D .
#####

module purge
module load <modules>

srun --ntasks=4 --exclusive --output=slurm-%J.out serial-program --input /path/to/input_1 &
srun --ntasks=4 --exclusive --output=slurm-%J.out serial-program --input /path/to/input_2 &
srun --ntasks=4 --exclusive --output=slurm-%J.out serial-program --input /path/to/input_3 &
srun --ntasks=4 --exclusive --output=slurm-%J.out serial-program --input /path/to/input_4 &

# wait until all background processes are ended
wait
```

24 Ejecución de arrays

La ejecución de arrays requerirá de una preparación previa por parte del usuario.

El usuario deberá estructurar los datos de entrada para las ejecuciones en carpetas o nombres de ficheros identificados mediante un número de trabajo, de manera que éste se pueda manipular de manera sencilla dentro del script de lanzamiento.

Una vez se tengan los datos de entrada organizados de esta manera, se puede proceder a la ejecución de los trabajos, lanzando en un script de submit tantas tareas como se necesiten.

24.1 Ejecución de un array de 10 trabajos a un nodo

```
# !/bin/bash

#SBATCH -J <job_name>
#SBATCH -p <partition>
#SBATCH -N 1
#SBATCH -N <nodes>
#SBATCH --task=<number>
#SBATCH --cpus-per-task=<number>
#SBATCH --constrains=<node arquitectura> # sandy, ilk (icelake)... arquitectura
#SBATCH -t <days-HH:MM:SS>
#SBATCH --array=0-10
## %A JobId de la tarea principal (script de lanzamiento del arrayjob)
## %a Índice de la tarea dentro del arrayjob
## %j JobId de la tarea dentro del arrayjob en el gestor de colas
#SBATCH -o <array-%A-%a-%j.out>
#SBATCH -D .
#SBATCH --mail-type END
#SBATCH --mail-user <mail to notify>
#####

module purge
module load <modules>

echo "Array job $SLURM_ARRAY_JOB_ID -> $SLURM_ARRAY_TASK_ID"
srun program
echo "End array job $SLURM_ARRAY_JOB_ID -> $SLURM_ARRAY_TASK_ID"
```

24.2 Ejecución de un array de 100 trabajos con sólo 10 trabajos simultáneos

```
# !/bin/bash

#SBATCH -J <job_name>
#SBATCH -p <partition>
#SBATCH -N 1
#SBATCH -t <days-HH:MM:SS>
#SBATCH -array=0-100%10
## %A JobId de la tarea principal (script de lanzamiento del arrayjob)
## %a Índice de la tarea dentro del arrayjob
## %j JobId de la tarea dentro del arrayjob en el gestor de colas
#SBATCH -o <array-%A-%a-%j.out>
#SBATCH -D .
#SBATCH --mail-type END
#SBATCH --mail-user <mail to notify>
#####

module purge
module load <modules>

echo "Array job $SLURM_ARRAY_JOB_ID -> $SLURM_ARRAY_TASK_ID"
srun program
echo "End array job $SLURM_ARRAY_JOB_ID -> $SLURM_ARRAY_TASK_ID"
```

25 Trabajos dependientes

Esta página describe la función de dependencia de trabajos en SLURM.

Es una función se utiliza cuando se encadenan varios trabajos dependientes unos de otros. Por ejemplo:

- Un trabajo de preprocesamiento con 1 núcleo debe ser seguido por una simulación con 16 núcleos. Los resultados deben procesarse posteriormente con un solo trabajo principal.
- Un trabajo de procesamiento posterior debe enviarse después de que finalicen todas las tareas de una matriz de trabajo.

25.1 Uso

El trabajo dependiente debe lanzarse utilizando el comando de `sbatch` especificando la opción `--dependency=<type>:<listOfJobIDs>`. El tipo puede ser:

- `after`: el trabajo se lanza después de que el primer trabajo entre a ejecutar, es decir, que esté RUNNING.
- `afterok`: el trabajo se lanza después de que el primer trabajo haya terminado con éxito.
- `afterany`: el trabajo se lanza después de que el primer trabajo haya terminado, independientemente de si ha fallado o no.
- `afternotok`: el trabajo se lanza después de que el primer trabajo haya terminado en fallo.
- `aftercorr`: una tarea de un array de trabajos puede empezar a ejecutarse después de que el ID de otro trabajo del array haya finalizado con éxito.
- `singleton`: el trabajo puede empezar a ejecutarse cuando cualquier otro trabajo con el mismo nombre y del mismo usuario haya terminado.

Ejemplo

```
sbatch --dependency=afterok:111111,111112 my_job_script.sh
```

El trabajo subyacente (del que depende este trabajo) debe enviarse primero. El ID de trabajo relacionado se puede capturar recopilando la salida del comando `sbatch` con la opción `--parsable`:

```
first_jobid=$(sbatch --parsable my_first_job.sh)
```

Info

[más información, en la página oficial](#)

26 Cómo Trabajar con la partición /local

Cada uno de los nodos de cómputo dispone de una partición de 300GB, llamada `/local`, que el usuario puede utilizar mientras está utilizando el nodo. Cuando un nodo es asignado a un usuario, en esta partición se crea una carpeta con el ID del trabajo y cuyo propietario es el usuario, de esta forma, en este directorio el usuario puede crear, modificar o borrar ficheros.

El uso de esta partición no está restringido a ninguna cola, se puede utilizar incluso cuando se utiliza el nodo en modo interactivo, a través del comando `sbatch`.

Note

Una vez termine por completo la ejecución, se eliminará el contenido de esta partición, por lo que se deben mover los datos resultantes de la ejecución a, por ejemplo, la partición `/data`, una partición destinada para que cada usuario pueda trabajar.

26.1 Ejemplo de ejecución utilizando la partición /local

```
#!/bin/bash

#SBATCH -J <job_name>
#SBATCH -p <partition>
#SBATCH -N 1
#SBATCH --constrains=<node architecture> # sandy, ilk (icelake)... architecture
#SBATCH -o <out.log>
#SBATCH -e <error.log>
#SBATCH -D .

#####

module purge
module load <modules>

# Copy the data to /local
cp file1 file2 file3 ... /local/$SLURM_JOBID/

# Move into the /local directory
cd /local/$SLURM_JOBID

# Run software
...

# Once everything is finished, we move the results to our working directory
mv result1 result2 result3 ... ~/data/work_dir/
```

27 Paralelismo

Si está trabajando en problemas complejos y de gran escala que requieren informática de alto rendimiento (HPC), es posible que se pregunte cómo elegir la mejor arquitectura para sus necesidades.

Los sistemas HPC se pueden clasificar en dos tipos principales: **memoria compartida (Shared memory)** y sistemas de **memoria distribuida (Distributed memory)**.

27.1 Paralelismo de memoria compartida (threading)

En el *Paralelismo de memoria compartida* o *shared-memory parallelism (SM)*, las aplicaciones logran el paralelismo al ejecutar más de un subproceso a la vez en los núcleos dentro de un nodo. Cada uno de los subprocesos ve y manipula la misma memoria que asignó el proceso inicial. Los subprocesos son ligeros y rápidos y procesan de forma independiente partes del trabajo. Con el paralelismo SM, **los trabajos están limitados a la cantidad total de memoria y núcleos en un nodo**.

Para ejecutar una aplicación con subprocesos correctamente a través de Slurm, deberá especificar una serie de restricciones de Slurm. Por ejemplo, para iniciar un trabajo de Slurm para su aplicación con subprocesos que utilizará 16 subprocesos, haga lo siguiente:

```
#!/bin/bash
#SBATCH -J my_job
# Numero de Nodos y cores por nodo
#SBATCH -N 1
#SBATCH --cpu-per-task=16
#SBATCH --mem=45000
#SBATCH --partition=batch

module load GCC/12.2.0
srun ./myapp
```

Si opta por no usar `srun` para iniciar su aplicación, debe agregar esto a su script de trabajo de Slurm: `export OMP_NUM_THREADS=16` Este número debe coincidir con `--cpu-per-task`.

27.2 Paralelismo de memoria distribuida (MPI)

En el *paralelismo de memoria de distribuida* o *distributed-memory parallelism (DM)*, una aplicación logra el paralelismo al ejecutar varias instancias de sí misma en varios nodos para resolver el problema. A cada instancia se le asigna su propia porción de memoria virtual y se comunica con otras instancias a través de una interfaz de paso de mensajes como MPI.

Por ejemplo, para usar 4 nodos y 16 procesadores por nodo,


```
#!/bin/bash
#SBATCH -J my_job
#SBATCH -N 4
#SBATCH --tasks=16
#SBATCH --mem=28000
#SBATCH --partition=batch

module load intel/2021b intel-compilers/2021.4.0 impi/2021.4.0
srun ./myapp
```

27.3 Paralelismo híbrido: DM + SM


En el paralelismo híbrido, las aplicaciones logran el paralelismo con el uso de subprocesos y tareas MPI. Este tipo de paralelización combina las características de los dos modelos anteriores, lo que permite que un programa ejecute subprocesos en varios nodos del clúster.

Los trabajos híbridos pueden ejecutarse potencialmente de manera más eficiente (consumiendo menos memoria y escalando más) al reducir la sobrecarga de comunicación MPI de un trabajo. Esto se puede hacer sustituyendo hilos livianos por rangos MPI.

Para configurar trabajos para usar DM+SM, primero asegúrese de que su aplicación lo admita.

```
#!/bin/bash
#SBATCH -cpus-per-task=12 #same as -c 12
#SBATCH -ntasks=4 #same as -n 4
#SBATCH -ntasks-per-node=2

module load gOMPI/2022b
srun ./myapp
```

 **Los trabajos de DM pueden afectar el rendimiento debido a los pasos de comunicación de MPI cuando se ejecuta la aplicación en demasiados nodos.**

Por ejemplo, si tiene un trabajo de DM que funciona mejor con 64 cores, puede solicitar que su trabajo se ejecute sólo en 4 nodos con `nodes=4`. Por supuesto, al hacer esto, es probable que su trabajo tarde más en comenzar debido a la restricción adicional del trabajo, pero podría valer la pena para sus análisis. Lo siguiente mejor que puede hacer es pedirle a Slurm que use preferiblemente 4 nodos, pero si no es posible, use 5 o 6. Debe especificar esto con `--nodes=4-6`. Sin especificar `--nodes`, su trabajo de MPI podría ejecutarse en cualquier cantidad de nodos en el clúster.

27.4 OpenMP/Multithreading vs. MPI

Si bien hay varias formas de solicitar una cierta cantidad de núcleos de CPU para su programa, observe la siguiente distinción:

- **Solicitar `ntask` (nº) cores CPU para MPI (procesos distintos)**

Estas tareas se pueden distribuir en varios nodos de cómputo.

```
#SBATCH --ntask= <ntask>
```

- Solicitar *nodes ntask-per-node* (nº) cores CPU para MPI (procesos distintos) en el mismo nodo

Esto distribuye sobre X nodos de cómputo Y tareas.

```
#SBATCH --nodes= <nodes>
#SBATCH --ntask-per-node= <ntas-per-node>
```

- Solicitar *cpu-per-task* (nº) de CPU cores para aplicaciones multithread (eg. OpenMP)

--cpus-per-task especifica cuántas CPU puede usar cada tarea. ¡Estos siempre se asignarán dentro de un sólo nodo de cómputo, nunca a varios nodos!

```
#SBATCH --cpus-per-task= <cpus-per-task>
```

27.4.1 Algunos consejos

Para una aplicación MPI simple, use

```
--ntasks,
esto usa memoria distribuida (entre nodos),
requiere MPI.
```

Para una aplicación simple de OpenMP/multiproceso, use

```
--ntasks=1
--cpus-per-task=<cpu_per-task>
Esto usa Memoria compartida (dentro de un solo nodo)."
```

Para una aplicación híbrida, use:

```
--ntasks=<no of nodes>
--cpus-per-task=<no of cores per node>,
Esto usa SM and DM y requiere MPI.
```

La opción SBATCH *--ntasks-per-core=#* sólo es adecuada para nodos de cómputo que tengan HyperThreading habilitado en hardware/BIOS

En TeideHPC, el HyperThreading está desactivado por defecto, es decir `#SBATCH --ntasks-per-core=1`

Haz tus propios test de escalabilidad

Comience con 2 nodos y estudie los resultados y aumente la cantidad de nodos o tareas poco a poco.

27.4.2 Ejemplo básico: *usar srun or no usarlo*

```
#!/bin/bash
#SBATCH --ntasks=8
## more options
echo hello
```

Esto siempre generará una línea, porque el script sólo se ejecuta en uno de los nodos

```
#!/bin/bash

#SBATCH --ntasks=8

## more options
srun echo hello
```

srun hace que el script ejecute su comando en los nodos trabajadores y, como resultado, debería obtener 8 líneas de saludo.

27.4.3 Ejemplo básico: *Usar ntask y una aplicación de 1 thread.*

- Usando el valor por defecto de `ntasks=1`

```
#!/bin/bash

#SBATCH --ntasks=1

srun sleep 10 &
srun sleep 12 &
wait
```

El número de tareas predeterminado se especificó en 1 y, por lo tanto, la segunda tarea no puede comenzar hasta que la primera tarea haya finalizado. Este trabajo tardará en alrededor de 22 segundos. Desglosando esto:

```
sacct -j1425 --format=JobID,Start,End,Elapsed,NCPUS
```

JobID	Start	End	Elapsed	NCPUS
1425	2023-07-13T20:51:44	2023-07-13T20:52:06	00:00:22	1
1425.batch	2023-07-13T20:51:44	2023-07-13T20:52:06	00:00:22	1
1425.0	2023-07-13T20:51:44	2023-07-13T20:51:56	00:00:12	1
1425.1	2023-07-13T20:51:56	2023-07-13T20:52:06	00:00:10	1

Aquí la tarea 0 comenzó y finalizó (en 12 segundos) seguida de la tarea 1 (en 10 segundos). Para hacer un tiempo de usuario total de 22 segundos.

- Usando `ntasks=2`

```
#!/bin/bash

#SBATCH --ntasks=2

srun --ntasks=1 sleep 10 &
srun --ntasks=1 sleep 12 &
wait
```

Ejecutando el mismo comando `sacct` como se especificó anteriormente:

```
sacct -j 515064 --format=JobID,Start,End,Elapsed,NCPUS
```

JobID	Start	End	Elapsed	NCPUS
515064	2023-07-13T21:34:08	2023-07-13T21:34:20	00:00:12	2
515064.batch	2023-07-13T21:34:08	2023-07-13T21:34:20	00:00:12	2
515064.0	2023-07-13T21:34:08	2023-07-13T21:34:20	00:00:12	1
515064.1	2023-07-13T21:34:08	2023-07-13T21:34:18	00:00:10	1

27.5 Más ejemplos

Visite nuestro repositorio en github https://github.com/hpciter/user_codes.git

28 GPUs en TeideHPC

GPU es el acrónimo de *Graphics Processing Unit* y representa precisamente el corazón de una tarjeta gráfica al igual que la CPU lo hace en un PC. Aparte del corazón, también es su cerebro, ya que es la encargada de realizar todos los cálculos complejos que permiten que algunos programas pueden ejecutarse mucho más rápido que en una CPU.

Entre las principales usos de las GPUs están los siguientes:

- Edición de vídeo
- Renderización de gráficos 3D
- Aprendizaje automático
- Aplicaciones científicas
- etc...

El clúster TeideHPC ofrece 2 modelos diferentes de GPU para usar con sus trabajos. Recomendamos echar un vistazo a la [descripción del clúster](#) para tener una idea de cómo se ve.

28.1 Modelos de GPU disponibles

Estas son las GPU disponibles actualmente en TeideHPC:

modelos	# nodos	# GPU/ nodo	tipo slurm	Cores CPU/ nodo	Memoria CPU/nodo	Capacidad de cómputo (*)	M G
Nvidia A100	16	4	a100	64	256GB	80	
Nvidia A100	1	8	a100	64	512GB	80	
Nvidia Tesla T4	4	1	t4	32	256GB	75	

(*) *Capacidad de cómputo o Compute Capability* es un término técnico creado por NVIDIA como una forma compacta de describir qué funciones de hardware están disponibles en algunos modelos de GPU y no en otros. No es una medida de rendimiento y solo es relevante si está compilando sus propios programas de GPU. Consulte la página sobre [programación CUDA](#) para obtener más información.

28.2 Computación GPU en TeideHPC

El clúster TeideHPC tiene una serie de nodos que tienen conectadas unidades de procesamiento de gráficos de propósito general (GPGPU) de NVIDIA. Es posible usar herramientas CUDA para ejecutar trabajo computacional en ellas y, en algunos casos de uso, ver aceleraciones muy significativas.

Como explicamos en la [sección cómo ejecutar trabajos](#) podemos utilizar 3 formas diferentes de enviar un trabajo a la cola de trabajos: utilizando una sesión interactiva, iniciando la aplicación en tiempo real o medio de un script de ejecución.

28.2.1 Cómo solicitar recursos de GPU

- **Para solicitar nodos de GPU o GPU primero debe comprender cómo está configurado su clúster así como las particiones**, por este motivo, le recomendamos que consulte la página inicial donde tenemos una [descripción del clúster](#) que le puede ayudar.
- Básicamente hay **2 arquitecturas: icelake (nodos de GPU) y sandy (nodos de CPU)**.
- En base a estas arquitecturas se solicitan los recursos dentro del cluster.

La forma de solicitar un nodo de GPU es utilizando la partición dedicada de `gpu`. Para poder utilizar esta partición hay que solicitar acceso. Además de especificar la partición, hay que utilizar las **gres** de SLURM

28.2.2 Recursos genéricos GRES (Generic Resource)

En Slurm, **GRES significa Recurso Genérico**. GRES es una función que le permite especificar y administrar varios tipos de recursos genéricos, como GPU.

La funcionalidad GRES de Slurm permite la asignación, la programación y el seguimiento eficientes de estos recursos para los trabajos enviados. Ayuda a garantizar que los recursos solicitados estén disponibles y sean utilizados correctamente por los trabajos que los requieren.

Para usar GRES de manera efectiva, debe comprender cómo está configurado su clúster, cómo se definen los tipos y cantidades de GRES disponibles y especificar los requisitos de GRES al enviar trabajos.

Para obtener los tipos de GRES definidos en el clúster, puede usar:

```
$ scontrol show config | grep Gres
GresTypes = gpu
```

Para conocer la lista de GRES en TeideHPC mira la columna GRES después de ejecutar este comando.

```
$ sinfo -o "%40N %10c %10m %35f %30G"
```

NODELIST	CPUS	MEMORY	AVAIL_FEATURES
GRES			
node18109-1	64	257214	ilk, gpu, a100
gpu:a100:8			
node2204-[3-4]	20	31906	ivy
(null)			
node17109-1, node17110-1, node18110-1, node	64	257214	ilk, viz, t4
gpu:t4:1			
node0303-2, node0304-[1-4], node1301-[1-4]	16	30000+	sandy
(null)			
node17101-1, node17102-1, node17103-1, node17104-1,	64	257214	ilk, gpu, a100
gpu:a100:4(S:0-1)			

Alternativamente, también se puede ver enumerando y filtrando el listado de nodos:

```
$ scontrol show nodes | egrep "nodeName|gres"
```

```
....
nodeName=node1315-4 Arch=x86_64 CoresPerSocket=8
nodeName=node2204-3 Arch=x86_64 CoresPerSocket=10
...
nodeName=node17101-1 Arch=x86_64 CoresPerSocket=32
  CfgGRES=cpu=64, mem=257214M, billing=64, gres/gpu=4, gres/gpu:a100=4
nodeName=node17102-1 Arch=x86_64 CoresPerSocket=32
  CfgGRES=cpu=64, mem=257214M, billing=64, gres/gpu=4, gres/gpu:a100=4
nodeName=node17103-1 Arch=x86_64 CoresPerSocket=32
  CfgGRES=cpu=64, mem=257214M, billing=64, gres/gpu=4, gres/gpu:a100=4
nodeName=node17104-1 Arch=x86_64 CoresPerSocket=32
  CfgGRES=cpu=64, mem=257214M, billing=64, gres/gpu=4, gres/gpu:a100=4
...

```

Como se puede observar, actualmente cada nodo de GPU dispone de 4 tarjetas A100 de Nvidia.

28.2.3 Ejemplos de uso de GRES

- Para un nodo de GPU (icelake) con una GPU (Nvidia A100).

```
salloc -n 1 --cpus-per-task 8 -p gpu --mem 8000 --gres=gpu:a100=1
```



Tip

Recuerde que para utilizar GPUs tiene que utilizar la partición `gpu`, para la cual tiene que haber solicitado acceso.

Para solicitar una sola GPU en slurm simplemente agregue la directiva `#SBATCH --gres=gpu:<modelo>` a su script de envío de trabajo y le dará acceso a una GPU. Para solicitar varias GPU, agregue `#SBATCH --gres=gpu:<modelo>:n`, donde 'n' es la cantidad de GPU.

Entonces, si desea 1 CPU y 2 GPU de nuestros nodos GPU de uso general en la partición 'gpu', debe especificar:

```
#SBATCH -p gpu
#SBATCH -n 1
#SBATCH --gres=gpu:a100:1
```

Si prefiere usar una sesión interactiva, puede usar:

```
salloc -p gpu --mem 8000 --gres=gpu:a100:1
```

Mientras está en el nodo GPU, puede ejecutar `nvidia-smi` para obtener información sobre las GPU asignadas.

28.2.4 Scripts de ejemplo para un trabajo de GPU

- Ejemplo usando una GPU Nvidia A100 completa

```
#!/bin/bash

#SBATCH --partition=gpu
#SBATCH --nodes=1
#SBATCH --ntasks=1
#SBATCH --cpus-per-task=4
#SBATCH --gres=gpu:a100:1
#SBATCH --mem=8G
#SBATCH --time=1:00:00

module purge
module load CUDA/12.0.0

nvidia-smi
sleep 20
```

```
sbatch 01_gpu_basic_a100.sbatch
```

28.2.5 Más ejemplos

Visita nuestro github https://github.com/hpciter/user_codes

29 Ejemplos de script

En esta sección, le proporcionamos un conjunto de ejemplos que le serán útiles para configurar scripts según sus necesidades de investigación. Si está interesado en echarles un vistazo, están todos disponibles en el directorio /*share/user_codes/*.

Además, estos ejemplos están en nuestro repositorio. https://github.com/hpciter/user_codes.git

Para ejecutar los scripts, cópielos en su carpeta *data*.

```
$ tree /share/user_codes/
├── arrays
│   ├── 01-2node-4cpupertask
│   │   ├── 01-2nodes-4cpupertask.sbatch
│   │   └── README.md
│   ├── 02-1node-4cpupertask
│   │   ├── 02-1node-4cpupertask.sbatch
│   │   └── README.md
│   └── 03-1node-6cpupertask-ilk
│       ├── 03-1node-6cpupertask.sbatch
│       └── README.md
├── basics
│   ├── 00-sandy-nodes.sbatch
│   ├── 01-icelake-nodes.sbatch
│   └── README.md
├── gpu_basics
│   ├── 01_gpu_basic_a100.sbatch
│   ├── 02_gpu_basic_mig_partition.sbatch
│   └── README.md
├── images
│   └── teidehpc_logo.png
├── LICENSE
└── README.md
```

29.1 Trabajos de uno y varios núcleos

Dependiendo del consumo de recursos, la simulación podría necesitar más de un procesador. En el script bash, es posible especificar cuántos núcleos se solicitan para ejecutar el trabajo.

Si la simulación usa más de un núcleo, debe agregar el parámetro `--cpu-per-task` o `-c` con la cantidad de núcleos de CPU utilizados por tarea.

Sin esta opción, el controlador asignará un núcleo. Recuerde especificar el número de tareas con el parámetro `--ntasks` o `-n`.

Single core	Multi-core
<code>#SBATCH --ntasks=X donde X => 1</code>	<code>#SBATCH --ntasks=X donde X => 1</code> <code>#SBATCH --cpus-per-task=Y donde Y > 1</code>

29.2 OMP, MPI y Trabajos Híbridos

Es posible encontrar diferentes tipos de paralelismo: OpenMP, OpenMPI o una solución híbrida combinando ambos. Por un lado, utilizará OpenMP para el paralelismo dentro de un nodo multinúcleo.

Dado que es una implementación de subprocesos múltiples, se debe definir una variable OMP_NUM_THREADS. En cambio, si el paralelismo es entre nodos, usarás OpenMPI. Entonces, en nuestro script de muestra, será necesario especificar la cantidad de nodos, la cantidad de tareas en cada nodo y cada CPU.

OpenMP	OpenMPI	Hybrid
<pre>#SBATCH --cpus-per-tasks=X donde X > 1 export OMP_NUM_THREADS=X donde X > 1</pre>	<pre>#SBATCH --ntasks=X donde X => 1 #SBATCH --cpus-per-task=Y donde Y>1 #SBATCH --nodes=Z donde Z =>2 #SBATCH --ntasks-per-node=W donde W>1 #SBATCH --ntasks-per-socket=U donde U>1 module load OpenMPI/2.0.2- GCC-6.3.0-2.27</pre>	Combine both options

29.3 Arrays de Jobs

Con el uso de arrays puede ejecutar múltiples trabajos con los mismos parámetros. En su script debe especificar la opción `--array`.

Array
<pre>#SBATCH --array=1-X donde X > 1</pre>

29.4 Trabajos de GPU

Es imprescindible utilizar el parámetro `--gres` para reservar un recurso GPU y cargar el módulo CUDA.

Nvidia A100	Nvida Tesla T4
<pre>#SBATCH --gres=gpu:a100:1 module load CUDA/8.0.61</pre>	<pre>#SBATCH --gres=gpu:t4:1 module load CUDA/8.0.61</pre>

II.II Transferencia de datos

30 Transferencia de datos masiva

Para realizar transferencias de datos se dispone de nodos de transferencia con un mayor ancho de banda y que disponen de IP pública, con lo que no es necesario utilizar la VPN. Estos nodos permiten copiar y descargar grandes cantidades de datos al espacio de usuario.

Actualmente tenemos dos nodos habilitados para tal efecto que proporcionan un mayor ancho de banda entre los propios nodos y el sistema de almacenamiento compartido. Estos nodos tienen las siguientes IPs públicas:

- 193.146.150.185
- 193.146.150.186

La transferencia de información a través de estos nodos solo se permite mediante el uso del protocolo seguro SFTP(SecureShell File Transfer Protocol o Secure File Transfer Protocol). SFTP nos permitirá descargar y subir archivos muy fácilmente, a la vez que nos proporciona confidencialidad y autenticación de los datos transmitidos, a diferencia de un servidor FTP donde no tenemos ningún tipo de seguridad, ya que las credenciales de usuario se envían sin cifrar, y todo el tráfico de datos también. Para ello es necesario disponer de un cliente de dicho protocolo en la máquina local, como pueden ser **sftp en linux o psftp en windows**.

Info

En estos nodos solo está permitido el uso de SFTP, no de SCP o Rsync. Estos últimos solo se pueden utilizar en los nodos de login, pero para copiar y descargar datos de gran tamaño **recomendamos encarecidamente** utilizar estos nodos por el mayor ancho de banda que tienen. Además, en los nodos de login, el SFTP está deshabilitado.

30.1 SFTP para usuarios Linux

Se puede abrir la conexión al servidor a través de su IP pública:

```
sftp miusuario@IPservidor
```

30.1.1 Comandos más usados

Una vez abierta la sesión, el comando help mostrará la lista de comandos que se pueden utilizar. A continuación se describen los más utilizados y son válidos para usuarios Linux como Windows:

```
sftp> help      -- Muestra la ayuda.

sftp> cd dir    -- Cambia el directorio de trabajo remoto.
sftp> lcd dir   -- Cambia el directorio de trabajo local.

sftp> pwd      -- Muestra el directorio de trabajo actual.
sftp> lpwd     -- Muestra el directorio de trabajo actual local.

sftp> put file1.zip ...      -- Sube un fichero desde el directorio de trabajo local al directorio
de trabajo remoto
sftp> get file1.zip ...      -- Descarga un fichero desde el directorio de trabajo remoto al
directorio de trabajo local

sftp> put -r directory      -- Sube un directorio desde el directorio de trabajo local al
directorio de trabajo remoto
sftp> get -r directory      -- Descarga un directorio desde el directorio de trabajo remoto al
directorio de trabajo local
```

30.2 SFTP para usuarios Windows

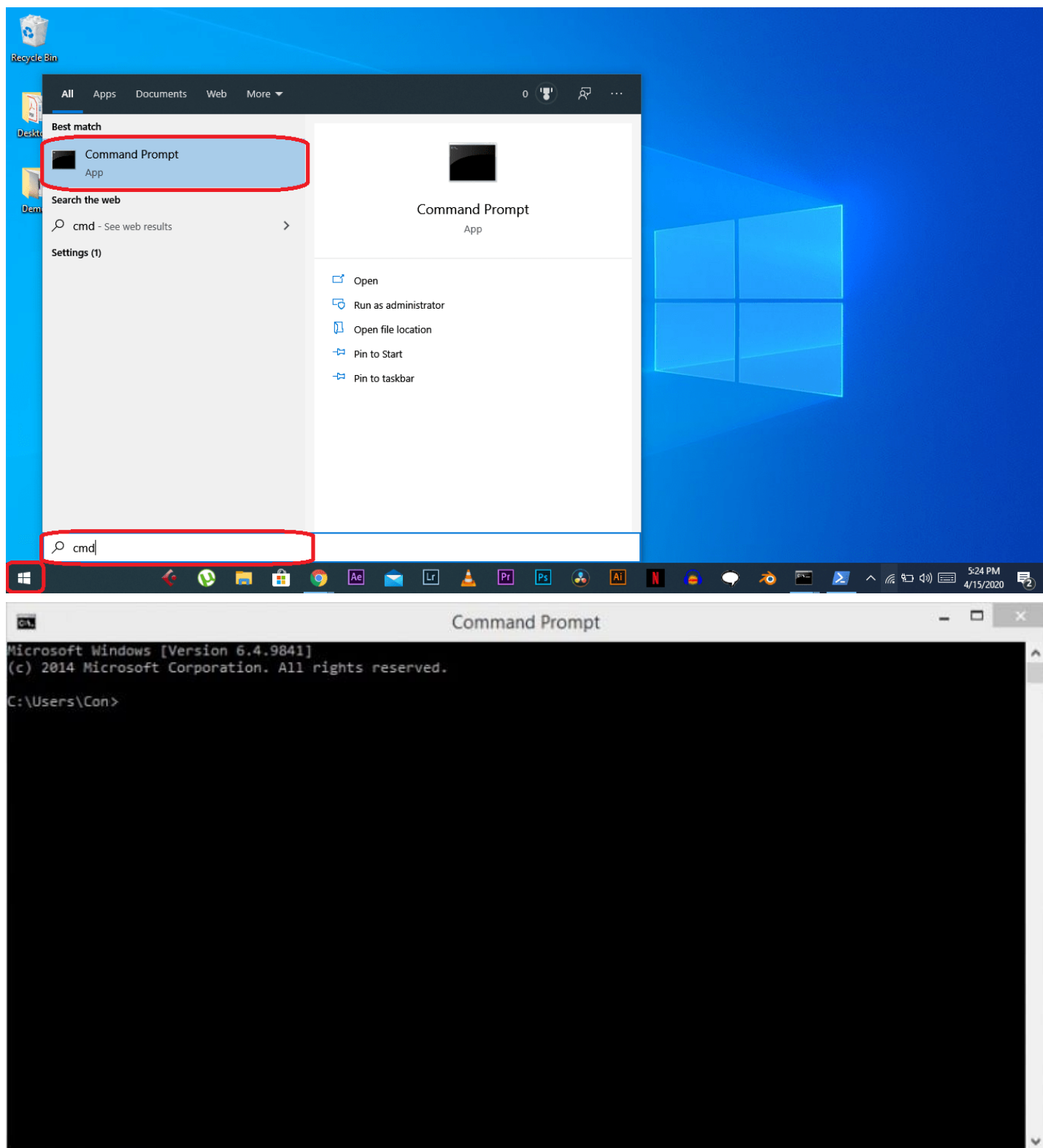
Para los usuarios windows existen multitud de aplicaciones tanto de linea de comandos como con interfaz gráfica. Entre ellas las más conocidas son: **psftp**, **Filezilla** y **WinSPC**.

30.2.1 psftp

PSFTP es el cliente de línea de comandos SFTP que se instala con la aplicación PuTTY, uno de los clientes SSH más populares para windows.

Hay tres formas de abrir PSFTP

1. Haciendo Click en inicio de Windows, buscar la aplicación Putty y luego PSFTP.
2. Ir al directorio "C:\Program Files (x86)\PuTTY" y hacer doble click en psftp.exe.
3. Iniciar la aplicación de línea de comandos de windows *cmd* o *PowerShell*.



Para esta última opción, es recomendable incluir el directorio en el que se ha instalado PSFTP dentro del PATH del usuario [aquí](#)

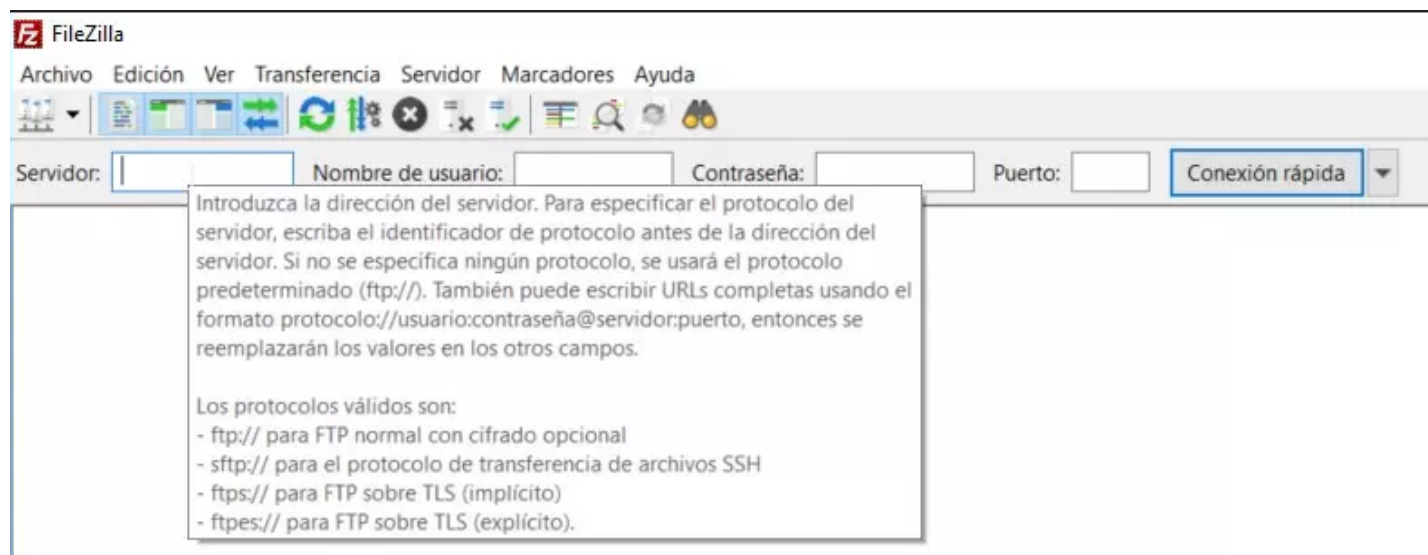
```
set PATH=C:\path\to\putty\directory;%PATH%
```

Se puede abrir la conexión al servidor a través de su IP pública.

```
C:\Users\Usuario>psftp usuario@IPservidor
```

30.2.2 Filezilla

FileZilla es uno de los programas más utilizados para usarlo como cliente FTP/FTPS y FTPES, pero también incorpora la posibilidad de conectarnos a un servidor SFTP. Tan solo tendremos que introducir en la barra de direcciones la siguiente sintaxis **Servidor: sftp://IP**, y seguidamente, el **usuario** y la **contraseña** que se le ha suministrado, e introducir el puerto de escucha que tengamos configurado en el servidor SSH al que pretendemos conectarnos.



30.2.3 WinSCP

WinSCP es un popular cliente para descargas gratuito que está disponible para Windows, soporta los protocolos SFTP, SCP, WebDAV y FTP, está enfocado principalmente para la transferencia de archivos, el uso de scripts y funcionalidades básicas de un administrador de archivos.

The screenshot shows the WinSCP interface with two panels. The left panel displays the local file system at D:\Documents\wiki\, listing files like interfaces.txt, introduction.txt, and languages.txt. The right panel displays the remote file system at /home/mprikryl/httpdocs/wiki/, listing files like administration.txt, after_installation.txt, and .htaccess. Below the panels is a Queue (2) section with a table of transfer operations.

Operation	Source	Destination	Transferred	Time	Speed	Progress
	/home/mprikryl/httpdocs/for...	D:\Documents\backup*.*	2 KB			Completed
	D:\Documents\wiki	/home/mprikryl/httpdocs...	29 KB	0:00:06	3,91 KB/s	52%
	D:\Documents\wiki\config.txt	/home/mprikryl/httpdocs...	5 KB			30%
	D:\Documents\movies\Movie\...	/home/mprikryl/httpdocs...	6 395 KB	0:07:49	44,6 MB/s	8%

30.2.4 Otros

- [Core FTP Client](#)
- [Cyberduck](#)
- [Smart FTP](#)

30.3 SFTP para usuario de MAC

Lo usuario de MAC pueden usar la terminal de la misma manera que los usuarios Linux o bien instalar una aplicación con interfaz gráfica como la siguiente:

- [Cyberduck](#)

30.4 Enlaces de interés

- [sftp -- Linux manual page](#)

III. Software y herramientas

31 Entorno de módulos (Environment Modules)

TeideHPC es un cluster compartido por cientos de usuarios, por lo que no es lo mismo la forma de proporcionar múltiples aplicaciones, en sus múltiples versiones, a multitud de usuarios, en un entorno de computación distribuida, que en un ordenador personal o en tu propio servidor.

Environment Modules es un sistema de gestión de software que permite a los usuarios de sistemas Unix-like, como Linux y macOS, y particularmente en entornos de computación de alto rendimiento (HPC), **dinámicamente modificar su entorno, acceder a diferentes compiladores, bibliotecas y software que pueden tener múltiples versiones y dependencias.**

Existen diferente implementaciones de Environment Modules pero la usada actualmente en TeideHPC es **Lmod**.

31.1 Lmod

Lmod es un acrónimo de **Lua Module System** es una implementación específica del concepto de Environment Modules para la gestión de entornos de software en sistemas de alto rendimiento (HPC).

Lmod permite a los usuarios:

- cargar y descargar módulos de software dinámicamente.
- cambia las variables de entorno como PATH y LD_LIBRARY_PATH.
- facilitar el uso de diferentes versiones de aplicaciones y bibliotecas sin interferencias entre ellas.

Puedes obtener más información sobre Lmod en <https://lmod.readthedocs.io/en/latest/>.

31.2 ¿Cómo se organiza el software en TeideHPC? Nomenclatura plana.

En TeideHPC existen 2 clusters, TeideHPC y AnagaGPU

Por este motivo, el conjunto de módulos disponibles **depende del Cluster que vaya a usar.**

Cada cluster tiene sus propios nodos de login y **no disponen del mismo software instalado.**

El esquema de módulos utiliza una nomenclatura específica para sus módulos, que normalmente sigue el formato:

```
<nombre_del_software>/<versión>-<toolchain>[-<variantes>]
```

- **Nombre_del_software:** Es el nombre del software o la biblioteca que se está instalando.
- **Versión:** Es la versión específica del software.
- **Toolchain:** Es la cadena de herramientas de compilación que incluye el compilador, la biblioteca MPI (si aplica), y otras bibliotecas de optimización de rendimiento que se usan para compilar el software. Por ejemplo, foss es una cadena de herramientas común que incluye GCC, OpenMPI, y otras.

- **Variantes:** Son modificadores adicionales que indican variantes específicas de la instalación, como optimizaciones específicas de hardware o la inclusión de ciertas características.

Por ejemplo, un módulo para la versión 6.0 de Unzip compilada con la cadena de herramientas GCCcore 12.2.0 podría tener una nomenclatura de módulo como

```
UnZip/6.0-GCCcore-12.2.0
```

Un ejemplo de módulo con distintas variantes y toolchain es el siguiente:

```
module spider 'Python'
-----
Python:
-----
Description:
  Python is a programming language that lets you work more quickly and integrate your systems
more
  effectively.

Versions:
  Python/2.7.18-GCCcore-11.2.0-bare
  Python/2.7.18-GCCcore-11.2.0
  Python/2.7.18-GCCcore-12.2.0-bare
  Python/3.8.6-GCCcore-11.2.0
  Python/3.8.6-intel-2021b
  Python/3.9.6-GCCcore-11.2.0-bare
  Python/3.9.6-GCCcore-11.2.0
  Python/3.9.6-intel-2021b
  Python/3.10.8-GCCcore-12.2.0-bare
  Python/3.10.8-GCCcore-12.2.0
```

31.2.1 ¿Qué es un *Toolchain*?

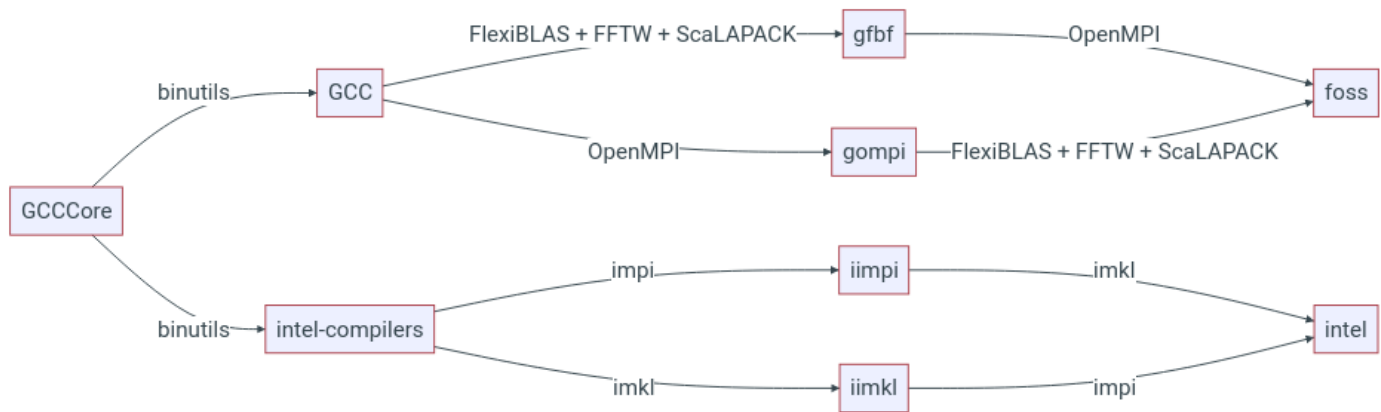
El término "toolchain" se refiere al conjunto de herramientas de software que se utilizan para compilar y enlazar programas.

Un toolchain típicamente **incluye compiladores para diferentes lenguajes de programación, bibliotecas de software matemáticas, y herramientas de enlazado y depuración.**

Los toolchains están definidos por un conjunto de componentes que incluyen:

- **Compiladores:** GCC (GNU Compiler Collection), Intel Compiler, etc.
- **Bibliotecas de comunicaciones MPI:** Por ejemplo, OpenMPI, MPICH, Intel MPI, etc.
- **Bibliotecas de matemáticas:** BLAS (Basic Linear Algebra Subprograms) y LAPACK (Linear Algebra Package).
- **Otras bibliotecas y herramientas:** Como GMP (GNU Multiple Precision Arithmetic Library), FFTW (Fastest Fourier Transform in the West), etc.

El siguiente diagrama puede ver cómo se divide el software en función del Toolchain elegido. **Básicamente se identifican 2 ramas en función del compilador principal: GNU GCC o Intel GCC**



Asimismo, **existen diferentes generaciones de estos toolchains que son incompatibles entre sí.**

31.2.2 Toolchain *foss*

El toolchain *foss* consta enteramente de software de código abierto (de ahí el nombre, derivado del término común 'FOSS', que es la abreviatura de "Free and Open Source Software").

En TeideHPC podemos identificar las siguientes generaciones para *foss*

foss	binutils	GCC	Open MPI	FlexiBLAS	OpenBLAS	LAPACK	Sc
2021b	2.37	11.2.0	4.1.1	3.0.4	0.3.18	(incl. with OpenBLAS)	2.1
2022a	2.38	11.3.0	4.1.4	3.2.0	0.3.20	(incl. with OpenBLAS)	2.1
2022b	2.39	12.2.0	4.1.4	3.2.1	0.3.21	(incl. with OpenBLAS)	2.1

31.2.3 Toolchain *intel*

El Toolchain *Intel* consta de compiladores y bibliotecas de Intel.

En TeideHPC podemos identificar las siguientes generaciones para *intel*

intel	binutils	GCC	Intel compilers	Intel MPI	Intel MKL
2021b	2.37	11.2.0	2021.4.0	2021.4.0	2021.4.0
2022a	2.38	11.3.0	2022.1.0	2021.6.0	2022.1.0

intel	binutils	GCC	Intel compilers	Intel MPI	Intel MKL
2022b	2.39	12.2.0	2022.2.1	2021.7.1	2022.2.1

En la siguiente sección puede ver [cómo usar modules para ver y cargar el software](#) disponible en TeideHPC.

```
module spider foss
```

```
-----  
foss:  
-----  
Description:  
  GNU Compiler Collection (GCC) based compiler toolchain, including OpenMPI for MPI support,  
OpenBLAS (BLAS  
  and LAPACK support), FFTW and ScaLAPACK.  
  
Versions:  
  foss/2021b  
  foss/2022b  
-----  
To find other possible module matches execute:  
  
$ module -r spider '.*GCC.*'
```

```
module spider intel
```

```
-----  
intel:  
-----  
Description:  
  Compiler toolchain including Intel compilers, Intel MPI and Intel Math Kernel Library (MKL).  
  
Versions:  
  intel/2021b  
  intel/2022b  
Other possible modules matches:  
  intel-compilers  
-----  
To find other possible module matches execute:  
  
$ module -r spider '.*intel.*'
```

32 Software disponible en los clusters TeideHPC y AnagaGPU

Antes de pasar a leer la siguiente sección, le recomendamos echar un vistazo a la sección "[Herramienta de módulos](#)"

El sistema de Environment Modules utiliza un conjunto de comandos que incluyen:

- `module overview` : Muestra una lista general resumida de los módulos disponibles.
- `module avail` : Muestra todos los módulos disponibles en el sistema para cargar.
- `module avail [módulo]` : Muestra los módulos disponibles con ese nombre.
- `module spider [módulo]` : Realiza una búsqueda de un módulo concreto.
- `module load [módulo]` : Carga un módulo específico, lo que establece las variables de entorno necesarias para utilizar el software que el módulo representa.
- `module unload [módulo]` : Descarga un módulo, eliminando las variables de entorno asociadas con el software de la sesión del usuario.
- `module swap [módulo_actual] [módulo_nuevo]` : Intercambia un módulo cargado por otro, lo que es útil para cambiar rápidamente de versiones de software.
- `module purge` : Elimina todos los módulos cargados, limpiando el entorno del usuario de cualquier configuración realizada por módulos.
- `module help [módulo]` : Proporciona información de ayuda para un módulo específico, lo que puede incluir información sobre cómo usar el software, sus variables de entorno, etc.
- `module whatis [módulo]` : Muestra una breve descripción de lo que hace un módulo.
- `module show [módulo]` : Muestra información sobre lo que hace el módulo al entorno del usuario; es decir, qué variables de entorno ajusta y cómo.

En TeideHPC existen 2 clusters, TeideHPC y AnagaGPU

Por este motivo, el conjunto de módulos disponibles **depende del Cluster que vaya a usar.**

Cada cluster tiene sus propios nodos de login y **no disponen del mismo software instalado.**

La ejecución de software está prohibida en los nodos de login

En los nodos de login TeideHPC y AnagaGPU está prohibido ejecutar cualquier aplicación salvo comandos comunes de linux.

ml es una abreviatura de *modules*

Para evitar errores comunes al escribir el comandos *modules* tales como *modles*, *moduls*, *mdules*, etc es posible usar la abreviatura **ml**.

```
ml spider ... ml load ... ml purge ...
```

32.1 Significado de algunos módulos significativos.

- **GCCcore**: conjunto básico de compiladores de la GNU Compiler Collection: C, C++, Objective-C, Fortran.
- **GCC**: colección de compiladores de GCCcore y librerías (GDB, binutils, glibc)
- **foss**: el módulo foss consiste en su totalidad en el término común *FOOS*, que es la abreviatura de *Free and Open source software*.
- **gomp**: toolchain que agrupa herramientas compiladas con GCC y OpenMPI
- **gompic**: toolchain que agrupa herramientas compiladas con GCC+CUDA y OpenMPI
- **intel-compilers**: Compiladores de Intel C, C++ & Fortran clasicos y oneAPI.
- **iimpi**: Compilador Intel C/C++ y Fortran con soporte MPI de intel.

32.2 Ejemplo de uso de modules

```
module overview  
or ml overview
```

```

-----/share/easybuild/software/common/modules/all -----
EasyBuild (2)  Go (1)  Mamba (1)  Miniconda3 (2)  Singularity (1)  Squashfs (1)  slurm (1)

-----/share/easybuild/software/x86_64/modules/all -----
ADMIXTURE      (1)  Imath          (1)  UCX            (2)  jupyter-server
(1)
ATK             (1)  Infernal      (1)  UDUNITS       (2)  libGLU
(2)
ATLAS          (1)  JasPer        (4)  USEARCH       (1)  libarchive
(3)
AdmixTools     (2)  Java          (5)  UnZip         (2)  libcerf
(1)
Armadillo      (2)  JsonCpp       (1)  WPS           (2)  libdeflate
(1)
Arrow          (2)  JupyterLab    (2)  WRF           (2)  libdrm
(2)
Autoconf       (4)  LAME          (1)  X11           (2)  libepoxy
(1)
Automake       (4)  LAMPLD        (2)  XALT          (1)  libevent
(1)
Autotools      (5)  LAPACK        (1)  XML-LibXML    (2)  libfabric
(2)
BCFtools       (1)  LERC          (1)  XZ            (2)  libffi
(2)
BEDTools       (1)  LLVM          (2)  Xerces-C++    (1)  libgd
(1)

```

```

module avail
or ml av

```

```

----- /share/easybuild/software/common/modules/all -----
EasyBuild/4.7.0      Go/1.18.3      Miniconda3/22.11.1-1      Singularity/3.11.0
slurm/teide
EasyBuild/4.8.2 (D)  Mamba/4.14.0-0  Miniconda3/23.5.2-0 (D)  Squashfs/4.3

----- /share/easybuild/software/x86_64/modules/all -----
ADMIXTURE/1.3.0      UCC/1.1.0-GCCcore-12.2.0
ATK/2.38.0-GCCcore-12.2.0  UCX/1.11.2-GCCcore-11.2.0
ATLAS/0.9.9-foss-2022b  UCX/1.13.1-GCCcore-12.2.0
(D)
AdmixTools/7.0.2-foss-2021b  UDUNITS/2.2.28-GCCcore-11.2.0
AdmixTools/7.0.2-foss-2022b (D) UDUNITS/2.2.28-GCCcore-12.2.0
(D)
Armadillo/10.5.3-foss-2022b  USEARCH/11.0.667-i86linux32
Armadillo/11.4.3-foss-2022b (D) UnZip/6.0-GCCcore-11.2.0
Arrow/6.0.0-foss-2021b      UnZip/6.0-GCCcore-12.2.0
(D)
Arrow/11.0.0-gfbf-2022b      (D) WPS/3.9.1-intel-2021b-dmpar
Autoconf/2.69-GCCcore-11.2.0  WPS/4.1-intel-2021b-dmpar
(D)
Autoconf/2.71-GCCcore-11.2.0  WRF/3.9.1.1-intel-2021b-dmpar
Autoconf/2.71-GCCcore-12.2.0  WRF/4.4.1-foss-2022b-dmpar
(D)
Autoconf/2.71 (D) X11/20210802-GCCcore-11.2.0
Automake/1.16.2-GCCcore-11.2.0  X11/20221110-GCCcore-12.2.0
(D)
Automake/1.16.4-GCCcore-11.2.0  XALT/3.0.1
...

```



```
module spider wrf
or ml spider wrf
```

```
-----
WRF:
-----
```

Description:

The Weather Research and Forecasting (WRF) Model is a next-generation mesoscale numerical weather prediction system designed to serve both operational forecasting and atmospheric research needs.

Versions:

```
WRF/3.9.1.1-intel-2021b-dmpar
WRF/4.4.1-foss-2022b-dmpar
```

```
-----
For detailed information about a specific "WRF" package (including how to load the modules) use
the module's full name.
```

Note that names that have a trailing (E) are extensions provided by other modules.
For example:

```
$ module spider WRF/4.4.1-foss-2022b-dmpar
```

```
module load WRF/3.9.1.1-intel-2021b-dmpar
module list
```

Currently Loaded Modules:

1) GCCcore/11.2.0	11) intel/2021b
2) zlib/1.2.11-GCCcore-11.2.0	12) libtirpc/1.3.1-GCCcore-11.2.0
3) binutils/2.37-GCCcore-11.2.0	13) JasPer/2.0.24-GCCcore-11.2.0
4) intel-compilers/2021.4.0	14) Szip/2.1.1-GCCcore-11.2.0
5) numactl/2.0.14-GCCcore-11.2.0	15) HDF5/1.10.7-iimpi-2021b
6) UCX/1.11.2-GCCcore-11.2.0	16) OpenSSL/1.1
7) impi/2021.4.0-intel-compilers-2021.4.0	17) cURL/7.78.0-GCCcore-11.2.0
8) imkl/2021.4.0	18) netCDF/4.7.4-iimpi-2021b
9) iimpi/2021b	19) netCDF-Fortran/4.5.3-iimpi-2021b
10) imkl-FFTW/2021.4.0-iimpi-2021b	20) WRF/3.9.1.1-intel-2021b-dmpar

• **foss**

El módulo foss consiste en su totalidad en el término común *FOOS*, que es la abreviatura de *Free and Open source software*. 2021a, 2021b, 2022b, etc indica la generación.

Este módulo consta de los siguientes paquetes y se encarga de precargarlos.

- binutils
- Compilador GNU GCC (C), g++ (C++) and gfortran (Fortran)
- Librería Open MPI
- Librería FlexiBLAS con OpenBLAS + LAPACK como backend
- Librería ScaLAPACK

- **Librería FFTW (Fourier transform (DFT))**

```
module spider foss
module load foss/2021b
module list
```

Currently Loaded Modules:

1) GCCcore/11.2.0 GCC-11.2.0	8) libpciaccess/0.16-GCCcore-11.2.0	15) OpenMPI/4.1.1-
2) zlib/1.2.11-GCCcore-11.2.0 GCC-11.2.0	9) hwloc/2.5.0-GCCcore-11.2.0	16) OpenBLAS/0.3.18-
3) binutils/2.37-GCCcore-11.2.0 GCC-11.2.0	10) OpenSSL/1.1	17) FlexiBLAS/3.0.4-
4) GCC/11.2.0	11) libevent/2.1.12-GCCcore-11.2.0	18) gOMPI/2021b
5) numactl/2.0.14-GCCcore-11.2.0 gOMPI-2021b	12) UCX/1.11.2-GCCcore-11.2.0	19) FFTW/3.3.10-
6) XZ/5.2.5-GCCcore-11.2.0 gOMPI-2021b-fb	13) libfabric/1.13.2-GCCcore-11.2.0	20) ScaLAPACK/2.1.0-
7) libxml2/2.9.10-GCCcore-11.2.0	14) PMIx/4.1.0-GCCcore-11.2.0	21) foss/2021b

- **gOMPI:**

El módulo *gOMPI* precarga todo el software compilado con GNU GCC y OpenMPI. 2021a, 2021b, 2022b, etc indica la generación.

```
$ module load gOMPI/2021b
$ module list
```

Currently Loaded Modules:

1) GCCcore/11.2.0 GCCcore-11.2.0	7) libxml2/2.9.10-GCCcore-11.2.0	13) libfabric/1.13.2-
2) zlib/1.2.11-GCCcore-11.2.0 GCCcore-11.2.0	8) libpciaccess/0.16-GCCcore-11.2.0	14) PMIx/4.1.0-
3) binutils/2.37-GCCcore-11.2.0 GCC-11.2.0	9) hwloc/2.5.0-GCCcore-11.2.0	15) OpenMPI/4.1.1-
4) GCC/11.2.0	10) OpenSSL/1.1	16) gOMPI/2021b
5) numactl/2.0.14-GCCcore-11.2.0	11) libevent/2.1.12-GCCcore-11.2.0	
6) XZ/5.2.5-GCCcore-11.2.0	12) UCX/1.11.2-GCCcore-11.2.0	

- **intel:**

Este módulo común precarga los compiladores y librerías de Intel. 2021a, 2021b, 2022b, etc indica la generación.

- Compiladores Intel C/C++/Fortran (icc, icpc and ifort)
- binutils
- GCC, el cual sirve como base de los compiladores de intel.
- Librería Intel MPI
- Librería Intel Math Kernel Library (MKL) para funcionalidades de BLAS/LAPACK/FFT

```
module load intel/2022b
module list
```

Currently Loaded Modules:

1) GCCcore/12.2.0	5) numactl/2.0.16-GCCcore-12.2.0	9) iimpi/2022b
2) zlib/1.2.12-GCCcore-12.2.0	6) UCX/1.13.1-GCCcore-12.2.0	10) imkl-FFTW/
2022.2.1-iimpi-2022b		
3) binutils/2.39-GCCcore-12.2.0	7) impi/2021.7.1-intel-compilers-2022.2.1	11) intel/2022b
4) intel-compilers/2022.2.1	8) imkl/2022.2.1	

33 Solicitud de instalación de nuevo software

Para cualquier duda o problema, así como casos en los que se requiera instalación de software específico, mayor capacidad de cómputo, mayor espacio de almacenamiento, mayor duración de trabajos en los nodos de cómputo, utilización de máquinas no accesibles al usuario, etc. puede ponerse en contacto con los administradores de TeideHPC correo electrónico support@hpc.iter.es indicando, si es necesario, la siguiente información de la aplicación:

- *Nombre de la aplicación*
- *Versión*
- *Enlace de descarga o repositorio donde resida*
- *Otra información necesaria si lo desea*

III.I Guías de uso

34 Conda

Conda es un gestor paquetes y entornos de software libre utilizado sobre todo para paquetes de Python y R del ámbito científico, aunque también soporte otros lenguajes como C, C++, FORTRAN, Java, Scale, Ruby y Lua.

34.1 Utilizar Conda en TeideHPC

Una vez nos hayamos conectados a los nodos de login, tendremos que cargar el módulo de `miniconda` para empezar a utilizar `conda`:

```
module load Miniconda2/4.7.10
```

Una vez cargado el software, podemos, por ejemplo, ver la información de la versión de `conda`:

```
conda info
  active environment : None
  user config file : /home/vjuidias/.condarc
  populated config files : /home/vjuidias/.condarc
  conda version : 4.7.10
  conda-build version : not installed
  python version : 2.7.16.final.0
  virtual packages :
  base environment : /opt/envhpc/slurm19/rhel7/Miniconda2/4.7.10 (read only)
  channel URLs : https://conda.anaconda.org/conda-forge/linux-64
                 https://conda.anaconda.org/conda-forge/noarch
                 https://conda.anaconda.org/bioconda/linux-64
                 https://conda.anaconda.org/bioconda/noarch
                 https://repo.anaconda.com/pkgs/main/linux-64
                 https://repo.anaconda.com/pkgs/main/noarch
                 https://repo.anaconda.com/pkgs/r/linux-64
                 https://repo.anaconda.com/pkgs/r/noarch
                 https://conda.anaconda.org/r/linux-64
                 https://conda.anaconda.org/r/noarch
  package cache : /opt/envhpc/slurm19/rhel7/Miniconda2/4.7.10/pkg
                  /home/vjuidias/.conda/pkg
  envs directories : /home/vjuidias/.conda/envs
                    /opt/envhpc/slurm19/rhel7/Miniconda2/4.7.10/envs
  platform : linux-64
  user-agent : conda/4.7.10 requests/2.22.0 CPython/2.7.16 Linux/3.10.0-327.el7.x86_64
               centos/7.2.1511 glibc/2.17
  UID:GID : 1136:1136
  netrc file : None
  offline mode : False
```

Ahora tenemos que inicializar la shell para utilizar Conda. **Este paso es solo es necesario la primera vez.** Si no lo hacemos, al utilizar `conda`, nos saldrá este mensaje:

```
CommandNotFoundError: Your shell has not been properly configured to use 'conda activate'.
To initialize your shell, run
```

```
$ conda init <SHELL_NAME>
```

```
Currently supported shells are:
```

- bash
- fish
- tcsh
- xonsh
- zsh
- powershell

```
See 'conda init --help' for more information and options.
```

```
IMPORTANT: You may need to close and restart your shell after running 'conda init'.
```

Por tanto, lo que tenemos que hacer es:

```
conda init bash
no change      /opt/envhpc/slurm19/rhel7/Miniconda2/4.7.10/condabin/conda
no change      /opt/envhpc/slurm19/rhel7/Miniconda2/4.7.10/bin/conda
no change      /opt/envhpc/slurm19/rhel7/Miniconda2/4.7.10/bin/conda-env
no change      /opt/envhpc/slurm19/rhel7/Miniconda2/4.7.10/bin/activate
no change      /opt/envhpc/slurm19/rhel7/Miniconda2/4.7.10/bin/deactivate
no change      /opt/envhpc/slurm19/rhel7/Miniconda2/4.7.10/etc/profile.d/conda.sh
no change      /opt/envhpc/slurm19/rhel7/Miniconda2/4.7.10/etc/fish/conf.d/conda.fish
no change      /opt/envhpc/slurm19/rhel7/Miniconda2/4.7.10/shell/condabin/Conda.psm1
no change      /opt/envhpc/slurm19/rhel7/Miniconda2/4.7.10/shell/condabin/conda-hook.ps1
no change      /opt/envhpc/slurm19/rhel7/Miniconda2/4.7.10/lib/python2.7/site-packages/xontrib/
conda.xsh
no change      /opt/envhpc/slurm19/rhel7/Miniconda2/4.7.10/etc/profile.d/conda.csh
modified       /home/vjuidias/.bashrc

==> For changes to take effect, close and re-open your current shell. <==
```

Al hacer esto se nos ha modificado la shell, en este caso, la bash. Se nos han añadido las siguientes líneas en el fichero ~/.bashrc`:

```
# >>> conda initialize >>>
# !! Contents within this block are managed by 'conda init' !!
__conda_setup="$( '/opt/envhpc/slurm19/rhel7/Miniconda2/4.7.10/bin/conda' 'shell.bash' 'hook' 2> /
dev/null)"
if [ $? -eq 0 ]; then
    eval "$__conda_setup"
else
    if [ -f "/opt/envhpc/slurm19/rhel7/Miniconda2/4.7.10/etc/profile.d/conda.sh" ]; then
        . "/opt/envhpc/slurm19/rhel7/Miniconda2/4.7.10/etc/profile.d/conda.sh"
    else
        export PATH="/opt/envhpc/slurm19/rhel7/Miniconda2/4.7.10/bin:$PATH"
    fi
fi
unset __conda_setup
# <<< conda initialize <<<
```

Para que tenga efecto los cambios, tenemos que salir del nodo de login y volver a entrar. Una vez hecho, ya podremos utilizar conda, instalar paquetes y crear entornos.

34.2 Crear un entorno en /data

Por defecto, Conda creará un entorno e instalará todo el software en el `/home` del usuario. Este almacenamiento está limitado en tamaño, así que deberíamos configurar los entornos e instalar el software en la partición de `/data`. Para hacer esto, simplemente tenemos que especificarle a Conda la ruta:

```
conda create -p data/myenvironment mysoftware

Collecting package metadata (current_repodata.json): done
Solving environment: done

## Package Plan ##

  environment location: /data/vjuidias/myenvironment
  ...
```

Y lo activamos de la misma manera, especificando la ruta:

```
conda activate /data/vjuidias/myenvironment
```

4 11 18

35 Python

35.1 Utilizar Python en TeideHPC

Una vez conectados a los nodos de login, disponemos de una versión de Python por defecto, que es la versión 2.6.6. Para ver las versiones de Python disponibles, utilizamos la herramienta de *modules*:

```
module ava python

----- /opt/envhpc/modulefiles/.rhel6
-----
python/2.7.18/gcc python/3.5.4/gcc python/3.7.9/gcc python/3.8.11/gcc
```

Para cargar alguna de las versiones es posible que tengamos que cargar algún módulo previamente:

```
module load python/3.8.11/gcc
python/3.8.11/gcc(10):ERROR:151: Module 'python/3.8.11/gcc' depends on one of the module(s)
'openssl/1.1.1k/gcc'
python/3.8.11/gcc(10):ERROR:102: Tcl command execution failed: prereq openssl/1.1.1k/gcc
```

Por tanto, cargamos los módulos necesarios, en el orden correspondiente:

```
module load openssl/1.1.1k/gcc python/3.8.11/gcc

module list
Currently Loaded Modulefiles:
  1) openssl/1.1.1k/gcc  2) python/3.8.11/gcc
```

Ahora podremos utilizar una versión específica de Python y utilizar entornos, donde instalar paquetes de Python de manera aislada.

35.2 Instalar paquetes de Python en /data

Por defecto, Python instalará los paquetes en el `/home` del usuario. Este almacenamiento está limitado en tamaño, así que deberíamos configurar los entornos e instalar el software en la partición de `/data`. Para hacer esto, tenemos que hacer lo siguiente:

```
pip3 install --target=$HOME/data/tutu --install-option="--install-scripts=$HOME/data/foo" Package
```

Con la opción `--target=DIR` le indicamos a pip donde se tienen que instalar tanto el paquete solicitado, como todas sus dependencias. Por otro lado, la opción `--install-option="--install-scripts=DIR`, indica donde irán instalados los binarios, en caso de utilizar el paquete de python desde la línea de comandos directamente.

Luego, tenemos que añadir a nuestro PATH la nueva ruta para utilizar los binarios directamente:

```
export PYTHONPATH="$HOME/data/foo"
```

Para hacerlo permanente, podemos añadir esa línea al final de nuestro `~/ .bashrc`.

36 R

R es un lenguaje y un entorno de programación diseñado para trabajar en el ámbito estadístico.

36.1 Utilizar R en TeideHPC

Una vez conectados a los nodos de login, para ver las versiones de R disponibles, utilizamos la herramienta de *modules*:

```
module ava R

----- /opt/envhpc/modulefiles/.rhel6
-----
R/3.4.4/gcc R/3.6.1/gcc R/4.0.3/gcc R/4.1.1/gcc
```

Para cargar alguna de las versiones es posible que tengamos que cargar algún módulo previamente:

```
module load R/4.1.1/gcc
R/4.1.1/gcc(10):ERROR:151: Module 'R/4.1.1/gcc' depends on one of the module(s) 'gcc/10.2.0'
R/4.1.1/gcc(10):ERROR:102: Tcl command execution failed: prereq gcc/10.2.0
```

Por tanto, cargamos los módulos necesarios, en el orden correspondiente:

```
module load gcc/10.2.0 R/4.1.1/gcc

module list
Currently Loaded Modulefiles:
  1) openssl/1.1.1k/gcc  2) python/3.8.11/gcc  3) gcc/10.2.0  4) R/4.1.1/gcc
```

Ahora ya podemos empezar a utilizar R:

```
R

R version 4.1.1 (2021-08-10) -- "Kick Things"
Copyright (C) 2021 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

>
```

36.2 Instalar paquetes de R en /data

Por defecto, R instalará los paquetes en el `/home` del usuario. Este almacenamiento está limitado en tamaño, así que deberíamos configurar los entornos e instalar el software en la partición de `/data`. Para hacer esto, tenemos que configurar una variable de entorno:

```
export R_LIBS_USER=$HOME/data/mylibrary
```

De esta forma, R utilizará esta ruta como librería por defecto en lugar del `/home` del usuario. Al utilizar las funciones de R como `install.packages()` y `library()`, los paquetes serán instalados en la ruta definida en la variable `R_LIBS_USER`. Si quiere evitar tener que definir esta variable cada vez que se conecta, puede añadir esa línea al final de su fichero `~/.bashrc`, de esta forma, será permanente.

36.3 Ejecutar un script de R en Slurm

Para utilizar un script de R en Slurm, tenemos que ejecutarlo de la siguiente manera:

```
Rscript myscript.R
```

III.1.1 Jupyter

37 Jupyter en el Clúster HPC

Proyecto Jupyter es un proyecto con el objetivo de desarrollar software de código abierto, estándares abiertos y servicios para computación interactiva a través de múltiples lenguajes de programación.

Los **Jupyter Notebooks** son herramientas web interactivas conocidas como cuadernos computacionales, que los investigadores pueden usar para combinar código de software, texto explicativo y recursos multimedia, y salida computacional, en un solo documento. Jupyter se ha convertido en un estándar de facto para los científicos de datos y otros dominios científicos.

Dentro del proyecto Jupyter se han desarrollado y respaldado los productos informáticos interactivos **Jupyter Notebook, JupyterLab y JupyterHub**.

37.0.1 Jupyter Notebook: la interfaz de notebook clásica

Jupyter Notebook es la aplicación web original para crear y compartir documentos computacionales. Ofrece una experiencia simple, optimizada y centrada en documentos.

Es una aplicación de servidor-cliente que permite editar y ejecutar documentos de cuaderno a través de un navegador web. Jupyter Notebook se puede ejecutar en un escritorio local que no requiera acceso a Internet o se puede instalar en un servidor remoto y acceder a través de Internet.

37.0.2 JupyterLab: una interfaz de notebook de próxima generación

JupyterLab es el último entorno de desarrollo interactivo basado en la web para notebooks, código y datos. Su interfaz flexible permite a los usuarios configurar y organizar flujos de trabajo en ciencia de datos, computación científica, periodismo computacional y aprendizaje automático.

37.0.3 JupyterHub

JupyterHub lleva el poder de los notebooks a grupos de usuarios. Brinda a los usuarios acceso a entornos y recursos computacionales sin sobrecargar a los usuarios con tareas de instalación y mantenimiento. Los usuarios, incluidos estudiantes, investigadores y científicos de datos, pueden realizar su trabajo en sus propios espacios de trabajo en recursos compartidos que los administradores del sistema pueden administrar de manera eficiente.

JupyterHub se ejecuta en la nube o en su propio hardware y permite brindar un entorno de ciencia de datos preconfigurado a cualquier usuario del mundo. Es personalizable y escalable, y es adecuado para equipos pequeños y grandes, cursos académicos e infraestructura a gran escala.

38 Ejecución remota de Jupyter Notebook con Slurm

Como sabemos, Jupyter Notebook se pueden iniciar localmente y acceder a los sistemas de archivos locales, o se pueden iniciar en una máquina remota y que proporcione acceso a los archivos de un usuario en el sistema remoto. En este último caso, Jupyter se inicia a través de un proceso que crea una URL única que se compone del nombre de host más un puerto disponible (elegido por la aplicación jupyter) más un token único. El usuario obtiene esta URL y la ingresa en un navegador web local, donde el servidor de Notebooks está disponible siempre que el proceso en la máquina remota esté activo y en ejecución. Hay que tener en cuenta que de forma predeterminada, estos servidor de Jupyter no es seguro y exponen potencialmente los archivos locales de los usuarios a usuarios no deseados.

Por ello debes tener en cuenta estos conceptos:

✗ No ejecutar Jupyter en los nodos de login

Los nodos de login del clúster es un recurso compartido por muchos usuarios. La ejecución de Jupyter en uno de estos nodos puede afectar negativamente a otros usuarios. **Utilice uno de los enfoques descritos en esta página para realizar su trabajo.**

✗ Internet no está disponible en los nodos de cómputo

Las sesiones de Jupyter deben ejecutarse en los nodos de cómputo los cuales no tienen acceso a Internet. Esto significa que no podrá descargar archivos, clonar un repositorio de GitHub, instalar paquetes, etc. **Deberá realizar estas operaciones en los nodos de login.** Cualquier archivo que descargue mientras está en el nodo de login estará disponible inmediatamente en los nodos de cómputo durante la sesión.

✗ No puede acceder directamente a los nodos de cómputo

Por motivos de seguridad, no se puede acceder directamente a los nodos de cálculo. Sin embargo, se puede acceder a ellos mientras ejecuta un trabajo desde cualquier nodo de login.

✗ Esta implementación no es el servidor multiusuario que andas buscando

Este documento describe cómo puede ejecutar un servidor público con un solo usuario. Esto solo debe hacerlo alguien que quiera acceder de forma remota a su cuenta personal. Aun así, hacer esto requiere una comprensión profunda de las limitaciones de las configuraciones y las implicaciones de seguridad. Si permite que varios usuarios accedan a un servidor portátil como se describe en este documento, sus comandos pueden colisionar, aplastarse y sobrescribirse entre sí.

Si quieres un servidor multiusuario, la solución oficial es JupyterHub. Para usar JupyterHub, necesita un servidor Unix (generalmente Linux) que se ejecute en algún lugar al que puedan acceder sus usuarios en una red. Esto puede ejecutarse a través de Internet público, pero hacerlo presenta problemas de seguridad adicionales.

38.0.1 Crear un entorno conda

En primer lugar tenemos que crear un entorno conda. Puede ver estos pasos [aquí] (how_to_conda.md).

Recordar:

- Use nodos de inicio de sesión para crear un entorno python/conda. Sólo estos tienen acceso a internet.
- **Instalar paquetes python o conda en /data**

```
$ module spider miniconda
$ module load Miniconda3/4.9.2
$ conda create --name jupyter-env [python=3.9.15] or
$ conda create --prefix /home/user/data/allenvironments/my_environment [python=3.9.15]
```

```
Collecting package metadata (current_repodata.json): done
Solving environment: done
...
# To activate this environment, use
#   $ conda activate jupyter-env
# To deactivate an active environment, use
#   $ conda deactivate
```

- Instalar los paquetes necesarios para Jupyter Notebooks

```
conda activate jupyter-env
conda install jupyter ipykernel matplotlib ipywidgets ipympl --channel conda-forge -y
```

38.0.2 Ejecutar Jupyter Notebook en un nodo de cómputo mediante *salloc*

Las tareas más grandes se pueden ejecutar en uno de los nodos de cómputo solicitando una sesión interactiva usando *salloc*.

Lo primero es lo primero: inicia una sesión de **screen** (o *tmux* si lo prefieres). Si busca tener un programa ejecutándose por más tiempo del que quiero mantener abierta una ventana de terminal, *screen* o *tmux* son excelentes opciones, ya que evitan que su sesión se agote en máquinas remotas.

```
screen -S jupyter
salloc --nodes=1 --partition batch --ntasks=1 --mem=20G --time=12:00:00
# or salloc -N 1 -p batch
# or srun -p batch --pty bash

conda activate jupyter-env
```

Ahora debería ver que el indicador de su terminal ha cambiado a algo como lo siguiente, lo que indica que ha iniciado sesión interactiva un nodo de cómputo y está trabajando dentro del entorno conda-env:

```
(jupyter-env) yourUser@nodeXXXX-X
```

Una vez que se ha asignado un nodo de cómputo, inicia Jupyter. En el nodo de cómputo ejecuta:

```
jupyter-notebook --no-browser --port=8889 --ip=127.0.0.1
```

Jupyter Notebook se inician a través de un proceso que crea una URL única que se compone de ip y puerto 8889 más un token de un solo uso. El usuario obtiene esta URL y la ingresa en un navegador web local, donde el servidor está disponible siempre que el proceso en la máquina remota esté activo y en ejecución.

```
[I 13:22:01.198 NotebookApp] Writing notebook server cookie secret to /home/youruser/.local/share/jupyter/runtime/notebook_cookie_secret
```

```
[I 13:22:12.448 NotebookApp] Serving notebooks from local directory: /home/youruser
```

```
[I 13:22:12.448 NotebookApp] Jupyter Notebook 6.5.2 is running at:
```

```
[I 13:22:12.448 NotebookApp] http://127.0.0.1:8889/?  
token=36229e08e0944c8d1b4df0174e23a7ee11e278ccbed5f967
```

```
[I 13:22:12.448 NotebookApp] or http://127.0.0.1:8889/?  
token=36229e08e0944c8d1b4df0174e23a7ee11e278ccbed5f967
```

```
[I 13:22:12.448 NotebookApp] Use Control-C to stop this server and shut down all kernels (twice to skip  
confirmation). [C 13:22:12.568 NotebookApp]
```

To access the notebook, open this file in a browser:

```
file:///home/youruser/.local/share/jupyter/runtime/nbserver-4762-open.html
```

Or copy and paste one of these URLs:

```
http://127.0.0.1:8889/?token=36229e08e0944c8d1b4df0174e23a7ee11e278ccbed5f967
```

```
or http://127.0.0.1:8889/?token=36229e08e0944c8d1b4df0174e23a7ee11e278ccbed5f967
```

Tip

Tenga en cuenta que de forma predeterminada Jupyter Notebook abre automáticamente un navegador, pero *no podemos hacerlo en un nodo de cómputo directamente por seguridad*, por lo que omitimos esa función con el indicador `--no-browser`.

Tip

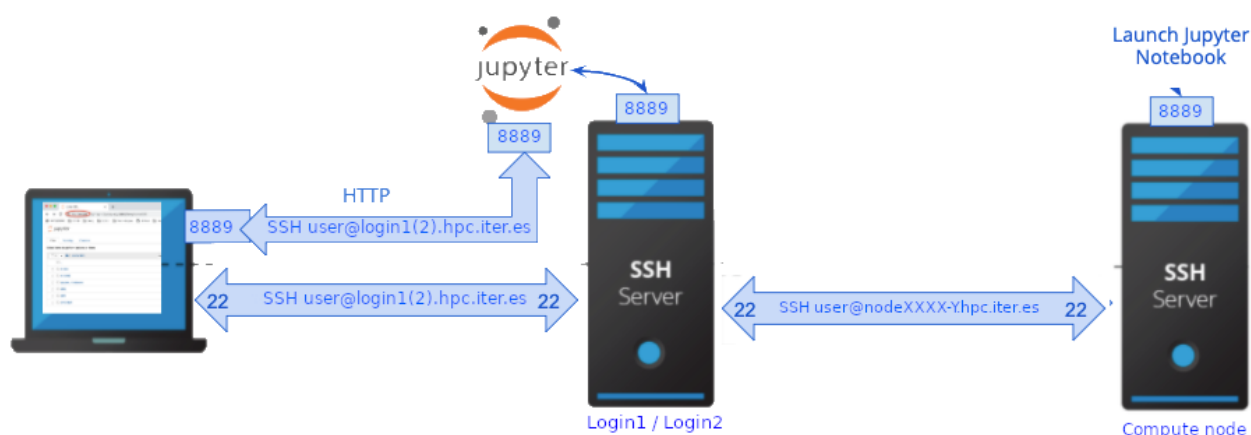
Tenga en cuenta que seleccionamos el puerto Linux `8889` para conectarse al servidor de Notebook. Si no especifica el puerto, se establecerá de forma predeterminada en el puerto `8888`, pero a veces este puerto puede estar ya en uso en la máquina remota o en la local (es decir, su PC). Si el puerto que seleccionó no está disponible, recibirá un mensaje de error, en cuyo caso debe elegir otro. Es mejor seleccionar un puerto mayor al 1024. Considere comenzar con `8888` e incrementarlo en 1 si falla, por ejemplo, intente `8888`, `8889`, `8890` y así sucesivamente. Si está ejecutando en un puerto diferente, sustituya su número de puerto por `8889`.

⚠ Por defecto, Jupyter Notebooks no es seguro y potencialmente exponen sus archivos a usuarios no deseados.

Ejecutar Notebook de la forma habitual utiliza conexiones HTTP inseguras. En esta guía, **presentamos una guía para ejecutarlo de forma más segura.**

38.0.2.1 Crear un tunnel SSH

Por razones de seguridad, recomendamos usar túneles ssh para conectarse de forma segura al servidor de Jupyter Notebook. Con esto creará una conexión ssh entre su host local y el puerto del servidor de Notebook ejecutándose en el nodo interactivo remoto. Cuando conecte su navegador al servicio de notebook, este canalizará todas las comunicaciones a través de la conexión SSH, que es segura y encriptada.



Inicie una segunda sesión de terminal en su máquina local, conéctese a cualquier nodo de login y configure el túnel SSH de la siguiente manera:

```
ssh yourUser@loginX.hpc.iter.es
```

En el nodo de login escribir:

```
ssh -N -L localhost:8889:localhost:8889 yourUser@nodeXXX-Y.hpc.iter.es
```

Podrías ver un mensaje de este tipo:

```
The authenticity of host 'node1710-1.hpc.iter.es (10.0.17.37)' can't be established.
ECDSA key fingerprint is SHA256:LqRpSE90hft08t047V2nx7jbIsU0X42SeKAgEVBPODo.

Are you sure you want to continue connecting (yes/no/[fingerprint])? yes

Warning: Permanently added 'node1710-1.hpc.iter.es' (ECDSA) to the list of known hosts.
```

Inicia un tercera consola en tu ordenador personal y establece el tunnel hacia el nodo de login de la siguiente manera.

```
ssh -N -L localhost:8889:localhost:8889 yourUser@loginX.hpc.iter.es
```

-N No ejecuta un comando remoto. Útil sólo para reenviar puertos.

-L Especifica que el puerto dado en el host local (cliente) se reenviará al puerto dado host y puerto en el lado remoto.

Puedes usar los DNS o las IPs de los nodos

Sólo hay que cambiar `nodeXXXX-X.hpc.iter.es` por la IP del nodo.

Finalmente, copie y pegue la dirección en su navegador favorito y **reemplace la parte “?token=x” de la URL con su token**

```
http://127.0.0.1:8889/?token=36229e08e0944c8d1b4df0174e23a7ee11e278ccbed5f967
```

Recomendamos leer sa sección [Notas importantes](#) below.



Files

Running

Clusters

To import a notebook, drag the file onto the listing below or [click here](#).

New ▾



<input type="checkbox"/>	 / examples
<input type="checkbox"/>	 ..
<input type="checkbox"/>	 Builtin Extensions
<input type="checkbox"/>	 Customization
<input type="checkbox"/>	 Embedding
<input type="checkbox"/>	 IPython Kernel
<input type="checkbox"/>	 Interactive Widgets
<input type="checkbox"/>	 Notebook
<input type="checkbox"/>	 Parallel Computing
<input type="checkbox"/>	 images
<input type="checkbox"/>	 utils
<input type="checkbox"/>	 Index.ipynb

38.0.3 Running Jupiter Notebook on a Compute Node via *sbatch*

La segunda forma de ejecutar Jupyter en el clúster es enviando un trabajo a mediante el commando *sbatch*.

Para hacer esto, necesitamos un script de envío como el siguiente que podemos llamar **jupyter.sh**:

```
#!/bin/bash
#SBATCH --nodes=1
#SBATCH --ntasks=1
#SBATCH --cpus-per-task=1
#SBATCH --mem=20G
#SBATCH --partition=batch
#SBATCH --constrains=sandy
#SBATCH --time=01:00:00
#SBATCH --job-name=jupyter-notebook
#SBATCH --output=jupyter-notebook-%j.out
#SBATCH --error=jupyter-notebook-%j.err
#####
# Enable modules profile

# Enable conda on bash console
eval "$(conda shell.bash hook)"
# load modules or conda environments here
module load Miniconda3/4.9.2
conda activate jupyter-env

# get tunneling info
XDG_RUNTIME_DIR=""
node=$(hostname -s)
user=$(whoami)
cluster="teide-hpc"
port=8889

# print tunneling instructions jupyter-log
echo -e "
Command to create ssh tunnel from login node to compute node:
ssh -N -L localhost:${port}:localhost:${port} ${user}@nodeXXXX-Y.hpc.iter.es

Command to create ssh tunnel from your local machine to any login node:
ssh -N -L localhost:${port}:localhost:${port} ${user}@login1(2).hpc.iter.es

Use a Browser on your local machine to go to:
localhost:${port} (prefix w/ https:// if using password)
"

# Run Jupyter
jupyter-notebook --no-browser --port=${port} --ip=127.0.0.1
```

Este trabajo inicia Jupyter en el nodo de cómputo asignado y podemos acceder a él a través de un túnel ssh como lo hicimos en la sección anterior.

Enviamos el trabajo a la cola:

```
sbatch jupyter-notebook.sh
```

Una vez que se ejecuta el trabajo, se creará un log llamado *jupyter-notebook-.log*. Este log contiene información sobre cómo conectarse a Jupyter y el token necesario.

Para conectarse a Jupyter que se ejecuta en el nodo de cómputo, configure un túnel entre el nodo de cómputo y el nodo de inicio de sesión como:

```
ssh -N -L localhost:8889:localhost:8889 yourUser@nodeXXXX-X.hpc.iter.es
```

y entre su máquina local y el nodo de login:

```
ssh -N -L localhost:8889:localhost:8889 yourUser@loginX.hpc.iter.es
```

Para acceder a Jupyter Notebook, navegue <http://localhost:8889/>

38.1 Notas Importantes

⚠ Establecer una contraseña segura en Jupyter Notebook

Antes de conectarse a un servidor remoto con jupyter notebook, asegúrese de haber configurado jupyter con información de contraseña. Puede hacerlo editando `jupyter-notebook_config.json` que generalmente se encuentra en `~/.jupyter` o escribiendo:

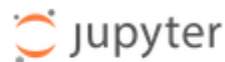
```
jupyter notebook password
```

```
Enter password:
```

```
Verify password:
```

```
[NotebookPasswordApp] Wrote hashed password to /home/yourUser/.jupyter/jupyter_notebook_config.json
```

Podrá acceder ahora con contraseña en lugar de autenticación de token en <http://localhost:8889/login>



Password:

Log in

⚠ Más consejos de seguridad aquí

Lea sobre la seguridad de Jupyter Notebook [aquí](#) y [aquí](#)

Cómo cambiar el directorio de inicio de Jupyter Notebook

De forma predeterminada, Jupyter usa su hogar como directorio de inicio predeterminado, pero es posible cambiarlo mediante dos métodos:

- Usando el argumento **--notebook-dir**:

```
jupyter-notebook --no-browser --port=8089 --ip=127.0.0.1 --notebook-dir=/home/my-user/data/my-dir
```

- Cambiar el directorio predeterminado generando un archivo de configuración
- Escriba el siguiente comando para crear una carpeta de configuración.

```
jupyter notebook --generate-config
```

- Abra el archivo `~/jupyter/jupyter_notebook_config.py*`.
- Busque el comentario, *The directory to use for notebooks and kernels*
- Descomente y reemplace la siguiente propiedad en el archivo con su directorio preferido.

```
## The directory to use for notebooks and kernels.  
# Default: ''  
c.NotebookApp.notebook_dir = '/home/my-user/data/my_notebooks'
```

- Vuelva a ejecutar Jupyter Notebook.

Asegúrese de cerrar su servidor de Jupyter cuando haya terminado

Para hacer esto, si está utilizando el modo *salloc*, puede volver a iniciar sesión en la sesión de *screen* que inició anteriormente donde se ejecuta el cuaderno jupyter y use *ctrl-C* para cerrar el cuaderno jupyter y *salir* para cerrar la sesión con el nodo. Si está utilizando el *sbatch*, puede usar el comando *scancel* .

38.2 Conceptos básicos de Jupyter Notebook

En el siguiente enlace puede encontrar una descripción de Jupyter Notebooks:

- [Notebook basics](#)

38.2.1 El dashboard Notebook



Files

Running

Clusters

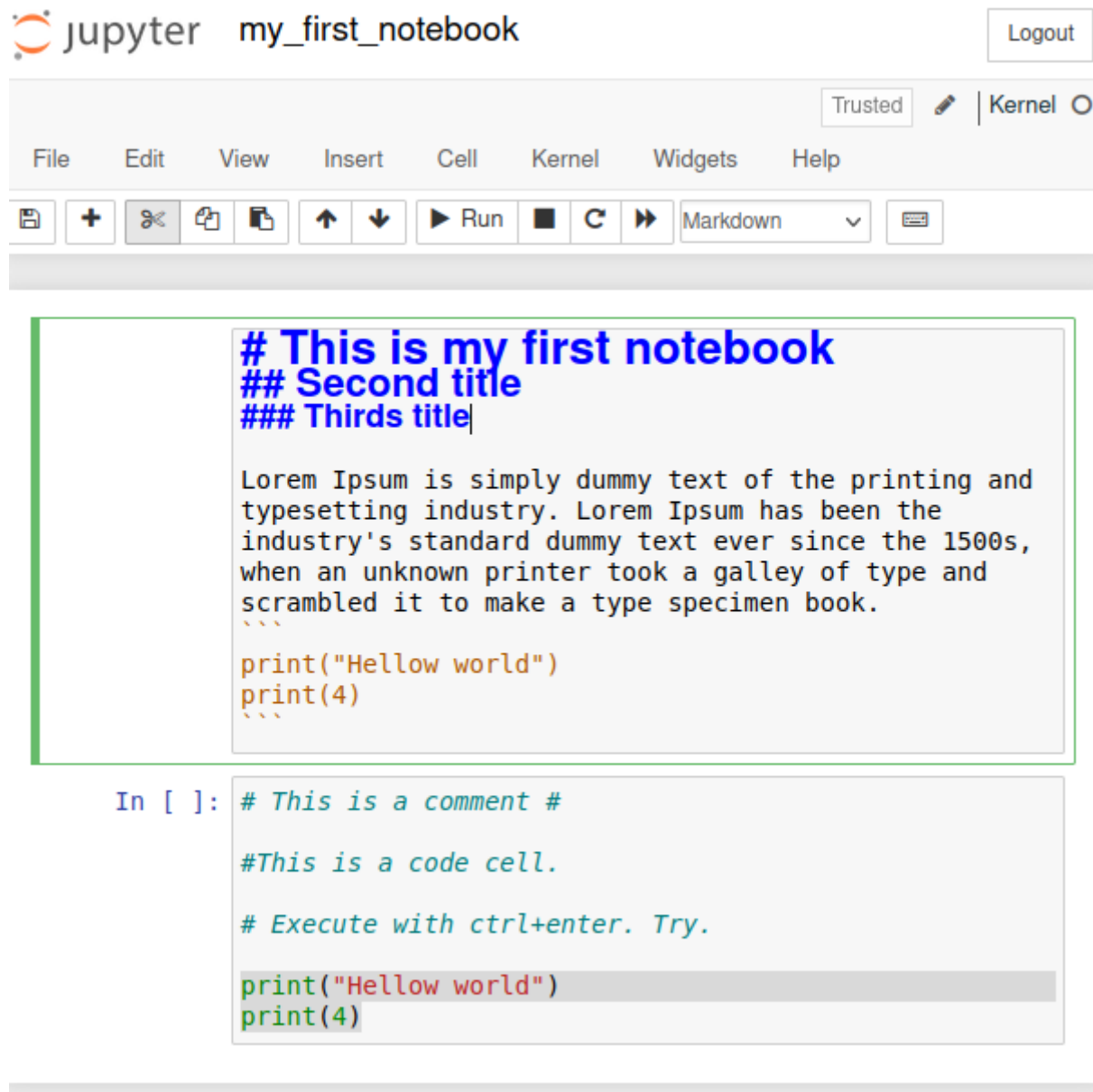
To import a notebook, drag the file onto the listing below or [click here](#).

New ▾



<input type="checkbox"/> ▾	🏠 / examples
<input type="checkbox"/>	..
<input type="checkbox"/>	Builtin Extensions
<input type="checkbox"/>	Customization
<input type="checkbox"/>	Embedding
<input type="checkbox"/>	IPython Kernel
<input type="checkbox"/>	Interactive Widgets
<input type="checkbox"/>	Notebook
<input type="checkbox"/>	Parallel Computing
<input type="checkbox"/>	images
<input type="checkbox"/>	utils
<input type="checkbox"/>	Index.ipynb

38.2.2 Ejemplo básico



The image shows a Jupyter Notebook interface. At the top, there is a header with the Jupyter logo, the text "jupyter my_first_notebook", and a "Logout" button. Below the header is a toolbar with a "Trusted" indicator, a pencil icon, and a "Kernel" status indicator. The main menu includes "File", "Edit", "View", "Insert", "Cell", "Kernel", "Widgets", and "Help". Below the menu is a secondary toolbar with icons for file operations, navigation, and execution, along with a "Markdown" dropdown menu.

The notebook content is displayed in a light gray box with a green border. It contains two cells:

- Markdown Cell:** Contains three levels of heading tags: `# This is my first notebook`, `## Second title`, and `### Thirds title`. Below the headings is a paragraph of Lorem Ipsum text, followed by a code block containing `print("Hellow world")` and `print(4)`.
- Code Cell:** Starts with `In []:` followed by three comment lines: `# This is a comment #`, `#This is a code cell.`, and `# Execute with ctrl+enter. Try.`. Below the comments is a code block containing `print("Hellow world")` and `print(4)`.

38.2.3 Ejecutar de un Notebook



This is my first notebook

Second title

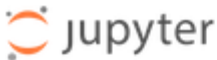
Thirds title

Lorem Ipsum is simply dummy text of the printing and typesetting industry. Lorem Ipsum has been the industry's standard dummy text ever since the 1500s, when an unknown printer took a galley of type and scrambled it to make a type specimen book.

ejemp

```
In [1]: # This is a comment #  
  
#This is a code cell.  
  
# Execute with ctrl+enter. Try.  
  
print("Hellow world")  
print(4)
```

```
Hellow world  
4
```


QuitLogout

Files Running Clusters


Currently running Jupyter processes

↻

Terminals ▾

There are no terminals running.

Notebooks ▾

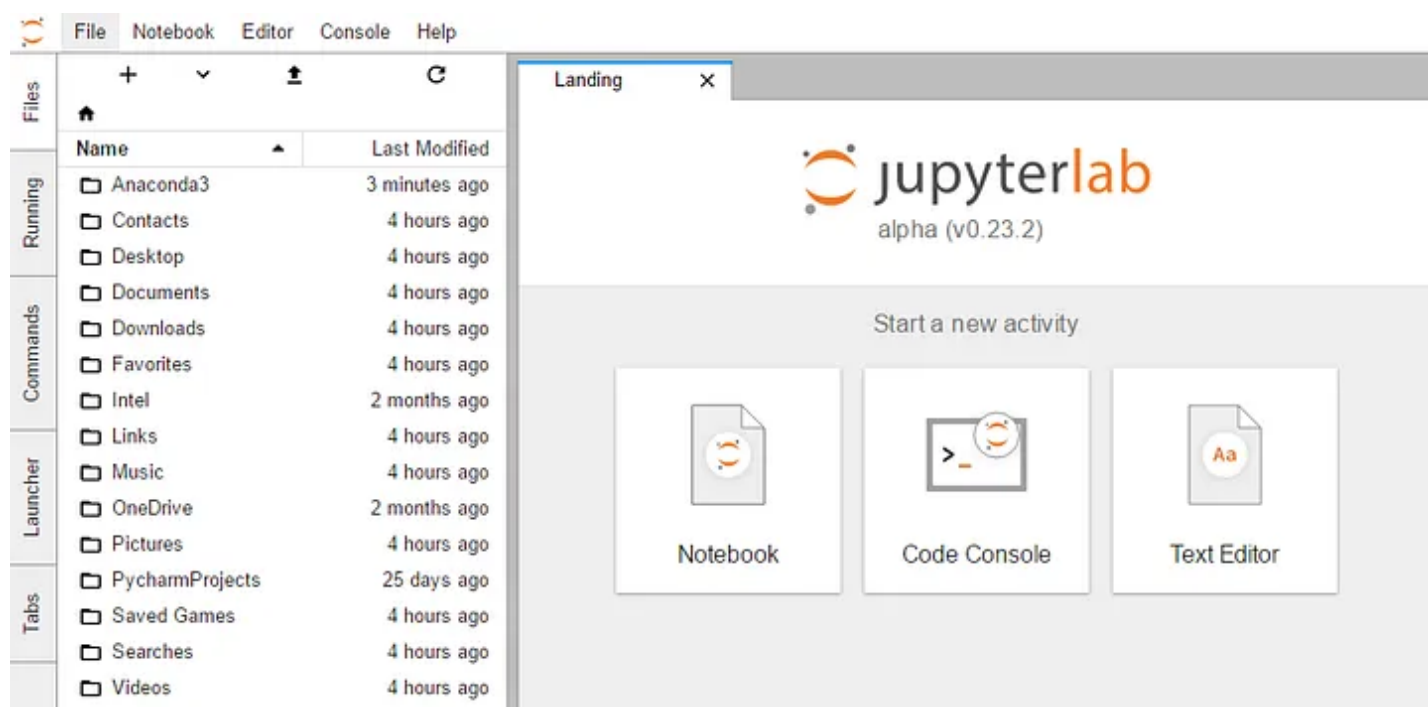
 data/jupyter_examples/my_first_notebook.ipynb

Python 3 (ipykernel)

Shutdown hace unos segundos

39 JupyterLab: La evolución de Jupyter Notebook

JupyterLab es la interfaz de usuario nueva generación para el proyecto Jupyter. Básicamente es un IDE con todas las funciones que tiene todo lo que siempre se quiso tener en los notebooks de Jupyter que le permite trabajar con documentos y actividades tal y como en Notebook, además de editores de texto, terminales y componentes personalizados de manera flexible, integrada y extensible.



Las principales características de JupyterLab son:

- *Arrastrar y soltar:*

La capacidad de reordenar celdas sin cortar y pegar es poderosa. También se siente más natural arrastrar y soltar dado que el código está organizado en celdas en notebooks.

- *Múltiples notebooks y kernels:*

La ejecución de varios notebooks al mismo tiempo ya existe con los Jupyter Notebook. Sin embargo, estos notebooks tenían que abrirse en varias ventanas del navegador. En JupyterLab, puede tener varios notebooks abiertos al mismo tiempo y en la misma ventana del navegador. Además, puede organizar sus Notebook como desee, lo que le brinda más flexibilidad. Otra buena característica es que es posible hacer que cada notebook se ejecute en su propio kernel, esto es poderoso cuando se ejecutan varios notebooks al mismo tiempo haciendo cosas diferentes.

- *Editor de markdown en tiempo real*

Con esta nueva función, puedo editar y ver en tiempo real la actualización de mis archivos markdown en JupyterLab. Esto acelera el proceso de edición y agiliza el trabajo.

The screenshot shows the JupyterLab interface with three main panels:

- File Explorer (Left):** Shows a directory structure with files like 'iris.csv', 'README.md', and 'Untitled.ipynb'.
- Code Editor (Center):** Displays a Python script in a notebook cell:


```
[4]: import pandas
df = pandas.read_csv('../data/iris.csv')
df.head(20)
```

 Below the code, the output is shown as a table:

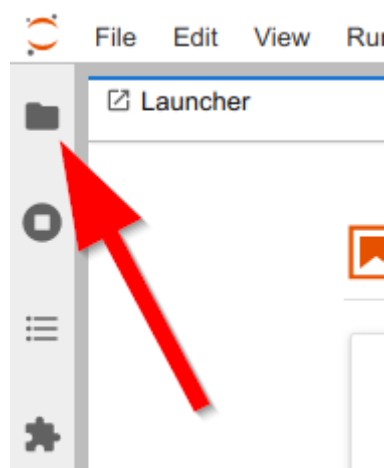
	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	se
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa
5	5.4	3.9	1.7	0.4	setosa
6	4.6	3.4	1.4	0.3	setosa
7	5.0	3.4	1.5	0.2	setosa
8	4.4	2.9	1.4	0.2	setosa
9	4.9	3.1	1.5	0.1	setosa
10	5.4	3.7	1.5	0.2	setosa
11	4.8	3.4	1.6	0.2	setosa
12	4.8	3.0	1.4	0.1	setosa
13	4.3	3.0	1.1	0.1	setosa
14	5.8	4.0	1.2	0.2	setosa
15	5.7	4.4	1.5	0.4	setosa
16	5.4	3.9	1.3	0.4	setosa
- Documentation Panel (Right):** Titled 'JupyterLab Demo', it provides information about the project, including its GitHub repository and a list of contributors: Project Jupyter, Bloomberg, and (then) Continuum.

- *Múltiples ventanas*

Con varias ventanas abiertas al mismo tiempo, puedo tener varios notebooks en las que estoy trabajando y luego usar una terminal dentro de JupyterLab

- *Explorador de archivos completo*

El explorador de archivos y el menú Archivo le permiten trabajar con archivos y directorios en su sistema. Esto incluye abrir, crear, eliminar, renombrar, descargar, copiar y compartir archivos y directorios.



- *Gestionar distintos kernel y terminales*
- *Búsqueda de comandos*
- *Visor rápido de archivos CSV*
- ...

39.1 Ejecutar JupyterLab de forma remota con Slurm

Al igual que Jupyter Notebook, JupyterLab se puede iniciar localmente y acceder a los archivos locales, o se pueden iniciar en una máquina remota.

De forma predeterminada, **Labs no es seguro y potencialmente expone los archivos locales de los usuarios a usuarios no deseados**. Recomendamos leer la sección [Ejecutar Jupyter Notebook con Slurm](#) donde puede encontrar varios consejos para usar Jupyter Notebook y que son de aplicación a JupyterLab.

Recuerda:

✗ No ejecute Jupyter en los nodos de login

La razón

✗ Internet no está disponible en los nodos de cómputo

¿Qué solución hay?

✗ No se puede acceder directamente a los nodos de cómputo

¿Por qué?

✗ Esta implementación no es el servidor multiusuario que andas buscando

¿Qué significa esto?

Puede utilizar JupyterLab de dos maneras diferentes:

- Crear un entorno conda o pip
- Uso como módulo

39.1.1 Crear un entorno Conda para instalar JupyterLab

✓ Crear un entorno Conda

Aquí tenemos varias recomendaciones para trabajar con entornos de conda que recomendamos leer. Para simplificar:

```
conda activate jupyterlabenv
conda install jupyterlab matplotlib ipykernel ipywidgets ipynb --channel conda-forge -y
```

39.1.2 Usar JupyterLab como módulo

Para buscar el módulo escribir:

```
module spider jupyter
```

```
-----  
JupyterLab: JupyterLab/3.1.6  
-----
```

Description:

JupyterLab is the next-generation user interface for Project Jupyter offering all the familiar building blocks of the classic Jupyter Notebook (notebook, terminal, text editor, file browser, rich outputs, etc.) in a flexible and powerful user interface.

JupyterLab will eventually replace the classic Jupyter Notebook.

You will need to load all module(s) on any one of the lines below before the "JupyterLab/3.1.6" module is available to load.

```
GCCcore/11.2.0
```

```
...
```

39.1.3 Cómo iniciar JupyterLab

Depende de cómo quieras iniciarlo. Si prefieres usarlo como módulo sólo tienes que cambiar la línea que activa el entorno por la que carga el módulo.

```
conda activate jupyter-labenv
```

with

```
module load GCCcore/11.2.0 JupyterLab/3.1.6
```

39.1.4 Ejecutar JupyterLab en un nodo de cómputo mediante *salloc*

```
salloc --nodes=1 --partition batch --ntasks=1 --mem=20G --time=12:00:00  
# or salloc -N 1 -p batch  
# or srun -p batch --pty bash
```

```
conda activate jupyter-labenv  
jupyter lab --no-browser --port=8890 --ip=127.0.0.1
```



Por defecto JupyterLab no es seguro y expone sus archivos locales a usuarios no deseados.

Ejecutar Labs de la forma habitual utiliza conexiones HTTP inseguras. En esta guía, presentamos una guía para ejecutarlo de forma más segura. Más info [aquí](#)

 **Tip**

Tenga en cuenta que el valor predeterminado es que JupyterLab abra automáticamente un navegador, pero no podemos hacerlo en un servidor remoto, por lo que omitimos esa función con el parámetro `--no-browser`).

39.1.5 Crear túnel SSH

Aquí hablamos sobre las razones por las que es necesario crear el túnel ssh y cómo hacerlo.

Inicia una segunda terminal, conecta a uno de los nodos de login y ejecuta crea un tunel entre el nodo login y el de ejecución.


```
ssh -N -L localhost:8889:localhost:8889 yourUser@nodeXXXX-Y.hpc.iter.es
```

En tu máquina local, establece el tunel contra el nodo de login como a continuación:

```
ssh -N -L localhost:8890:localhost:8890 yourUser@loginX.hpc.iter.es
```

!!! tip "Puedes usar los DNS del nodo o la IP

```
Change _nodeXXXX-X.hpc.iter.es_ or loginX.hpc.iter.es by node ip.
```

 **Introduce la dirección que te da Lab en tu navegador favorito**

<http://127.0.0.1:8890/lab?token=4c7eddd5770e27195808f4615d8e6f3de48b45c57169b69e>

Recomendamos securizar Jupyter Labs

39.1.6 Ejecutar JupyterLab en un nodo de cómputo mediante `sbatch`

```
#!/bin/bash
#SBATCH --nodes=1
#SBATCH --ntasks=1
#SBATCH --cpus-per-task=1
#SBATCH --mem=20G
#SBATCH --partition=batch
#SBATCH --constrains=sandy          # sandy, ilk (icelake)... architecture
#SBATCH --time=01:00:00
#SBATCH --job-name=jupyter-lab
#SBATCH --output=jupyter-lab-%j.out
#SBATCH --error=jupyter-lab-%j.err
#####
# Enable modules profile

# Enable conda on bash console
eval "$(conda shell.bash hook)"

# load modules or conda environments here
module load Miniconda3/4.9.2
conda activate jupyter-labenv

# get tunneling info
XDG_RUNTIME_DIR=""
node=$(hostname -s)
user=$(whoami)
cluster="teide-hpc"
port=8890

# print tunneling instructions jupyter-log
echo -e "
Command to create ssh tunnel from login node to compute node:
ssh -N -L localhost:${port}:localhost:${port} ${user}@nodeXXXX-Y.hpc.iter.es

Command to create ssh tunnel from your local machine to any login node:
ssh -N -L localhost:${port}:localhost:${port} ${user}@login1(2).hpc.iter.es

Use a Browser on your local machine to go to:
localhost:${port} (prefix w/ https:// if using password)
"

# Run Jupyter
jupyter lab --no-browser --port=${port} --ip=127.0.0.1
```

Para ejecutarlo

```
sbatch jupyter-lab.sbatch
```

Para acceder a JupyterLab navegar a <http://localhost:8890/>

39.2 Notas importantes

🔥 Establecer una contraseña segura en JupyterLab

De forma predeterminada, Lab inicia el servidor con la autenticación de token habilitada y este token se registra en el terminal, de modo que puede copiar y pegar la URL en su navegador. **Este token se puede usar sólo una vez** y se usa para configurar una cookie para su navegador una vez que se conecta. Después de que su navegador haya realizado su primera solicitud con este token único, el token se descarta y se establece una cookie en su navegador.

Como alternativa a la autenticación por token puede establecer una contraseña para su servidor. En posteriores inicios se le pedirá una contraseña y esta se almacenará cifrada en su archivo de configuración `jupyter_server_config.json` que generalmente se encuentra en `~/.jupyter`. Puede configurar su contraseña escribiendo:

```
jupyter server password
```

```
Enter password:  
Verify password:  
[JupyterPasswordApp] Wrote hashed password to /home/vdominguez/.jupyter/jupyter_server_config.json
```

You can access now with password instead of token authentication on <http://localhost:8890/lab>

🔥 Más consejos de seguridad aquí

<https://jupyter-server.readthedocs.io>

🔥 Cómo cambiar el directorio de inicio de Jupyter

Si los archivos de su notebook no están en el directorio donde está ejecutando Jupyter, puede pasar la ruta del directorio de trabajo como argumento al iniciar JupyterLab de la siguiente forma:

```
jupyter lab --notebook-dir=/home/yourUser/data/notebooks/ --preferred-dir /home/yourUser/data/myapp
```

Hay que tener en cuenta que la sesión de Jupyter siempre reside en un workspace. El workspace por defecto es la URL `/lab`:

```
bash http(s)://<server:port>/<lab-location>/lab
```

🔥 Asegúrese de cerrar JupyterLab cuando haya terminado

Para hacer esto, si está utilizando el modo `salloc`, puede volver a iniciar sesión en la sesión de `screen` que inició anteriormente donde se ejecuta el cuaderno jupyter y use `ctrl-C` para cerrar el servidor de Jupyter y `exit` para cerrar la sesión con el nodo.

Si está utilizando el modo `sbatch`, puede usar el comando `scancel`.

39.3 JupyterLab Extensions Manager

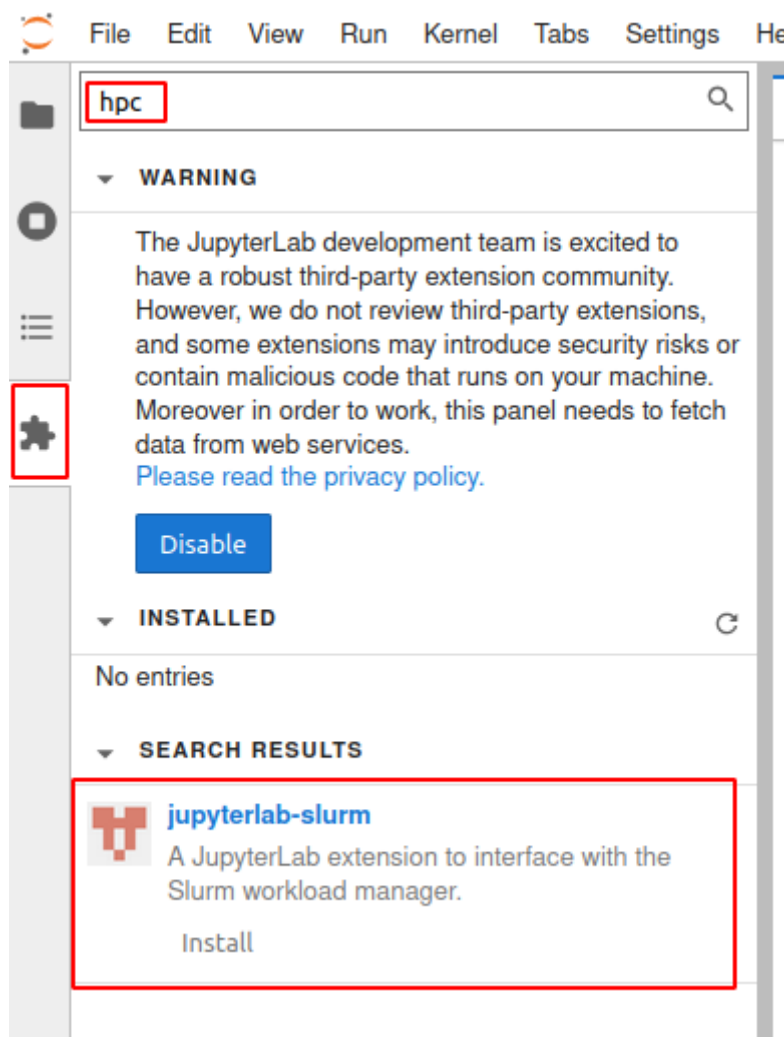
El gestor de extensiones de JupyterLab es simplemente un complemento plug-and-play que hace posibles más cosas que puede que necesite.

Técnicamente, la extensión JupyterLab es un paquete de JavaScript que puede agregar todo tipo de características interactivas a la interfaz de JupyterLab.

Hay un montón de extensiones de JupyterLab que tal vez quieras usar. Entre las más conocidas podemos destacar:

39.3.1 Jupyterlab-slurm

Una extensión de JupyterLab para interactuar con el administrador de carga de trabajo de Slurm.



39.3.2 Neptune-notebooks

Neptune es una herramienta para el seguimiento de experimentos, el registro de modelos, el control de versiones de datos y la supervisión de modelos en vivo.

Name	Owner	Latest checkpoint	Description	Compare
Untitled	neptuner	2019/10/26 19:25:55		↔ ↓ 🔗
tf-training	neptuner	2019/05/29 19:11:44		↔ ↓ 🔗
neural_style_tutorial	neptuner	2019/05/29 18:40:53		↔ ↓ 🔗
Multitask_GP_Regression	kamil	2019/05/29 18:28:14		↔ ↓ 🔗
dcgan_faces_tutorial	kamil	2019/05/29 18:05:22		↔ ↓ 🔗
AUC_Derivation	kamil	2019/05/29 16:47:56		↔ ↓ 🔗

Untitled.ipynb

[Details](#) [Checkpoints](#)

ID e095b056-59d7-4e9c-9005-c537aee8e327 [🔗](#)

Size 1.05 KB

Created 2019/10/26 19:25:53

Latest checkpoint 2019/10/26 19:25:55

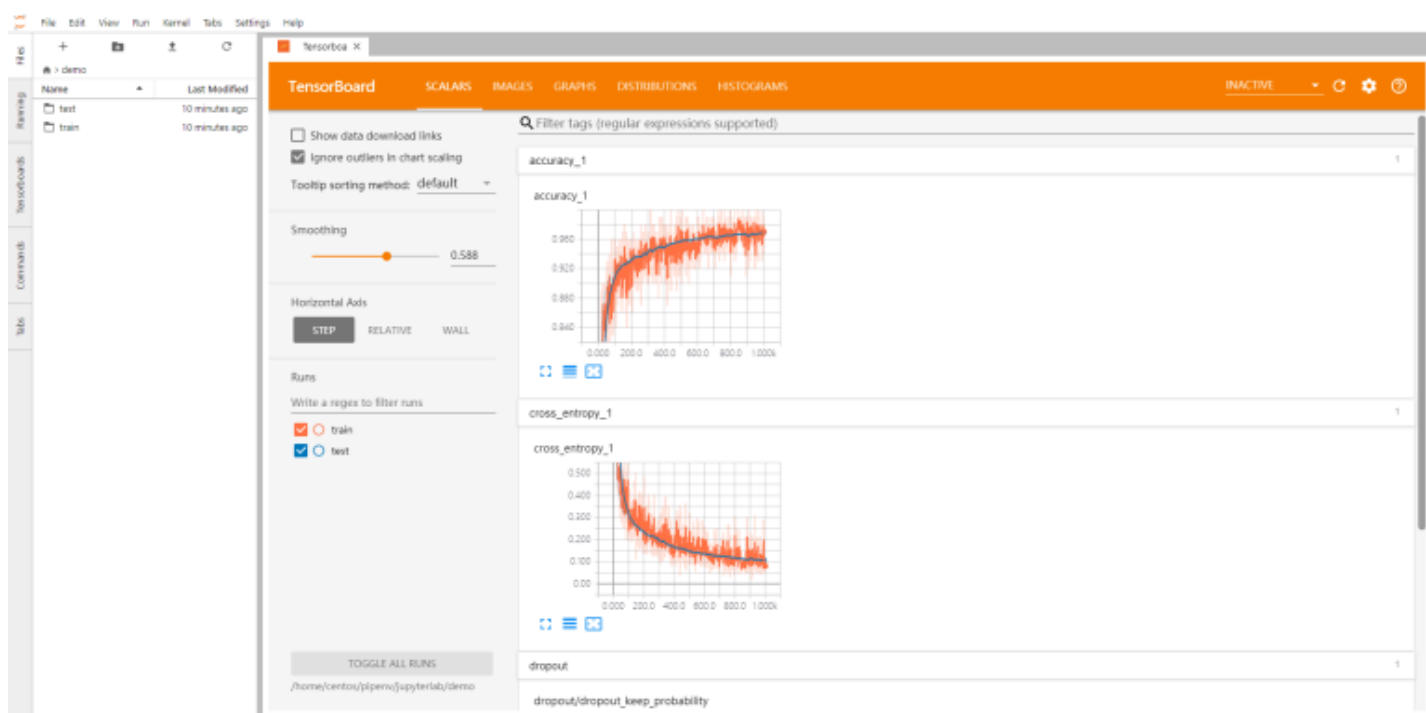
Owned by [neptuner](#)

Description

Experiments [Experiments started in notebook](#)

39.3.3 JupyterLab TensorBoard

JupyterLab TensorBoard es una extensión de interfaz para tensorboard en jupyterlab. Ayuda a colaborar entre jupyter notebook y tensorboard (una herramienta de visualización para tensorflow) al proporcionar una interfaz gráfica de usuario para iniciar, administrar y detener tensorboard en la interfaz de jupyter.



39.3.4 Jupyter ML-workspace

ML workspace es un entorno de desarrollo integrado todo en uno basado en la web dedicado al aprendizaje automático y la ciencia de datos.

Workspace Open Tool ▾

Files Running IPython Clusters Nbextensions

Select items to perform actions on them.

0 ▾ / tutorials

- ..
- images
- jupyter-basics-tutorial.ipynb
- jupyter-tips.ipynb
- numpy-tutorial.ipynb
- pandas-tutorial.ipynb
- python-basics-tutorial.ipynb
- visualization-tutorial.ipynb

VNC
Desktop GUI for the workspace

Ungit
Interactive Git interface

JupyterLab
Next-gen user interface for Jupyter

VS Code
Visual Studio Code webapp

Netdata
Monitor hardware resources

Glances
Monitor hardware resources

Filebrowser
Browse and manage workspace files

Access Port
Access any workspace internal port

SSH
Setup SSH connection to the workspace

39.3.5 JupyterLab Debugger

Debugger es una extensión de JupyterLab que funciona como un depurador visual para notebooks, consolas y archivos fuente de Jupyter. Puede ayudarlo a identificar y corregir errores.

File Edit View Run Kernel Tabs Settings Help

Launcher

Notebook

Python 3 xpython

Console

Python 3 xpython

Other

Terminal Text File Markdown File Show Contextual Help

VARIABLES

CALLSTACK

BREAKPOINTS

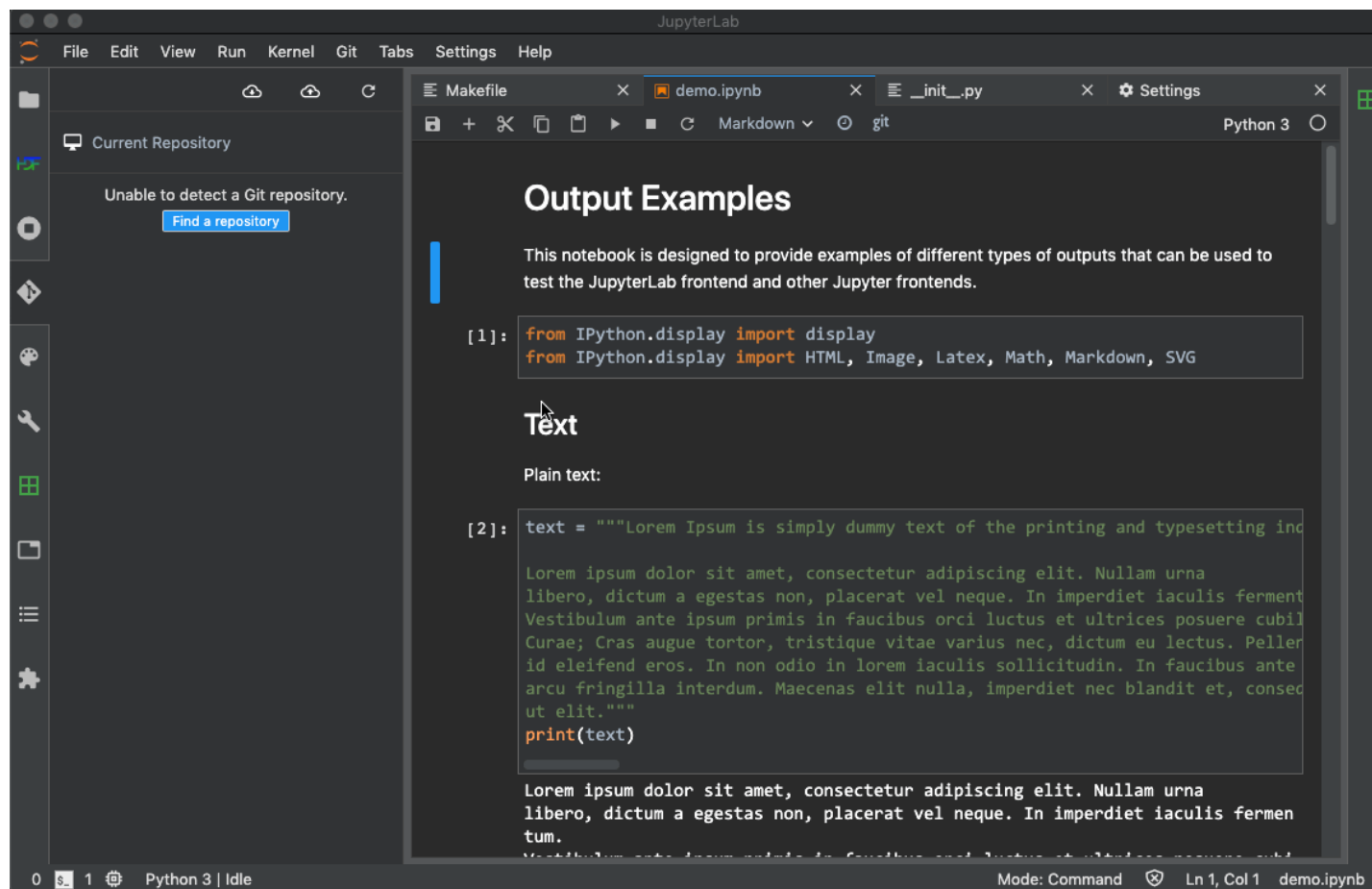
SOURCE

0 0

Launcher

39.3.6 JupyterLab Git

Esta es una extensión de JupyterLab para Git, un sistema de control de versiones distribuido gratuito y de código abierto. Le permite controlar la versión. Simplemente lo usa abriendo la extensión Git desde la pestaña Git en el panel izquierdo.



Otras extensiones destacables:

- JupyterLab LaTeX
- JupyterLab variableInspector
- JupyterLab plotly
- JupyterLab bokeh
- Jupyter Dash
- JupyterLab Table of Contents
- JupyterLab SQL

40 JupyterHub

Como decíamos en páginas previas, JupyterHub ofrece el poder de los Jupyter Notebook a grupos de usuarios. Si crees que es realmente lo que necesitas, contactanos por email en support@hpc.iter.es.

41 Singularity

Singularity es una plataforma que permite la ejecución de contenedores en entornos HPC. Docker es la herramienta más popular para ejecutar aplicaciones en contenedores, pero tal y como está diseñado, ponerlo en producción con la posibilidad de que sean los propios usuarios quiénes gestionen los contenedores, supone un riesgo de seguridad muy importante. Es por eso que nacieron alternativas como Singularity y otras.

Singularity cuenta con soporte para utilizar MPI y GPU para ejecutar contenedores y se pueden integrar en un script de Slurm sin problemas.

41.1 Utilizar Singularity en TeideHPC

Info

Actualmente disponemos de la versión v3.11 en TeideHPC.

Utilizando la herramienta de *modules* podemos cargar el software:

```
$ module load Singularity/3.11.0
$ singularity -h
```

Linux container platform optimized for High Performance Computing (HPC) and Enterprise Performance Computing (EPC)

Usage:

```
singularity [global options...]
```

Description:

Singularity containers provide an application virtualization layer enabling mobility of compute via both application and environment portability. With Singularity one is capable of building a root file system that runs on any other Linux system where Singularity is installed.

Options:

```
-c, --config string    specify a configuration file (for root or
                        unprivileged installation only) (default
                        "/share/easybuild/software/common/software/Singularity/3.11.0/etc/
singularity/singularity.conf")
-d, --debug            print debugging information (highest verbosity)
-h, --help            help for singularity
  --nocolor           print without color output (default False)
-q, --quiet           suppress normal output
-s, --silent          only print errors
-v, --verbose         print additional information
  --version           version for singularity
```

Available Commands:

```
build      Build a Singularity image
cache      Manage the local cache
capability  Manage Linux capabilities for users and groups
completion Generate the autocompletion script for the specified shell
config     Manage various singularity configuration (root user only)
delete     Deletes requested image from the library
exec       Run a command within a container
help       Help about any command
inspect    Show metadata for an image
instance   Manage containers running as services
key        Manage OpenPGP keys
oci        Manage OCI containers
overlay    Manage an EXT3 writable overlay image
plugin     Manage Singularity plugins
pull       Pull an image from a URI
push       Upload image to the provided URI
remote     Manage singularity remote endpoints, key servers and OCI/Docker registry credentials
run        Run the user-defined default command within a container
run-help   Show the user-defined help for an image
search     Search a Container Library for images
shell      Run a shell within a container
sif        Manipulate Singularity Image Format (SIF) images
sign       Add digital signature(s) to an image
test       Run the user-defined tests within a container
verify     Verify digital signature(s) within an image
version    Show the version for Singularity
```

Examples:

```
$ singularity help <command> [<subcommand>]
$ singularity help build
$ singularity help instance start
```

For additional [help](#) or support, please visit <https://www.sylabs.io/docs/>

41.2 Contenedores Docker

Singularity utiliza su propio formato de contenedores, `.sif`, teniendo que transformar los contenedores de Docker para que puedan ser utilizados, pero es algo que hace el propio programa sin necesidad de que el usuario tenga que intervenir.

41.2.1 Descargar un contenedor desde DockerHUB

Info

A la hora de descargar contenedores hay que hacerlo siempre desde los nodos de login que son los que tienen acceso a Internet y no desde los de cómputo.

En este ejemplo descargaremos una imagen desde DockerHub para ser utilizada en el clúster:

```
singularity pull docker://hello-world
INFO:   Converting OCI blobs to SIF format
INFO:   Starting build...
Getting image source signatures
Copying blob 8a49fdb3b6a5 done
Copying config 689808b082 done
Writing manifest to image destination
Storing signatures
2023/06/01 15:06:41 info unpack layer:
sha256:8a49fdb3b6a5ff2bd8ec6a86c05b2922a0f7454579ecc07637e94dfd1d0639b6
INFO:   Creating SIF file...
```

Nos descargará el archivo de la imagen del contenedor en el directorio actual:

```
hello-world_latest.sif
```

Si queremos, al igual que hacemos con docker, podemos especificar una versión en concreto:

```
singularity pull docker://hello-world:latest
```

Una vez descargado, podríamos probar a ejecutar el contenedor:

Warning

Recordamos a los usuarios que no se puede ejecutar software en los nodos de login y esto incluye contenedores. Para ello, tienen disponible el comando de slurm `salloc` para solicitar un nodo de manera interactiva y poder trabajar sin problemas.


```
singularity run hello-world_latest.sif
INFO:   Converting SIF file to temporary sandbox...
WARNING: passwd file doesn't exist in container, not updating
WARNING: group file doesn't exist in container, not updating

Hello from Docker!
This message shows that your installation appears to be working correctly.

To generate this message, Docker took the following steps:
 1. The Docker client contacted the Docker daemon.
 2. The Docker daemon pulled the "hello-world" image from the Docker Hub.
    (amd64)
 3. The Docker daemon created a new container from that image which runs the
    executable that produces the output you are currently reading.
 4. The Docker daemon streamed that output to the Docker client, which sent it
    to your terminal.

To try something more ambitious, you can run an Ubuntu container with:
$ docker run -it ubuntu bash

Share images, automate workflows, and more with a free Docker ID:
https://hub.docker.com/

For more examples and ideas, visit:
https://docs.docker.com/get-started/

INFO:   Cleaning up image...
```

En caso de necesitarlo, también disponemos del comando `build` para descargar contenedores de docker. La utilidad principal del comando `build` está en poder crear nuestros propios contenedores a partir de otros ya existentes o a partir de un fichero de definición.

```
singularity build tutu.sif docker://hello-world
INFO:   Starting build...
2023/06/01 14:58:33 info unpack layer:
sha256:719385e32844401d57ecfd3eacab360bf551a1491c05b85806ed8f1b08d792f6
INFO:   Creating SIF file...
INFO:   Build complete: tutu.sif
```

Y lo ejecutaríamos de la misma manera:

```
singularity run tutu.sif
INFO:   Converting SIF file to temporary sandbox...
WARNING: passwd file doesn't exist in container, not updating
WARNING: group file doesn't exist in container, not updating

Hello from Docker!
This message shows that your installation appears to be working correctly.

To generate this message, Docker took the following steps:
 1. The Docker client contacted the Docker daemon.
 2. The Docker daemon pulled the "hello-world" image from the Docker Hub.
    (amd64)
 3. The Docker daemon created a new container from that image which runs the
    executable that produces the output you are currently reading.
 4. The Docker daemon streamed that output to the Docker client, which sent it
    to your terminal.

To try something more ambitious, you can run an Ubuntu container with:
$ docker run -it ubuntu bash

Share images, automate workflows, and more with a free Docker ID:
https://hub.docker.com/

For more examples and ideas, visit:
https://docs.docker.com/get-started/

INFO:   Cleaning up image...
```

Si por lo que sea nuestra aplicación no está disponible en DockerHub y tenemos que construir el contenedor desde la fuente, lo podemos hacer en nuestro ordenador local, utilizando docker, y luego subir esa imange de Docker a nuestro /home en TeideHPC y crear el contenedor con singularity.

41.2.2 Ejecución de un contenedor

Como se ha visto, para ejecutar un contenedor con singularity tenemos el comando run:

```
singularity run mycontainer.sif <arg-1> <arg-2> ... <arg-N>
```

Pero también podemos ejecutar un contenedor pasándole un comando que se ejecute dentro de éste y argumentos:

```
singularity exec mycontainer.sif <command> <arg-1> <arg-2> ... <arg-N>
singularity exec tutu.sif python3 myscript.py 42
```

También podemos trabajar con el contenedor de manera interactiva al igual que hacemos con los contenedores de Docker. Para esto tenemos el comando shell:

```

singularity shell alpine_latest.sif
INFO:   Converting SIF file to temporary sandbox...
Singularity> cat /etc/os-release
NAME="Alpine Linux"
ID=alpine
VERSION_ID=3.18.0
PRETTY_NAME="Alpine Linux v3.18"
HOME_URL="https://alpinelinux.org/"
BUG_REPORT_URL="https://gitlab.alpinelinux.org/alpine/aports/-/issues"

Singularity> pwd
/home/vjuidias

Singularity> exit
INFO:   Cleaning up image...

```

Como vemos, podemos trabajar con el entorno del contenedor, pero seguimos en nuestro directorio, muy útil si queremos para trabajar con ficheros sin necesidad de copiarlos al contenedor.

41.3 Ejecutar un contenedor en Slurm

Para ejecutar singularity en slurm se ejecuta como cualquier otro software, cargando el módulo correspondiente y ejecutándolo:

```

#!/bin/bash -l

# Job name
#SBATCH -J singularity_job

# Partitiion to run the job
#SBATCH -p batch

# Number of nodes
#SBATCH --nodes=1

# Output files
#SBATCH -o out.log
#SBATCH -e err.log

#####

module load Singularity/3.11.0

singularity run $HOME/hello-world_latest.sif

```

Slurm tratará singularity como un software más, es decir, que le aplicará las mismas restricciones de recursos, en cuanto a cpu, memoria y tiempo que el resto del software.

41.3.1 MPI

 **Tip**

El soporte para MPI dependerá del software que vayamos a ejecutar, no de Singularity. Por tanto, pedimos a los usuarios que lean la documentación de su software detenidamente antes de ejecutar cualquier cosa.

Para poder utilizar MPI con singularity tenemos que cargar el módulo correspondiente y utilizar el comando de srun:

```
#!/bin/bash -l

# Job name
#SBATCH -J singularity_mpi

# Partition to run the job
#SBATCH -p batch

# Number of nodes
#SBATCH --nodes=2

# Output files
#SBATCH -o out.log
#SBATCH -e err.log
#####

module purge
module load Singularity/3.11.0
module load GCC/12.2.0 OpenMPI/4.1.4

srun singularity run $HOME/hello-world_latest.sif
```

41.3.2 GPU

 **Tip**

El soporte para GPU dependerá del software que vayamos a ejecutar, no de Singularity. Por tanto, pedimos a los usuarios que lean la documentación de su software detenidamente antes de ejecutar cualquier cosa.

Para utilizar GPU en la ejecución del software hay que utilizar el parámetro `--nv`:

```
singularity run --help
...
--nv          enable Nvidia support
```

Y para el caso de Slurm sería:

```
#!/bin/bash -l

# Job name
#SBATCH -J singularity_gpu

# Partitiion to run the job
#SBATCH -p gpu

# Number of nodes
##SBATCH --nodes=1

# Number of task
#SBATCH --cpus-per-task=4
#SBATCH --gpus=a100:1

# Output files
#SBATCH -o out.log
#SBATCH -e err.log

#####

module purge
module load Singularity/3.11.0

singularity run --nv $HOME/hello-world_latest.sif
```

41.4 Otras opciones

Al igual que con docker, tenemos la opción `-B` con la que podemos hacer un *bind* de un directorio de la máquina host en un directorio del contenedor:

```
singularity run -B /usr/lib/locale/:/usr/lib/locale -B "${PWD}/input":/input mycontainer.sif
<command> <arg-1> <arg-2> ... <arg-N>
```

IV. Infraestructura en la nube (IaaS)

IV.1 OpenNebula

42 Infraestructura como Servicio (IaaS)

IaaS (Infrastructure as a Service) es un modelo de pago por uso, escalable en función de las necesidades de almacenamiento y procesamiento. Proporciona infraestructuras de forma rápida y económica. El cliente paga exclusivamente por los servicios que utilice, sin realizar potentes inversiones en equipamiento de IT. Permite a las empresas aumentar la eficiencia, redundancia, seguridad y control de sus infraestructuras pero olvidándose de su instalación y mantenimiento de la cual se encarga TeideHPC.

Un proveedor de servicios informáticos en la nube, como TeideHPC, administra la infraestructura, mientras que usted compra, instala, configura y administra su propio software (sistemas operativos, middleware y aplicaciones).

Como características principales se pueden destacar las siguientes:

- En lugar de adquirir hardware directamente, los usuarios pagan por IaaS bajo demanda.
- La infraestructura es escalable, en función de las necesidades de almacenamiento y procesamiento.
- Ahorra a las empresas el coste de comprar y mantener su propio hardware.

42.1 OpenNebula

Es una plataforma cloud computing para administrar infraestructuras centro de datos heterogéneas distribuidas. La plataforma OpenNebula (ONE) gestiona la infraestructura virtual de un centro de datos para construir implementaciones privadas, públicas e híbridas de infraestructura como servicio (IaaS).

42.1.1 Acceso a OpenNebula

Después de conectarse al servidor VPN de TeideHPC, se le proporcionará una ruta de red privada para acceder a la interfaz web de OpenNebula donde podrá gestionar sus máquinas virtuales.

- <https://one.iter.es>
- <https://10.5.22.19>

Recuerde

Para acceder a cualquier dominio de TeideHPC como *one.iter.es* debe **estar conectado a la VPN**

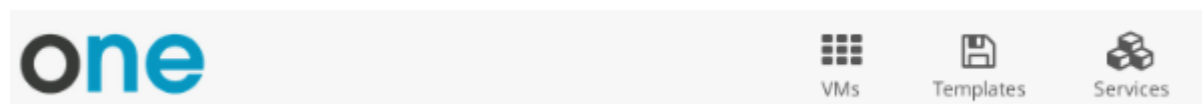


The image shows the OpenNebula Sunstone login interface. At the top, the logo "OpenNebula Sunstone" is displayed in blue and black. Below the logo is a login form with a light gray background and rounded corners. The form contains the following elements:

- Username:** A text input field containing the text "teide-x01".
- Password:** A password input field with a black dot mask and a vertical cursor line.
- Keep me logged in:** A checkbox with the text "Keep me logged in" next to it.
- Login:** A rectangular button with the text "Login" in bold black font.

42.1.2 Interfaz de OpenNebula

Después de iniciar sesión, la vista predeterminada brinda información sobre las máquinas virtuales (VM) y las cuotas del usuario.



El *menú de máquinas virtuales* proporciona información sobre las máquinas virtuales realmente definidas.

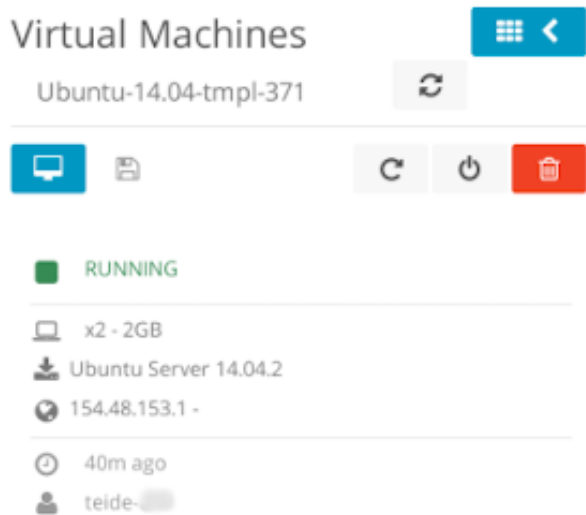
El *menú Plantillas* enumera las plantillas creadas a partir de máquinas virtuales que se utilizarán como fuente para nuevas máquinas virtuales.

El *menú Servicios* es la vista del usuario de una herramienta llamada *OneFlow*, que está diseñada para permitir la escalabilidad de las aplicaciones mediante el lanzamiento de nuevas máquinas virtuales basadas en diferentes parámetros, como la carga de los servicios en ejecución. Actualmente no está habilitado.

42.1.3 Máquinas Virtuales

El menú Máquinas virtuales enumera todas las máquinas virtuales asociadas con la cuenta o el grupo. Para cada VM se muestran los parámetros de capacidad y las interfaces de red.

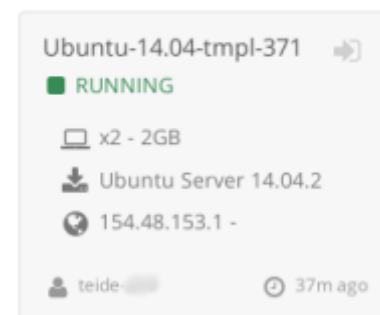
Se puede acceder a las máquinas virtuales a través de sus interfaces de red cuando están configuradas y a través de la consola VNC proporcionada por la interfaz web:



42.1.4 Plantillas

Se

Virtual Machines



pueden lanzar nuevas instancias de máquinas virtuales de dos formas, utilizando plantillas previamente definidas en el sistema o tomando instantáneas de máquinas virtuales que ya están en ejecución.

42.1.4.1 Crear instancias desde una plantilla

Virtual Machines



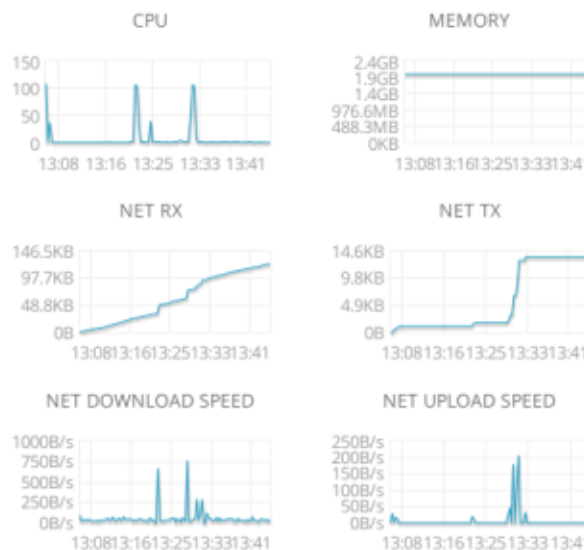
Usando el botón *más (+)* en el menú de VMs es posible crear una nueva máquina virtual con la configuración definida en una plantilla.

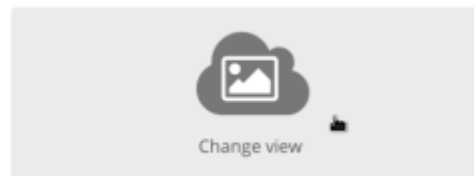
Las plantillas previamente definidas en la interfaz de la nube se enumeran en el menú *Sistema*, mientras que las plantillas creadas a partir de máquinas virtuales en ejecución se enumeran en el Menú *guardado*.

Con ambas opciones para crear una nueva VM, después de seleccionar la plantilla de origen, puede establecer el nombre de la VM e iniciar la creación con el botón *Crear*.

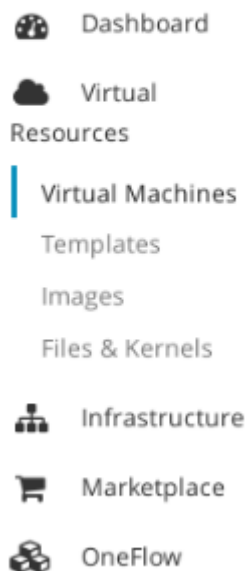
42.1.5 Vista de usuario

La interfaz web de OpenNebula (llamada Sunstone) proporciona vistas para diferentes roles de usuario en los sistemas de gestión de la nube. El usuario puede acceder a una vista extendida (vista de usuario) con la opción *Cambiar vista* en la configuración del perfil de usuario.





OpenNebula



The Dashboard gives a glimpse on how resources are used.

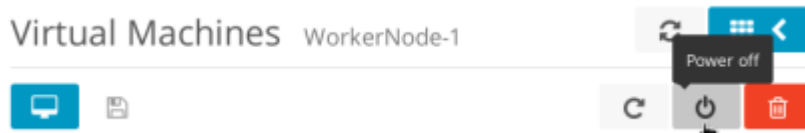
Bajo el menú *Virtual Resources* se puede acceder a la lista de máquinas virtuales que realmente se ejecutan en la nube, a las Plantillas que definen los parámetros de las nuevas máquinas virtuales e Imágenes asociadas a las máquinas virtuales o datos relacionados como imágenes en CDRROM de otros sistemas operativos.

The *Marketplace* es una herramienta que nos brinda acceso rápido a imágenes y dispositivos listos para ejecutar para probar una variedad de software y sistemas operativos, proporcionados por diferentes entidades editoras.

Finalmente, *OneFlow* es una herramienta avanzada para mejorar la escalabilidad de las aplicaciones mediante el lanzamiento de nuevas máquinas virtuales en función de diferentes parámetros, por ejemplo, la carga de los servicios en ejecución.

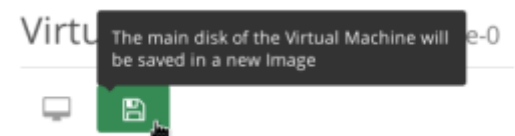
42.1.6 Crear Instantáneas

Es posible tomar una instantánea de una máquina virtual y usarla como plantilla para nuevas máquinas virtuales si tiene permisos para ello. Para hacer esto, el primer paso es apagar la VM. Puede enviar un comando de apagado con el botón de encendido dentro del menú de la máquina virtual o usando el comando de apagado en el shell de la máquina virtual.



Cuando la máquina virtual está apagada, el botón de guardar está habilitado y se puede realizar la copia de la imagen. Este procedimiento también crea una nueva plantilla en *Plantillas guardadas*.

El tiempo necesario para tomar la instantánea depende del tamaño de la imagen de la máquina virtual.



43 Gestión de Máquinas Virtuales

- Acceso a OpenNebula






[ver datos de conexión aquí](#)



The image shows the OpenNebula Sunstone login interface. At the top, the logo "OpenNebula Sunstone" is displayed, with "Open" in black and "Nebula" in blue. Below the logo is a rounded rectangular login form. Inside the form, there are two input fields: "Username" with the text "teide-x01" and "Password" with eight black dots. Below the password field is a checkbox labeled "Keep me logged in" which is unchecked. To the right of the checkbox is a "Login" button.

Al seleccionar una máquina virtual verá el siguiente menú:

OpenNebula

-  Dashboard
-  Virtual Resources
- Virtual Machines**
- Templates
- Images
- Files & Kernels
-  Infrastructure
-  Marketplace
-  OneFlow

- Acceso remoto a una MV mediante VNC

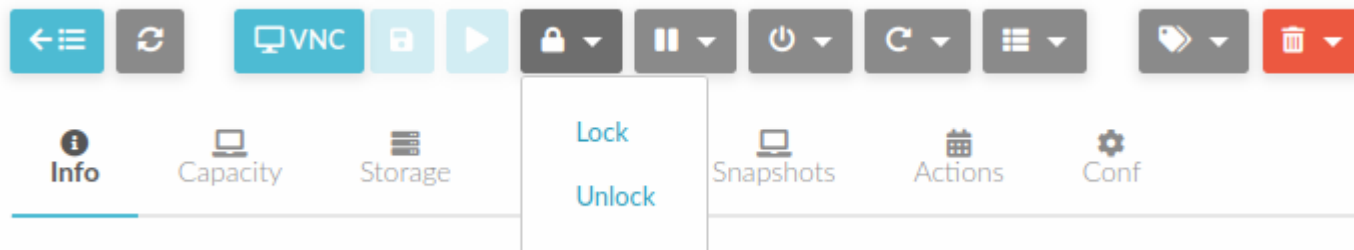
VNC Connected (unencrypted) to: QEMU (one-5088)

Send CtrlAltDel

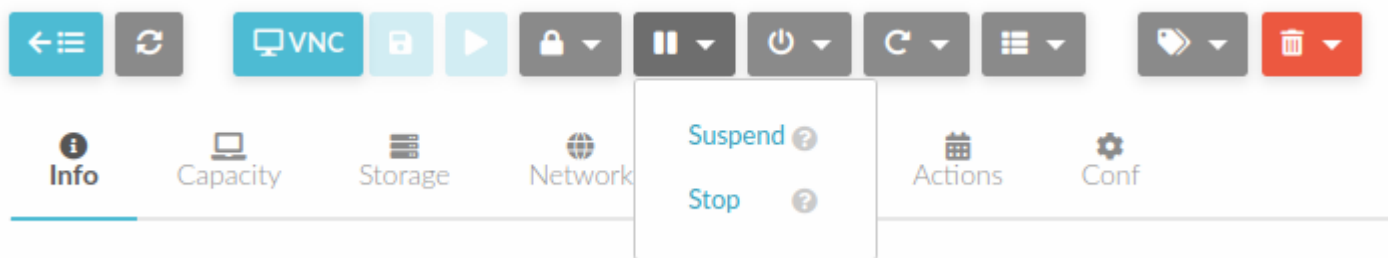


```
Ubuntu 20.04.3 LTS hpc-doc tty1
hpc-doc login: _
```

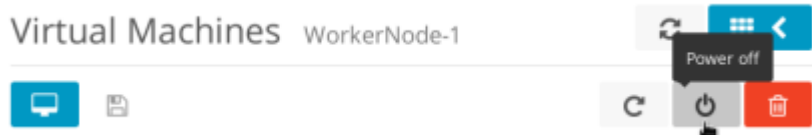
- Bloqueo/Desbloqueo de la MV



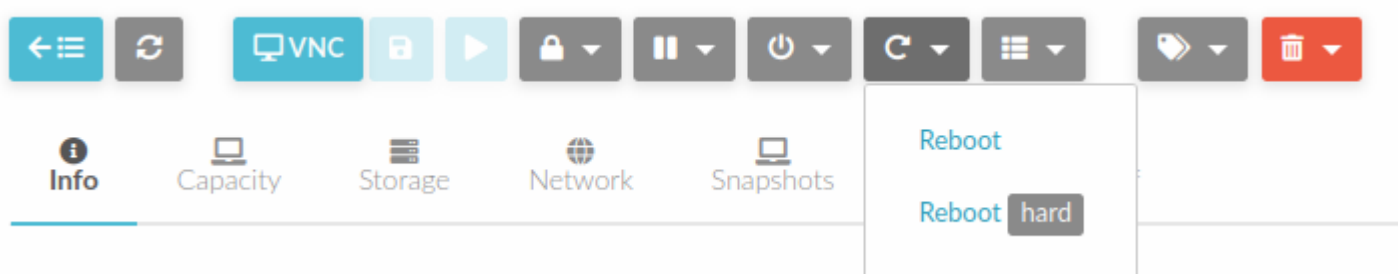
- Suspender/Parar la MV



• Apagar la MV



• Reiniciar la MV



V. Servicios

V.I NextCloud

44 Introducción a NextCloud



Nextcloud nace en 2016 como un fork de ownCloud. El motivo se sospecha que fueron ciertas diferencias culturales y de valores entre varios desarrolladores del proyecto, más próximos a la cultura del software libre, y otros miembros más centrados en el negocio en si, y no tanto en la comunidad.

Información

El proyecto está mantenido actualmente por Nextcloud GmbH junto con una gran comunicad de usuarios y desarrolladores que colaboran activamente en él. Además, como ocurre con otros tantos proyectos de software libre, en Github puede visualizarse su código fuente, y hacer un seguimiento de su desarrollo.

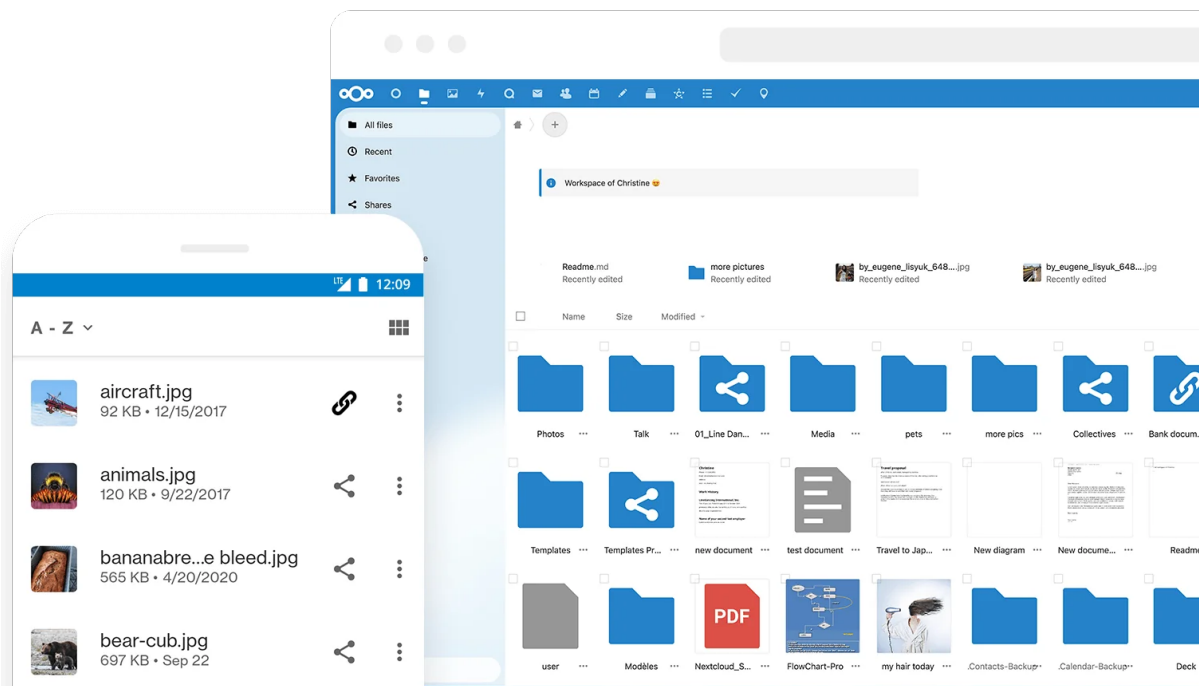
La visión de NextCloud es **recuperar el control de tu información** pero manteniendo a la vez todo el conjunto de funcionalidades que han popularizado otras plataformas y/o servicios SaaS privativos como Dropbox, Google Drive u Microsoft Office 365, por mencionar algunas de las más populares.

A diferencia de las plataformas SaaS centralizadas, lo que nos ofrece Nextcloud es directamente el software que nos permitirá construir nuestra propia plataforma, ya sea a nivel de usuarios individuales, o también a nivel de empresas o instituciones que deseen optar por una solución que pueda ser controlada, gestionada y por ellos.

Cita

Nextcloud pretende ser un completo centro de ocio y trabajo. Para ello, ofrece multiples capacidades de almacenamiento, colaboración, chat, comunicaciones audio/vídeo, calendario, contactos, y un largo etcétera.

44.1 ¿Qué es NextCloud?



Nextcloud es un software de código abierto que nos permite alojar archivos en la Nube y, además permite visualizarlos directamente desde un sitio web o apps, compartirlos, etc.

Con él, cualquier usuario con una cuenta puede subir información y se sincronizará con los demás usuarios en cualquiera de sus dispositivos.

” Cita

La diferencia más notable con respecto a OwnCloud es que NextCloud ofrece una moderna plataforma de colaboración de contenidos in situ con edición de documentos en tiempo real, videochat y trabajo en grupo en el móvil, el escritorio y la web.

44.2 ¿Para que sirve NextCloud?

i Información

La función básica de NextCloud es la de sustituir a OwnCloud y mejorar la experiencia de usuario.

El sistema de almacenamiento en la nube NextCloud tiene funcionalidades mejoradas con respecto a OwnCloud:

Mayor seguridad: Nextcloud incluye opciones de cifrado avanzado y autenticación de dos factores para proteger su información y evitar el acceso no autorizado.

Mayor flexibilidad: Nextcloud tiene una mayor cantidad de integraciones con otras aplicaciones y servicios, lo que significa que puede acceder a sus archivos de manera más rápida y fácil.

Mejoras en la colaboración: Nextcloud incluye herramientas de colaboración mejoradas, como la opción de dejar comentarios en documentos y ver la actividad reciente de otros usuarios. Esto puede hacer que sea más fácil trabajar en proyectos en equipo y mantenerse al día con el progreso de los demás.

Mejoras en la productividad: Nextcloud incluye aplicaciones y complementos que pueden ayudar a mejorar la productividad, como un calendario compartido, un administrador de tareas y una aplicación de correo electrónico.

Código abierto: Nextcloud es una plataforma de código abierto, lo que significa que está respaldada por una comunidad de desarrolladores y usuarios que contribuyen constantemente a su mejora y actualización. Esto garantiza que estaremos utilizando una herramienta de vanguardia y siempre estará a la vanguardia de las últimas tendencias y tecnologías.

Para más información tiene a su disposición la [guía oficial](#)

45 Características de la cuenta



45.1 ¿Cómo se accede?

Para acceder a Nextcloud, sigue estos pasos:

1. **Acceso a la Plataforma:** Abre tu navegador web e ingresa la URL de tu instancia de Nextcloud <https://nextcloud.iter.es>.



2. **Iniciar Sesión:** Introduce tus credenciales de inicio de sesión de ITER, generalmente consisten en un nombre de usuario y una contraseña.
3. **Panel de Control:** Una vez que hayas iniciado sesión, estarás en el panel de inicio de Nextcloud, desde donde podrás acceder y gestionar tus archivos y aplicaciones.

45.2 ¿Qué políticas de seguridad hay?

Nextcloud se esfuerza por garantizar la seguridad de tus datos. Algunas de las políticas de seguridad incluidas son:

Cifrado: Nextcloud permite cifrar tus archivos durante el almacenamiento y la transmisión para proteger tu privacidad.

Autenticación de Dos Factores (2FA): Puedes habilitar la autenticación de dos factores para una capa adicional de seguridad en el inicio de sesión.

Auditoría de Actividades: Nextcloud registra y mantiene un registro de actividades para rastrear los cambios en tus archivos y la actividad de los usuarios.

Control de Acceso: Configura políticas de acceso y permisos para garantizar que solo las personas autorizadas puedan acceder a los archivos y carpetas.

45.3 ¿Qué sucede cuando elimino un archivo?

Borrado de Archivos: Cuando borras un archivo, se envía a la papelera de reciclaje y se puede restaurar o eliminar permanentemente según tus necesidades.

Acceso desde Múltiples Ubicaciones: Puedes acceder y sincronizar tus archivos desde varias ubicaciones y dispositivos sin problemas.

Información

Normalmente si hay un acceso compartido y ambos acceden al mismo archivo el último en modificar ese archivo prevalecerá sobre los demás.

45.4 ¿Qué cuota de espacio tengo disponible?

Límite de Almacenamiento: Cada usuario tiene un límite predefinido de 20GB de espacio disponible para almacenar sus archivos

Visualización de la Cuota: Los usuarios pueden verificar cuánto espacio de almacenamiento han utilizado y cuánto les queda disponible en su cuenta. Esta información generalmente se muestra en el panel de control de Nextcloud.

Advertencias de Cuota Llena: Cuando un usuario se acerca al límite de su cuota de almacenamiento, Nextcloud puede mostrar advertencias para informar al usuario que está cerca de agotar su espacio asignado.

46 Características de la cuenta



46.1 ¿Qué tipos de cuentas hay disponibles?

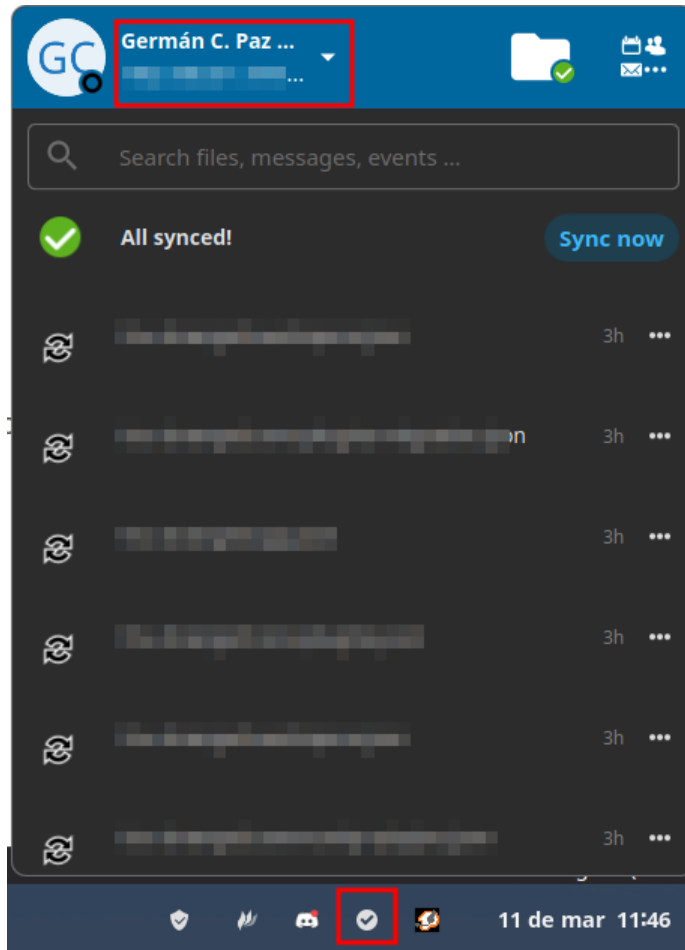
En el marco actual del ITER, se proporciona a cada integrante una cuenta de tipo **usuario**, la cual se ve limitada a una capacidad de almacenamiento de 20GB. Adicionalmente, se facilita una cuenta **departamental** por cada departamento, caracterizada por disponer de una cuota de almacenamiento superior en comparación con la cuenta de usuario.

La gestión de la cuenta departamental recae sobre una única persona designada dentro del departamento, la cual posee la responsabilidad de administrar el acceso y la distribución de los contenidos almacenados, garantizando así una adecuada coordinación y compartición de los recursos entre los miembros del departamento.

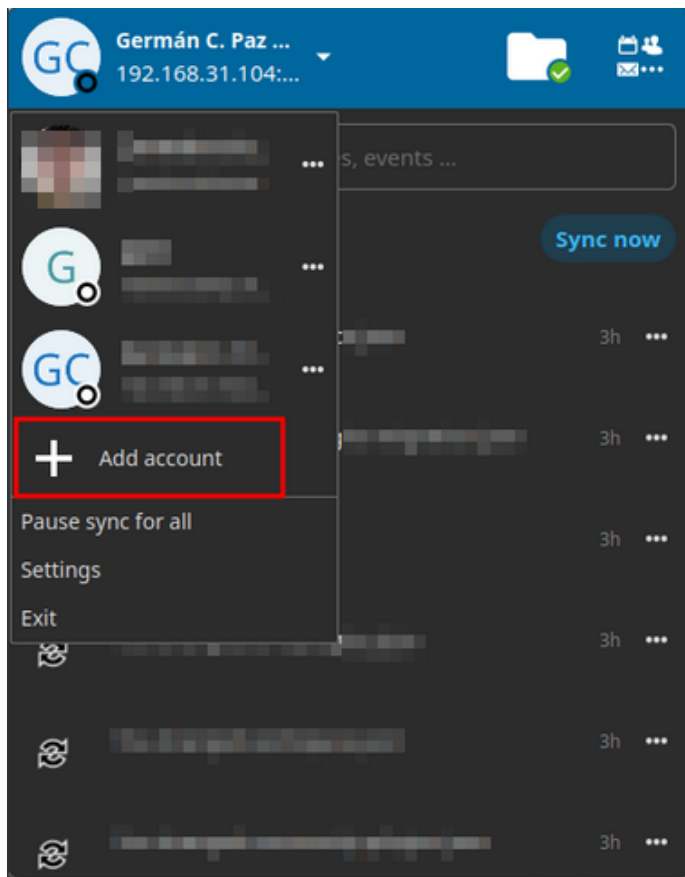
46.2 ¿Cómo configurar más de una cuenta en el cliente de Nextcloud?

Para tener disponibles ambas cuentas en el cliente de Nextcloud debes seguir los siguientes pasos descritos a continuación:

1. Haz clic con el botón izquierdo en el icono del sistema del cliente de Nextcloud y abre el dialogo principal. Ahora haz clic en el menú desplegable donde aparece el nombre de usuario. Deberías poder ver lo siguiente:

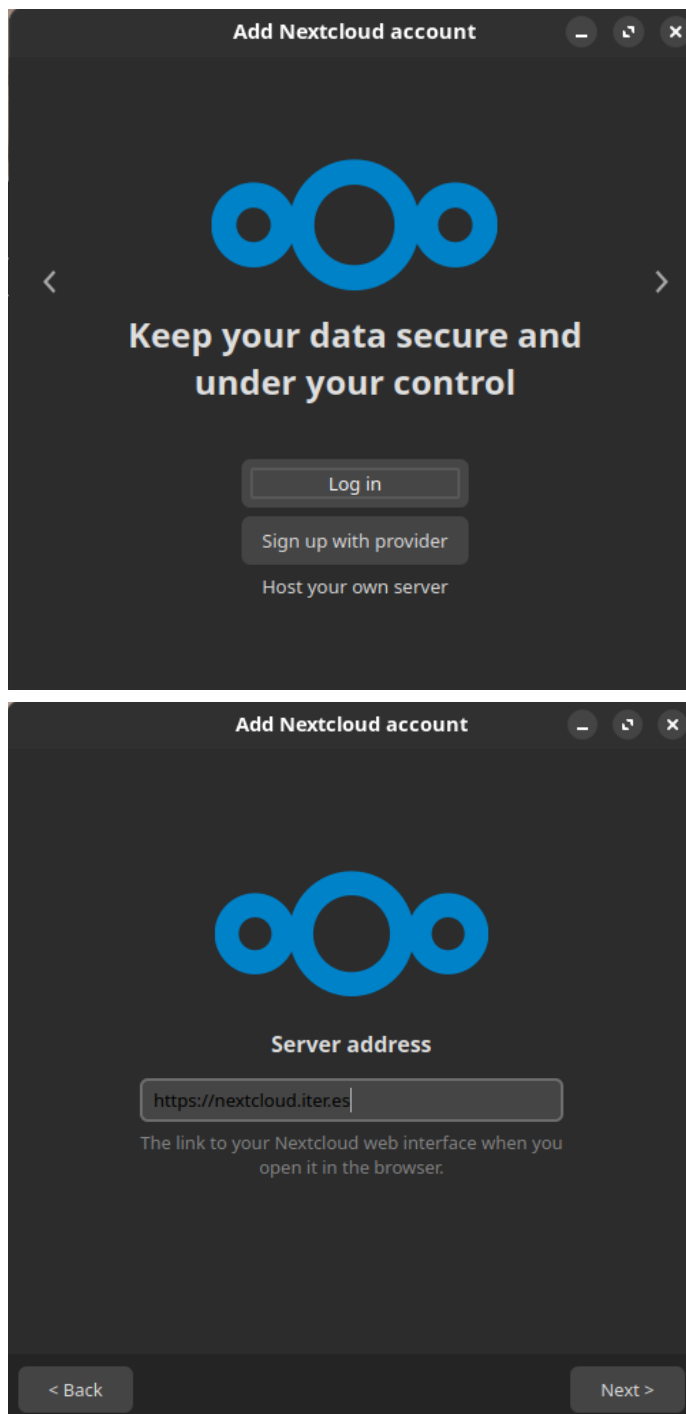


2. Haz clic sobre **agregar cuenta**:

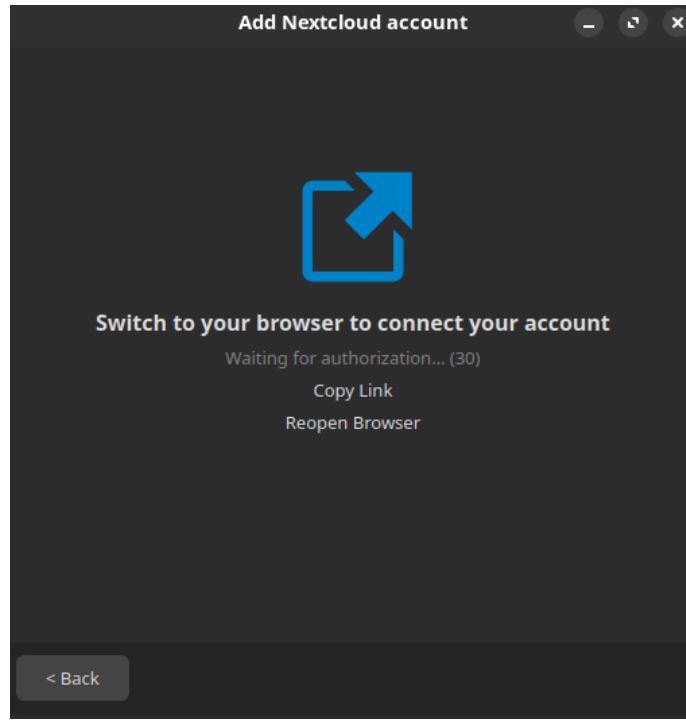


3.

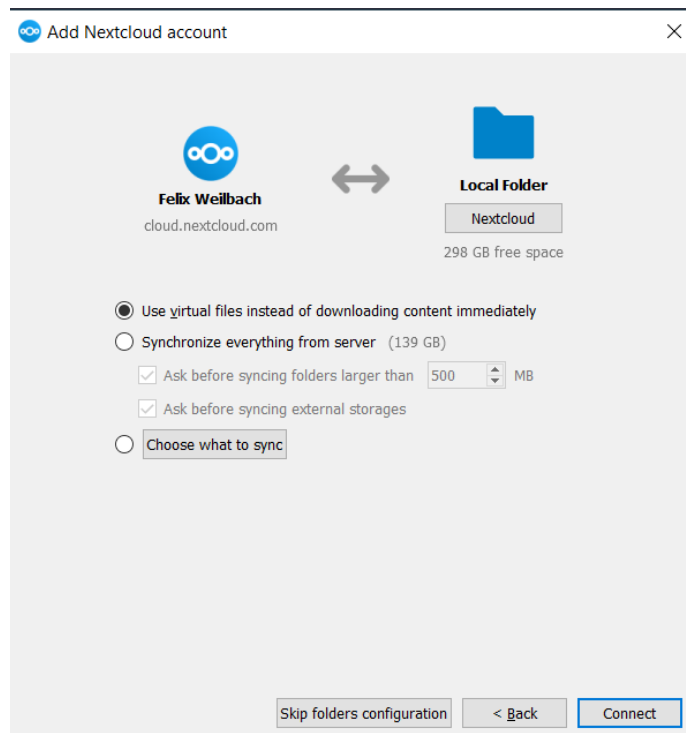
Deberías ver lo siguiente. Aquí deberás pulsar sobre "Log in" he introducir la URL. En nuestro caso la URL sería <https://nextcloud.iter.es>:



4. Ahora te aparecera la siguiente pantalla y te redirigirá al navegador para que allí inicies sesión con las credenciales del **usuario** o del **departamento** según corresponda.



5. Por último, elige la carpeta con la que el cliente de NextCloud debe sincronizar el contenido de tu cuenta de NextCloud.



Tras pulsar "Conectar", NextCloud Client comenzará con el proceso de sincronización.

6. Si todo ha ido bien deberías ver el siguiente icono en tus iconos del sistema:



47 Interfaz web de NextCloud

La interfaz web de usuario de NextCloud viene con un diseño renovado para adaptarse mejor a las necesidades de cada empresa y permite así establecer un fondo y logo corporativo.



Información

Desde aquí podrás acceder a tus archivos, así como: crear, previsualizar, editar, eliminar, compartir y recompartir archivos.

Nota

Tu administrador de NextCloud tiene la opción de desactivar algunas funciones. Si ves que alguna de las siguientes falta en tu sistema, pregunte directamente al administrador de tu NextCloud.

47.1 Iniciando sesión desde la interfaz web

Deberemos iniciar sesión con nuestras credenciales de ITER en la siguiente url: <https://nextcloud.iter.es>

The image shows a dark-themed login interface for Nextcloud. At the top, the title "Ingrese a Nextcloud" is displayed in white. Below the title, there are two input fields: "Nombre de cuenta o correo electrónico" and "Contraseña". The password field includes a toggle icon for visibility. A prominent blue button with a white arrow and the text "Iniciar sesión" is positioned below the fields. Underneath the button, there is a link that says "¿Contraseña olvidada?". At the bottom of the interface, the text "Iniciar sesión con dispositivo" is visible.

Ingrese a Nextcloud

Nombre de cuenta o correo electrónico

Contraseña

→ **Iniciar sesión**

[¿Contraseña olvidada?](#)

Iniciar sesión con dispositivo

Información

Actualmente NextCloud usa LDAP para la autenticación de usuarios por lo que únicamente necesitaremos iniciar sesión para estar registrados.

Una vez hemos iniciado sesión nos encontraremos con la siguiente pantalla:

Good afternoon, Bryan

Set status

Recommended files

- local_meeting_scratch.md
- nat_meeting_scratch.md
- keepass.kdbx
- call_script.md
- welcome_to_lincoln_ccl.md
- local_leaders.md
- salesforce_onboarding.md

Upcoming events

No more events today

- Group Leader Meeting
Tomorrow at 7:30 PM
- Date Night
Friday
- Pick Up Book
Friday
- Skype with Mom
in 8 days

Notes

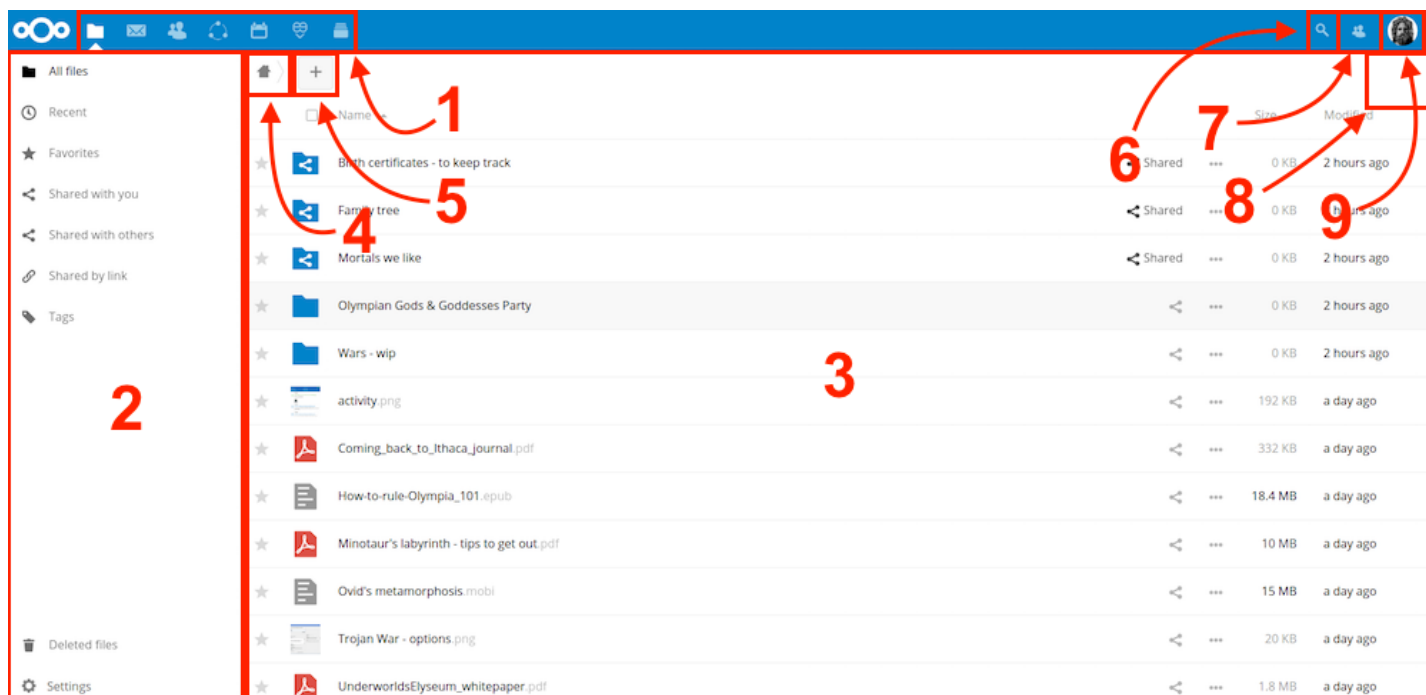
- local_meeting_scratch
GL call 6/2 "During these times, you sh...
- nat_meeting_scratch
UX Testing feedback What's the differe...
- call_script
6/29/2020 Fortenberry 202-225-4806...
- welcome_to_lincoln_ccl
Initial Welcome Letter Welcome to Lin...
- local_leaders
Dear , We are contacting you becaus...
- salesforce_onboarding
Jacob - Marketing / Campaigns Dashbo...
- typing
Here's the thing: it's practically imposib...

Customize

Información

Por defecto, la interfaz web de NextCloud se abre en el panel de control. Aquí podrás obtener las notificaciones y los cambios más recientes.

47.1.1 Navegando por la interfaz de usuario principal



Situándonos en "archivos" podemos añadir, eliminar y compartir archivos, y el administrador del servidor puede cambiar los privilegios de acceso.

La interfaz de usuario de Nextcloud contiene los siguientes campos y funciones:

- **Menú de selección de Apps (1):** Se encuentra en la esquina superior izquierda, y en él encontrará las aplicaciones que tiene disponibles en su instancia de Nextcloud. Al hacer clic en el icono de una aplicación se le redirigirá a ella.
- **Campo de información de la aplicación (2):** Situado en la barra lateral izquierda, ofrece filtros y tareas asociadas con la aplicación actual. Por ejemplo, en la aplicación Archivos aparecen una serie de filtros para encontrar archivos fácilmente, como por ejemplo archivos que han compartido con usted, y archivos que usted ha compartido. Cada aplicación tendrá elementos distintos.
- **Vista de la aplicación (3):** El campo central y principal de la interfaz de usuario de Nextcloud. Este campo muestra el contenido o características de la aplicación seleccionada.
- **Barra de navegación (4):** Situada sobre la ventana principal (la vista de la aplicación), esta barra muestra la ruta actual, que le permite moverse rápidamente a carpetas superiores, hasta el nivel raíz (la carpeta principal).
- **Botón nuevo (5):** Situado en la barra de navegación, el botón Nuevo permite crear nuevos archivos o carpetas, y subir archivos.

Información

También es posible arrastrar y soltar archivos desde su gestor de archivos a la vista de la aplicación Archivos para subirlos a su instancia (si su navegador soporta «drag and drop»)

- **Campo de búsqueda (6):** Haga clic en la lupa de la esquina superior derecha para buscar entre sus archivos
- **Menú de Contactos (7):** Le ofrece un resumen de sus contactos y usuarios en su servidor. En función de los detalles y aplicaciones disponibles, es posible empezar una videoconferencia con ellos o enviarles un correo electrónico.

- **Botón de vista en cuadrícula (8):** Este botón está formado por cuatro pequeños cuadrados, y alterna la vista en cuadrícula para carpetas y archivos.
- **Menú de Ajustes (9):** Para abrir el menú desplegable de Ajustes, haga clic en su foto de perfil, situada a la izquierda del campo de Búsqueda. Su página de Ajustes le ofrece las siguientes características y configuraciones:
 - Enlaces para la descarga de aplicaciones de escritorio y móviles
 - Uso del servidor y espacio disponible.
 - Gestión de contraseñas
 - Configuración de su nombre, correo electrónico y foto de perfil
 - Gestionar los navegadores y dispositivos conectados
 - Membresías en grupos
 - Configuración del idioma de la interfaz
 - Gestionar notificaciones
 - Identificador de nube federada y botones para compartir en redes sociales
 - Gestor de certificados SSL/TLS para almacenamiento externo
 - Configuración de verificación en dos pasos
 - Información sobre la versión de Nextcloud

Para más información tiene a su disposición la [guía oficial](#)

48 Compartiendo archivos en NextCloud

Los usuarios de Nextcloud pueden compartir archivos y carpetas. Los objetivos posibles son:

- Enlaces públicos
- Usuarios
- Grupos
- Círculos
- Conversaciones
- Usuarios o grupos en servidores Nextcloud federados

Al hacer clic en el icono de compartir en cualquier archivo o carpeta, se abre la vista de detalles a la derecha, donde la pestaña Compartir tiene el foco.

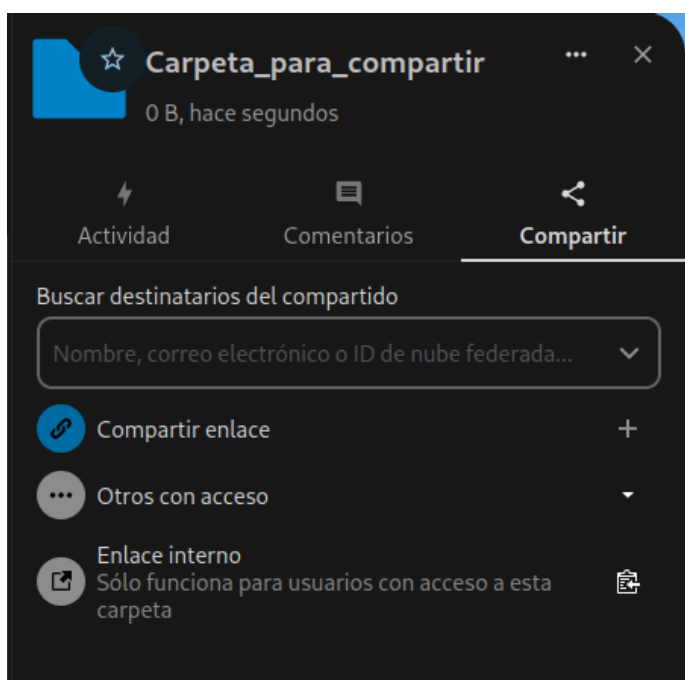
48.1 ¿Cómo se crean Espacios Compartidos?

Para crear espacios compartidos en Nextcloud y compartirlos con otros usuarios, sigue estos pasos:

1. **Selecciona un Archivo o Carpeta:** En tu panel de control, selecciona un archivo o carpeta que desees compartir.

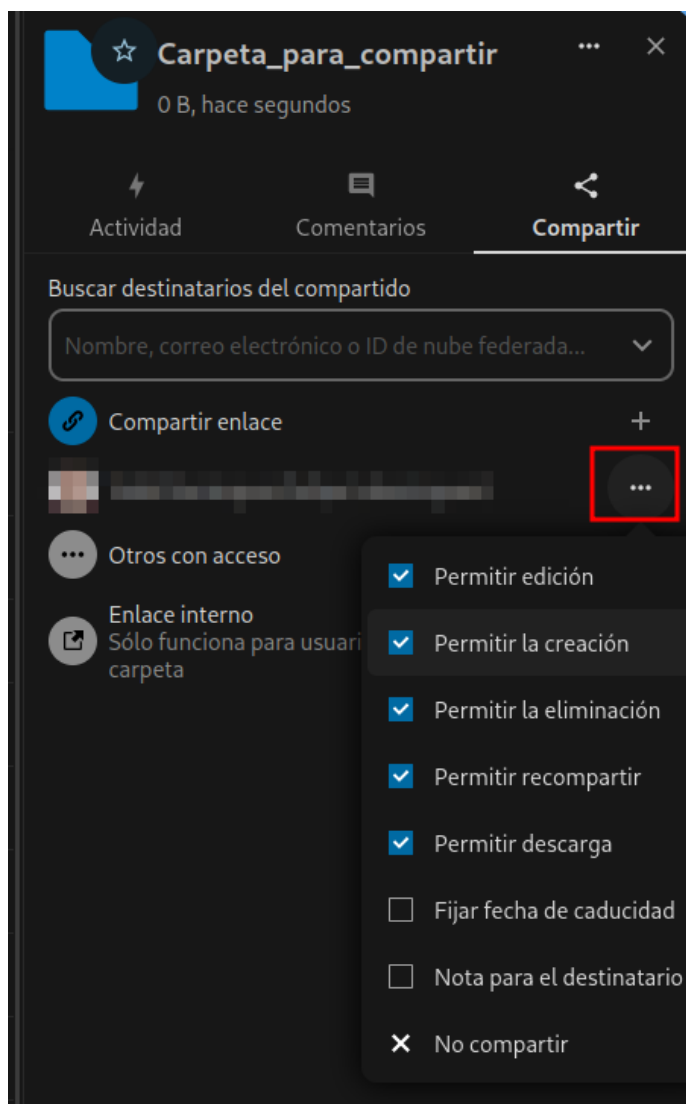


2. **Opciones de Compartir:** Haz clic en la opción "Compartir" y elige con quién desees compartir el archivo o carpeta. Puedes compartirlo con usuarios específicos o generar un enlace de acceso público.

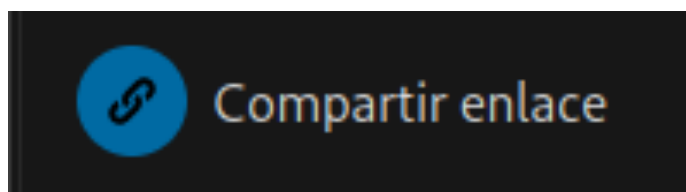


- 3.

Configura los Permisos: Define los permisos de acceso, como "Lectura" o "Edición", y establece una fecha de vencimiento si es necesario.



4. **Envía la Invitación:** Notifica a los usuarios con quienes compartes el archivo mediante correos electrónicos o compartiendo el enlace generado.



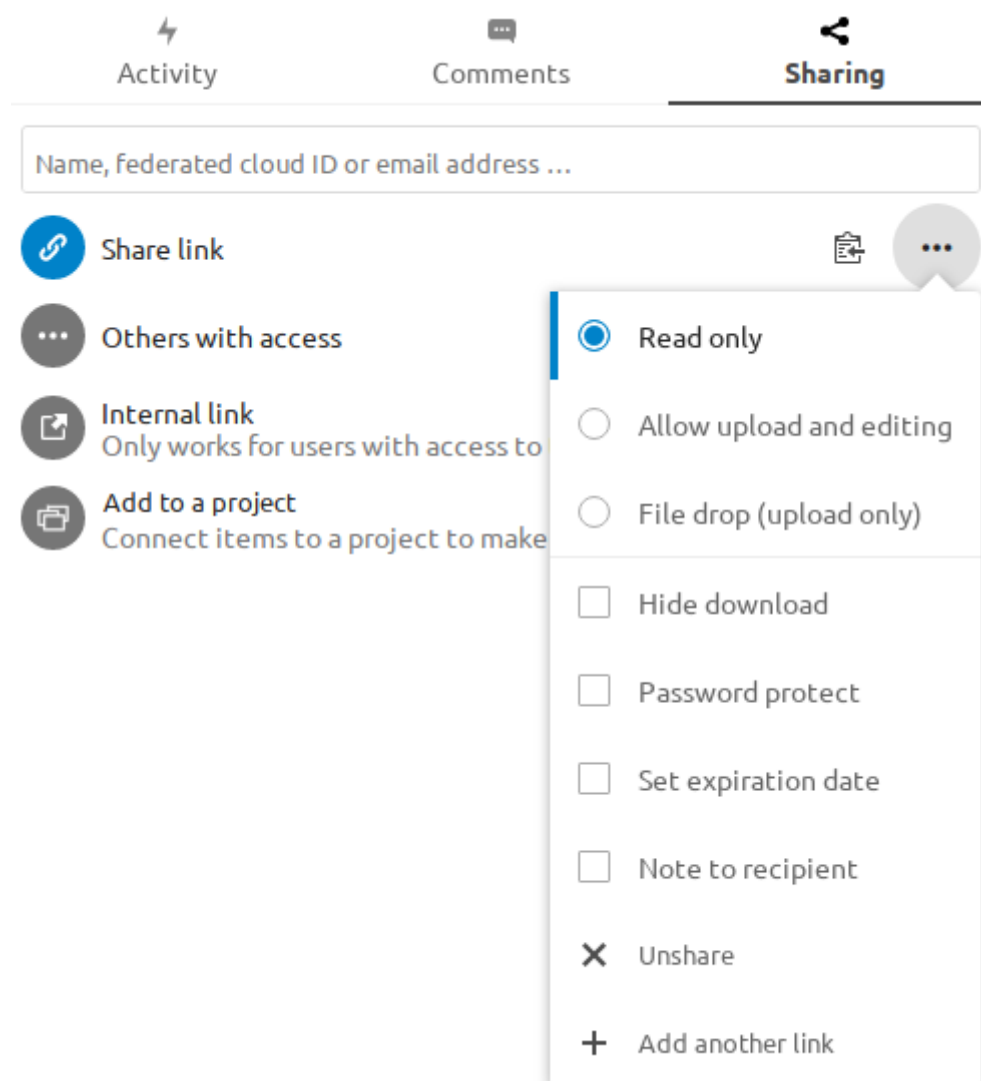
48.2 Formas de compartir los archivos

48.2.1 Enlaces públicos compartidos

Puedes compartir archivos y carpetas a través de enlaces públicos.

Se creará un token aleatorio de 15 dígitos. El enlace tendrá el aspecto de <https://cloud.example.com/s/yxcFKRWBjQYYzp4>.

Existen varias opciones para compartir carpetas públicas:



- **Sólo lectura** para permitir la visualización y descarga
- **Permitir subir y editar**
- Con **Subir archivos**, el compartidor sólo puede subir archivos a una carpeta sin ver los archivos que ya están en esa carpeta.
- **Ocultar descarga** oculta los botones de descarga y las opciones predeterminadas del botón derecho del navegador para dificultar la descarga para el sharee
- **Proteger con contraseña**
- **Establecer fecha de caducidad** desactivará automáticamente el recurso compartido
- **Nota para el destinatario**
- **Descompartir** para revertir la acción
- **Añadir otro enlace** para crear varios enlaces públicos con diferentes derechos

Nota

La protección por contraseña y la caducidad de los archivos no se propagan mediante el uso compartido de archivos federados en las versiones actuales de Nextcloud. Esto se ha ajustado en Nextcloud 22.

48.2.2 Uso compartido interno con usuarios y grupos

Al compartir con usuarios, grupos, o círculos, los derechos para los archivos o el contenido de las carpetas son ajustables:

The screenshot displays the sharing interface in Nextcloud. At the top, there is a search bar labeled "Search for share recipients" with the placeholder text "Name, email, or Federated Cloud ID ...". Below the search bar, there are several sharing options: "Share link" (with a plus sign), "Marie-Jeanne" (with a three-dot menu icon), "Others with access" (with a three-dot menu icon), and "Internal link" (with a note: "Only works for users with access to the resource").

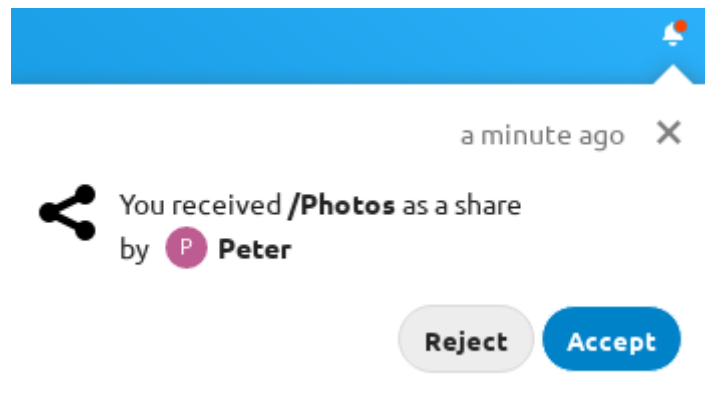
Below these options is a section titled "Related resources" with the text "Anything shared with the same group here". Underneath, there is a list of resource icons and names, which are partially blurred.

A permissions dropdown menu is open over the "Marie-Jeanne" option. The menu contains the following items, all with checked boxes:

- Allow editing
- Allow creating
- Allow deleting
- Allow resharing
- Allow download
- Set expiration date

The expiration date is set to "01 / 08 / 2023" with a close icon. Below the expiration date, there is an unchecked checkbox for "Note to recipient" and an "X Unshare" option at the bottom.

Como compartidor, puedes configurar si quieres aceptar automáticamente todas las comparticiones entrantes y que se añadan a tu carpeta raíz, o si quieres que se te pregunte cada vez si quieres aceptar o rechazar la compartición.



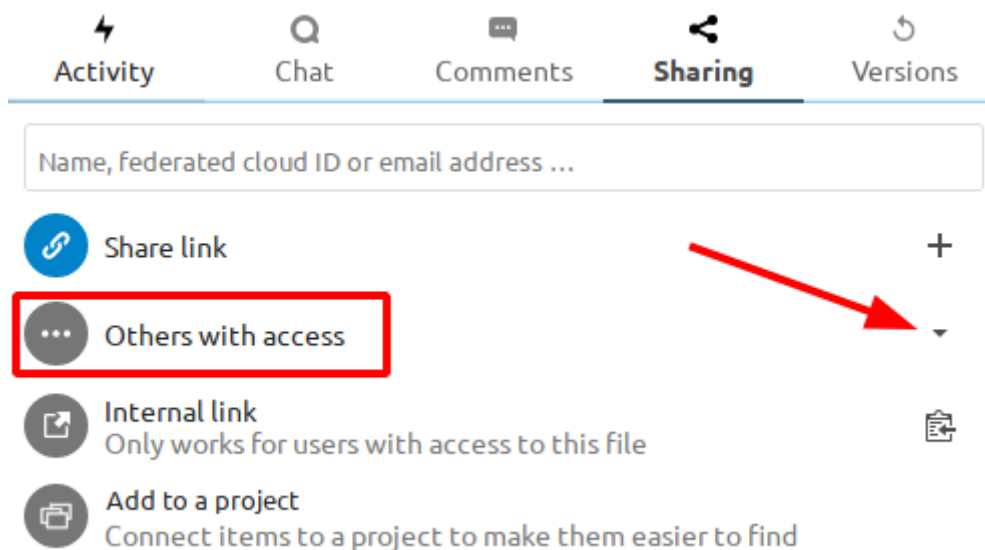
Para ajustar la aceptación, ve a **Ajustes > Personal > Compartir**:

Sharing

Accept user and group shares by default

48.2.3 Otros con acceso

Para saber si un archivo o carpeta es accesible a otros a través de la compartición de un nivel jerárquico de carpeta superior, haz clic en Otros con acceso en la pestaña de compartición:



La lista muestra todos los usuarios, grupos, etc. a los que se ha dado acceso al objeto actual mediante el uso compartido de una carpeta superior en la jerarquía:

Activity Chat Comments **Sharing** Versions

Name, federated cloud ID or email address ...

- Share link +
- Others with access ^
- Marketing ...
- Company chat ...
- (Shared link) ...
- Peter ...
- Internal link
Only works for users with access to this file
- Add to a project
Connect items to a project to make them easi...

Added by Jim

Via "Company"

Unshare

Haz clic en los tres puntos para:

- ver quién inició el intercambio
- ver dónde se inició la compartición (haz clic para navegar hasta la carpeta, siempre que tengas acceso a ella)
- anular la compartición inicial (sólo accesible para el propietario de la compartición)

Nota

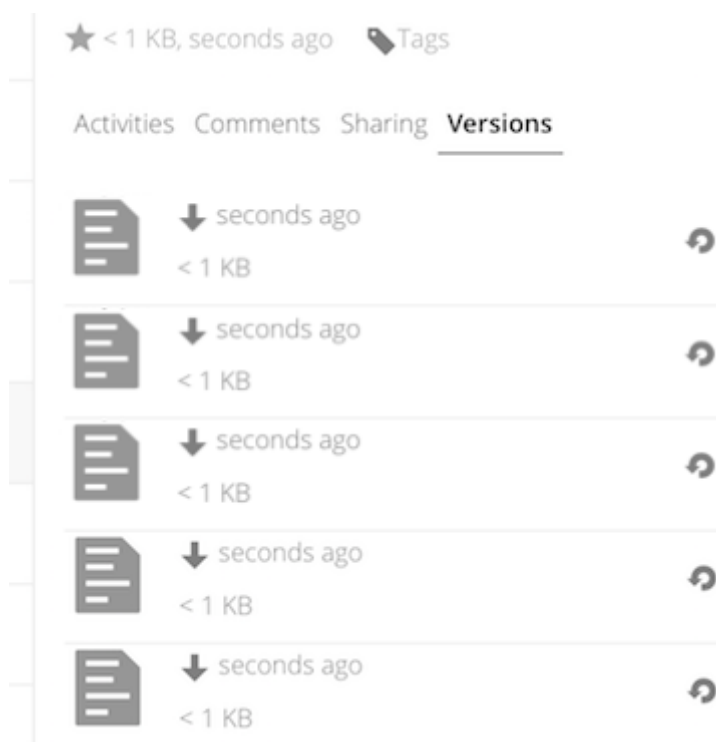
Esta información sólo es visible para el propietario de un archivo/carpeta o para los usuarios con derechos para volver a compartir.

Para más información tiene a su disposición la [guía oficial](#)

49 Control de versiones de NextCloud

Nextcloud admite un sencillo sistema de control de versiones para los archivos. **El control de versiones crea copias de seguridad de los archivos a las que se puede acceder a través de la pestaña "Versiones" de la barra lateral "Detalles".**

Esta pestaña contiene el historial del archivo, en el que se puede hacer retroceder un archivo a cualquier versión anterior. Los cambios realizados en intervalos superiores a dos minutos se guardan en **data/[usuario]/files_versions**.



Para restaurar una versión concreta de un archivo, haz clic en la flecha circular de la derecha. Haz clic en la marca de tiempo para descargarla.

La aplicación de versionado expira las versiones antiguas automáticamente para asegurarse de que el usuario no se queda sin espacio. Este patrón se utiliza para eliminar versiones antiguas:

- Para el primer segundo mantenemos una versión
- Durante los primeros 10 segundos Nextcloud mantiene una versión cada 2 segundos
- Durante el primer minuto Nextcloud guarda una versión cada 10 segundos
- Durante la primera hora, Nextcloud conserva una versión cada minuto.
- Durante las primeras 24 horas Nextcloud guarda una versión cada hora
- Durante los primeros 30 días Nextcloud guarda una versión cada día
- Después de los primeros 30 días, Nextcloud conserva una versión cada semana.

Las versiones se ajustan siguiendo este patrón cada vez que se crea una nueva versión.

La aplicación de versiones nunca utiliza más del 50% del espacio libre disponible del usuario. Si las versiones almacenadas superan este límite, Nextcloud borra las versiones más antiguas hasta que vuelve a alcanzar el límite de espacio en disco.

Para más información tiene a su disposición la [guía oficial](#)

50 Aplicación de escritorio en NextCloud

Con esta herramienta **pueden mantenerse sincronizados las carpetas y los archivos de su cuenta de NextCloud con una carpeta en el equipo local del usuario**, de modo similar a como funcionan otras herramientas de software (por ejemplo Dropbox).

Información

La sincronización continua hacia y desde el servidor NextCloud proporciona facilidad de uso combinada con un amplio control de acceso.

50.1 Instalación de la aplicación de escritorio

El cliente de Nextcloud es multiplataforma, puede ser instalado fácilmente en sistemas operativos Linux, Windows y macOS.

50.1.1 Instalación en Linux

En sistemas operativos Linux el cliente de Nextcloud puede instalarse ejecutando los siguientes comandos en una terminal:

```
$ sudo add-apt-repository ppa:nextcloud-devs/client
$ sudo apt update
$ sudo apt install nextcloud-client
```

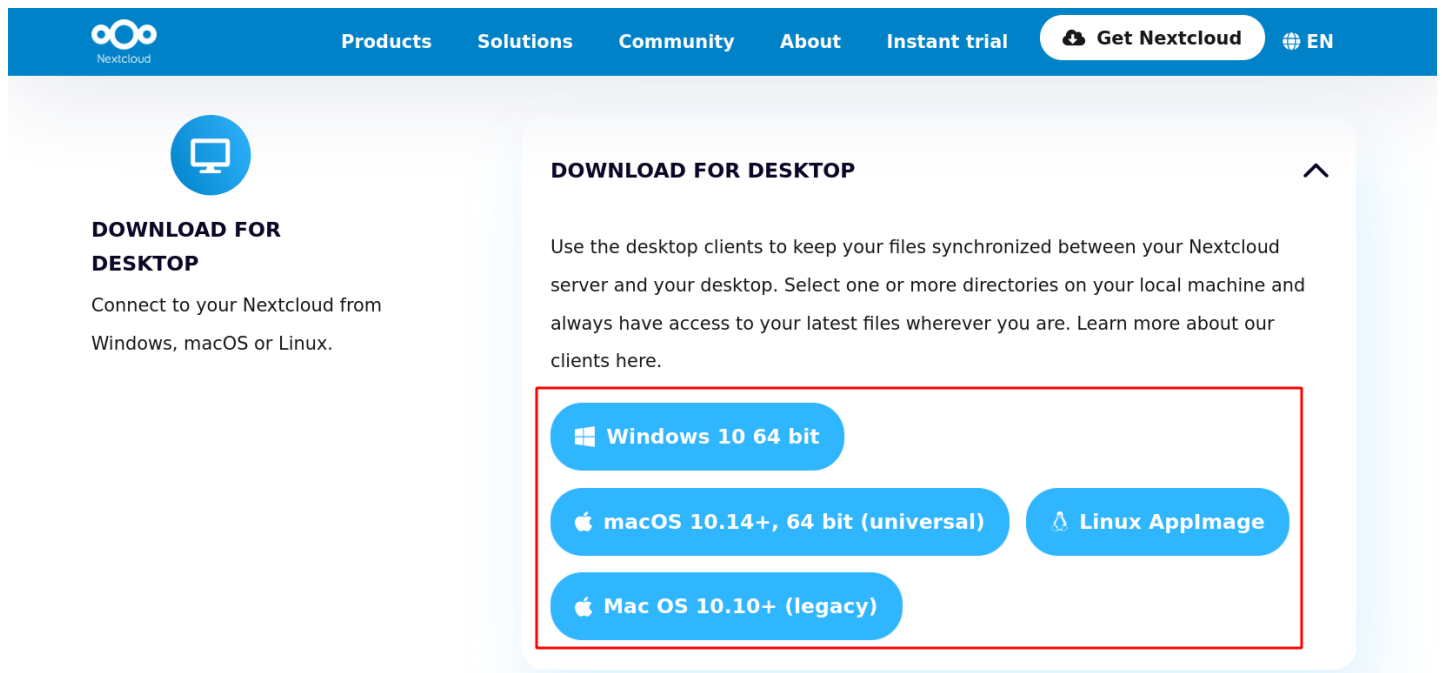
También es posible descargar el instalador del programa para Linux desde la página de descarga del cliente de Nextcloud: <https://nextcloud.com/install/#install-clients>

Información

Una vez que se instala, es necesario configurar la herramienta en la primera ejecución.

50.1.2 Instalación en Mac OS X y Windows

La instalación en Mac OS X y Windows es la misma que para cualquier aplicación de software:



DOWNLOAD FOR DESKTOP

Use the desktop clients to keep your files synchronized between your Nextcloud server and your desktop. Select one or more directories on your local machine and always have access to your latest files wherever you are. Learn more about our clients here.

- Windows 10 64 bit
- macOS 10.14+, 64 bit (universal)
- Linux AppImage
- Mac OS 10.10+ (legacy)

1. Descargue el instalador en la siguiente [URL](#)
2. Haga doble clic en él para iniciar la instalación y siga el asistente de instalación.

50.2 Uso de la aplicación de escritorio

La aplicación de escritorio NextCloud permanece en segundo plano y es visible como un icono en la bandeja del sistema (Windows, KDE), en la barra de menús (macOS) o en el área de notificación (Linux).

50.2.1 Iconos de aplicación utilizados



El indicador de estado utiliza iconos para indicar el estado actual de su sincronización. El círculo verde con la marca de verificación blanca te indica que tu sincronización está en curso y que estás conectado a tu servidor Nextcloud.



El icono azul con los semicírculos blancos significa que la sincronización está en curso.



El icono amarillo con las líneas paralelas le indica que la sincronización se ha pausado. (Lo más probable es que por usted).



El icono gris con tres puntos blancos significa que tu aplicación de escritorio ha perdido la conexión con tu servidor de NextCloud.



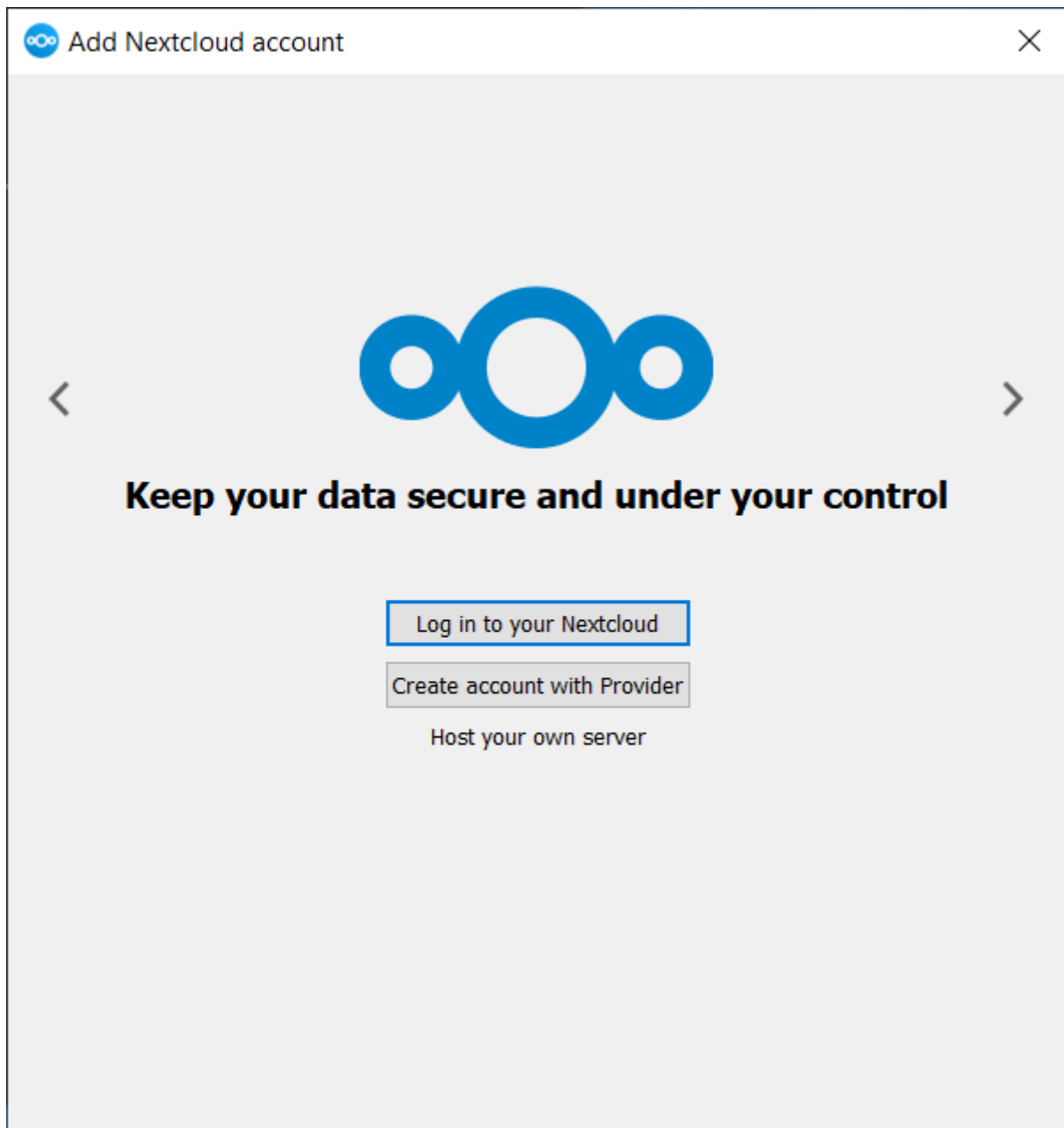
Cuando veas un círculo amarillo con el signo "!", ése es el icono informativo, así que debes hacer clic en él para ver qué tiene que decirte.



El círculo rojo con la "x" blanca indica un error de configuración, como un inicio de sesión o una URL de servidor incorrectos.


50.3 Configurar una cuenta


El asistente de instalación te guía paso a paso a través de las opciones de configuración y la configuración de la cuenta. En primer lugar, debes pulsar el botón "logueate em tú NextCloud".



Ahora deberás introducir la URL del servidor de NextCloud.

En nuestro caso la URL sería <https://nextcloud.iter.es> o <https://192.168.53.13>

 Add Nextcloud account ✕

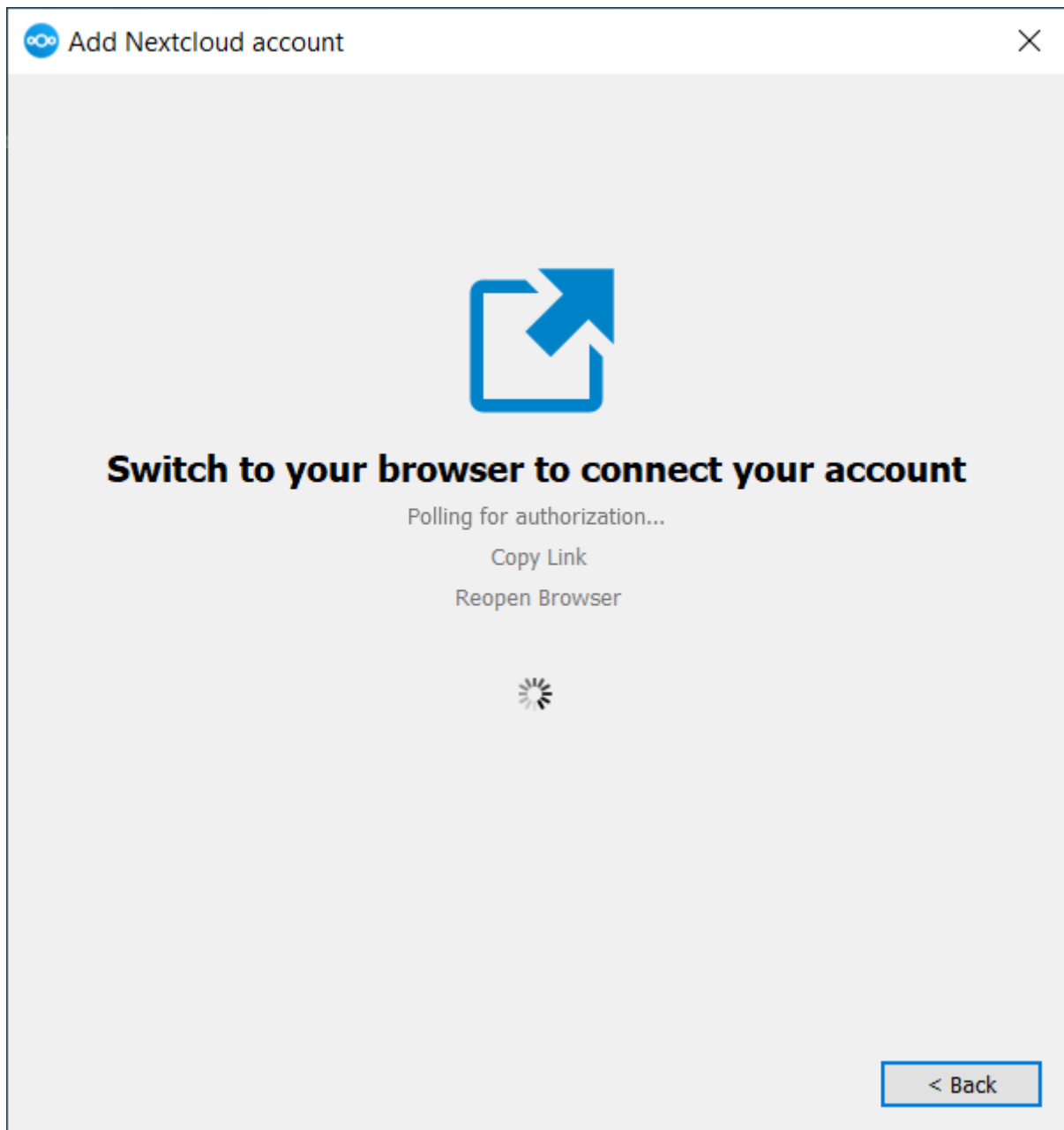


Server address

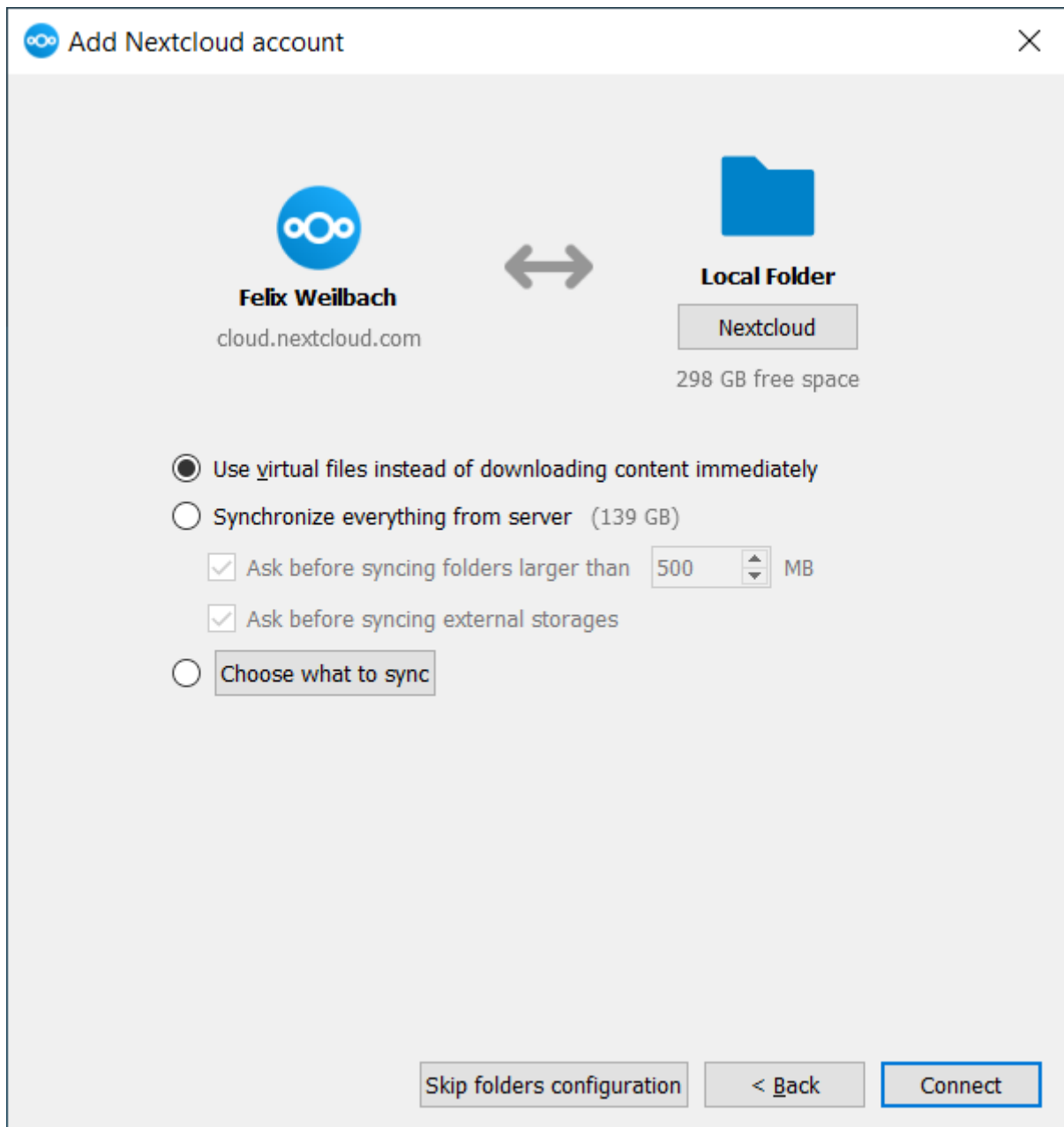
The link to your Nextcloud web interface when you open it in the browser.

< Back Next >

Ahora te aparecera la siguiente pantalla y te dirigirá al navegador para que allí inicies sesión con las mismas credenciales que utilizaría para iniciar sesión a través de la interfaz web.



Por último, elige la carpeta con la que el cliente de NextCloud debe sincronizar el contenido de tu cuenta de NextCloud.



Tras pulsar "Conectar", NextCloud Client comenzará con el proceso de sincronización.

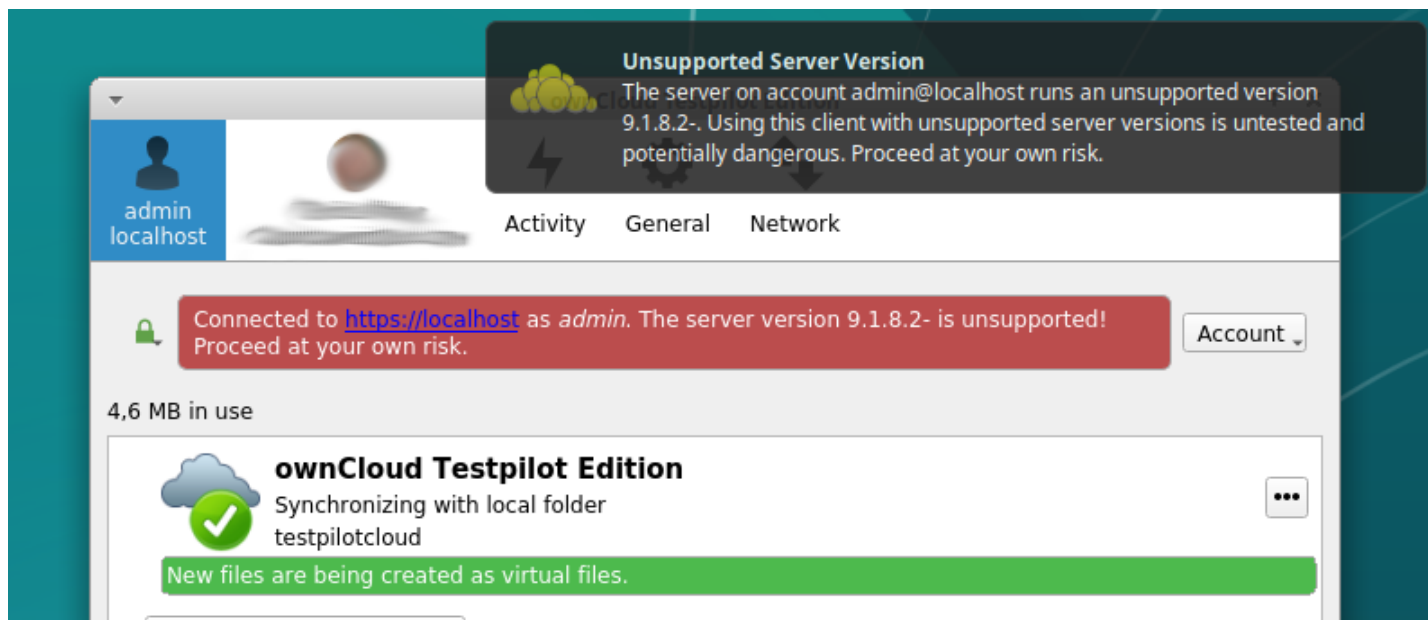
Si todo ha ido bien deberías ver el siguiente icono en tus iconos del sistema:



Para más información tiene a su disposición la [guía oficial](#)

51 Migración desde el cliente de escritorio de OwnCloud a NextCloud

Actualmente si seguís usando la aplicación de escritorio de OwnCloud os encontraréis con el siguiente mensaje: **!La versión del servidor de no es compatible!. Continúa bajo tu responsabilidad.**



No os preocupéis, podéis seguir trabajando sin problemas. Pero es recomendable hacer la migración hacía la aplicación de escritorio de NextCloud.

51.1 Pasos a seguir

51.1.1 1. Detener el cliente owncloud

La forma de detener el cliente ownCloud desde la aplicación de escritorio depende de la plataforma que esté utilizando.

51.1.1.1 Windows

Haga "click derecho" en el icono de ownCloud en la barra de tareas y seleccione "Cerrar" o "Salir" para detener el cliente.

51.1.1.2 MacOS

Haga clic en el icono ownCloud en la barra de menús superior y seleccione "Salir" para detener el cliente.

51.1.1.3 Linux

Haga clic en el icono ownCloud en la barra de tareas y seleccione "Salir" o "Cerrar" para detener el cliente.

51.1.2 2. Eliminar el cliente owncloud

La forma de eliminar el cliente de escritorio de ownCloud depende de la plataforma que esté utilizando.

51.1.2.1 Windows

Ir a "Agregar o quitar programas" en el Panel de control y buscar ownCloud, luego seleccione "Desinstalar" para eliminar el cliente.

51.1.2.2 MacOS

Arrastrar el icono de ownCloud desde la carpeta Aplicaciones a la papelera, luego vacíe la papelera para eliminar el cliente.

51.1.2.3 Linux

Utilice el administrador de paquetes de su distribución para buscar y eliminar ownCloud.

Por ejemplo, si está utilizando Ubuntu, puede abrir el terminal y ejecutar el siguiente comando:

```
$ sudo apt-get remove owncloud-client
```


51.1.3 3. Instalar el cliente nextcloud


Los pasos de indican en la siguiente URL:

51.1.4 4. Inicie el cliente nextcloud

Inicielo una única vez y ciérrelo justo cuando el asistente le pregunte por el servidor, etc.

Debe llegar hasta la siguiente ventana:

 Add Nextcloud account ✕


Server address

The link to your Nextcloud web interface
when you open it in the browser.

< Back Next >

51.1.5 5. Mueve tú fichero owncloud.cfg

Copie su fichero oculto owncloud.cfg localizado en el directorio principal de OwnCloud al directorio principal de NextCloud como nextcloud.cfg creado por el cliente en el paso 3.

Dependiendo de la plataforma la ruta por defecto será diferente.

51.1.5.1 Windows

En Microsoft Windows: %LOCALAPPDATA%\ownCloud\owncloud.cfg

51.1.5.2 MacOS

En MAC OS X: \$HOME/Library/Application Support/ownCloud

51.1.5.3 Linux

En distribuciones Linux: `$HOME/.local/share/data/ownCloud/owncloud.cfg` o `$HOME/.config/owncloud/owncloud.cfg`

51.1.6 6. Inicie el cliente nextcloud

Te pedirá la contraseña y después continuará donde terminó tu cliente owncloud.

Información

No se deberían transferir datos nuevos una vez hecho el procedimiento.

52 Copias de seguridad en Nextcloud

Nextcloud es una excelente solución para el almacenamiento y sincronización de archivos en la nube, **pero no está diseñado específicamente para ser utilizado como una solución de respaldo.**

Peligro

Se recomienda encarecidamente no usar Nextcloud para hacer copias de seguridad.

52.1 Razones

Existen varias razones por las cuales no se recomienda utilizar Nextcloud como backup:

52.1.1 No es una solución de backup completa

No es una solución de backup completa: Nextcloud es una solución de sincronización y almacenamiento de archivos, no una solución de backup completa. Una solución de backup completa debe incluir características como programación de copias de seguridad automáticas, verificación de integridad de datos y posibilidad de restaurar archivos a una fecha específica.

52.1.2 No se pueden hacer copias de seguridad incrementales

No se pueden hacer copias de seguridad incrementales: En Nextcloud, cada vez que se guarda un archivo, se sobrescribe la versión anterior del archivo. Esto significa que no se pueden hacer copias de seguridad incrementales, lo que es esencial para una solución de backup efectiva.

52.1.3 Riesgo de corrupción de datos


Riesgo de corrupción de datos: Si un archivo se corrompe o se elimina accidentalmente en Nextcloud, se eliminará en todas las instancias de la nube. Esto significa que si se produce un error en Nextcloud, se puede perder toda la información.

52.1.4 Limitaciones de almacenamiento

Limitaciones de almacenamiento: Nextcloud puede estar limitado en cuanto a la cantidad de almacenamiento que se puede utilizar para el almacenamiento de archivos. Si se utiliza para hacer copias de seguridad, puede ser difícil o costoso ampliar el almacenamiento necesario.

Información

Es importante asegurarse de que los datos estén bien respaldados para evitar la pérdida de información valiosa en caso de una falla en el sistema o una eliminación accidental.

 **Información**

Asegúrese de elegir una solución de backup que se adapte a sus necesidades y que cuente con características como **programación de copias de seguridad automáticas, verificación de integridad de datos y posibilidad de restaurar archivos a una fecha específica**. Además, siempre es recomendable realizar pruebas regulares de restauración para garantizar que los datos se puedan recuperar correctamente en caso de una emergencia.

Por estas razones, se recomienda utilizar una solución de copia de seguridad dedicada para proteger adecuadamente los datos importantes.

V.II OnDemand

53 Introducción a Open OnDemand

Open OnDemand es una plataforma revolucionaria basada en web diseñada para proporcionar a los usuarios un acceso fácil y seguro a recursos de cómputo de alto rendimiento (HPC). Esta interfaz amigable simplifica el proceso de usar sistemas HPC complejos al permitir a los usuarios realizar una amplia gama de tareas a través de un navegador web estándar, eliminando la necesidad de interacciones tradicionales de línea de comandos. Open OnDemand democratiza el acceso a recursos informáticos potentes, haciéndolos accesibles para una gama más amplia de disciplinas y niveles de experiencia.



OnDemand provides an integrated, single access point for all of your HPC resources.

powered by
OPEN OnDemand

OnDemand version: 3.0.3

53.1 Características Clave

Open OnDemand ofrece una variedad de características diseñadas para mejorar la experiencia del usuario y mejorar el acceso a los recursos TeideHPC:

- **Acceso Basado en Web:** Los usuarios pueden acceder a los recursos TeideHPC desde cualquier navegador web estándar, sin necesidad de VPNs o software cliente especializado.
- **Aplicaciones Interactivas:** Open OnDemand permite a los usuarios lanzar aplicaciones GUI interactivas (como Jupyter Notebooks, RStudio y Matlab) directamente en su navegador, facilitando la realización de cálculos y análisis complejos.
- **Gestión de Archivos:** Un navegador de archivos integrado permite a los usuarios cargar, descargar, editar y gestionar fácilmente sus archivos y directorios en el sistema TeideHPC.
- **Gestión de Trabajos:** Los usuarios pueden enviar, monitorear y gestionar trabajos por lotes directamente desde la

interfaz web, con soporte para scripts de trabajo y una variedad de sistemas de envío de trabajos.

- **Acceso a Shell:** Para los usuarios avanzados que prefieren interfaces de línea de comandos, Open OnDemand proporciona acceso integrado a terminales shell, permitiendo la interacción directa con el entorno Linux subyacente del sistema TeideHPC.

53.2 Ventajas de Open OnDemand

La principal ventaja de Open OnDemand es su capacidad para hacer los recursos TeideHPC más accesibles y fáciles de usar. Esta accesibilidad fomenta la innovación y colaboración en una amplia gama de disciplinas científicas, de ingeniería y análisis de datos. Otras ventajas incluyen:

- **Curva de Aprendizaje Reducida:** Al proporcionar una interfaz gráfica para tareas tradicionalmente realizadas a través de la línea de comandos, Open OnDemand reduce significativamente la barrera de entrada para nuevos usuarios de TeideHPC.
- **Productividad Mejorada:** El flujo de trabajo simplificado que ofrece Open OnDemand permite a los investigadores concentrarse más en su investigación y menos en dominar las complejidades de los sistemas HPC.
- **Colaboración y Compartición:** Open OnDemand facilita la colaboración entre investigadores al hacer más fácil compartir aplicaciones, datos y resultados.

53.3 Casos de Uso

Open OnDemand es versátil y puede soportar una amplia gama de casos de uso, incluyendo, pero no limitado a:

- **Análisis y Visualización de Datos:** Los investigadores pueden realizar análisis de datos utilizando sus herramientas preferidas y visualizar resultados directamente en su navegador web.
- **Simulación y Modelado:** Ingenieros y científicos pueden ejecutar simulaciones y modelos en recursos computacionales potentes, ajustando parámetros y analizando resultados en tiempo real.
- **Fines Educativos:** Los educadores pueden usar Open OnDemand para proporcionar a los estudiantes experiencia práctica con recursos HPC, habilitando aprendizaje práctico en cursos que requieren un poder computacional significativo.

53.4 Cómo Comenzar

Para comenzar con Open OnDemand, los usuarios típicamente necesitan:

1. Tener una **cuenta en TeideHPC** desde la cual tendrán acceso a Open OnDemand.
2. Navegar al portal de Open OnDemand en <https://ondemand.hpc.iter.es/> usando un navegador web.
3. Autenticarse usando sus credenciales de cuenta TeideHPC.

A partir de ahí, los usuarios pueden explorar las aplicaciones disponibles, gestionar archivos, enviar trabajos y más, todo a través de una interfaz web intuitiva.

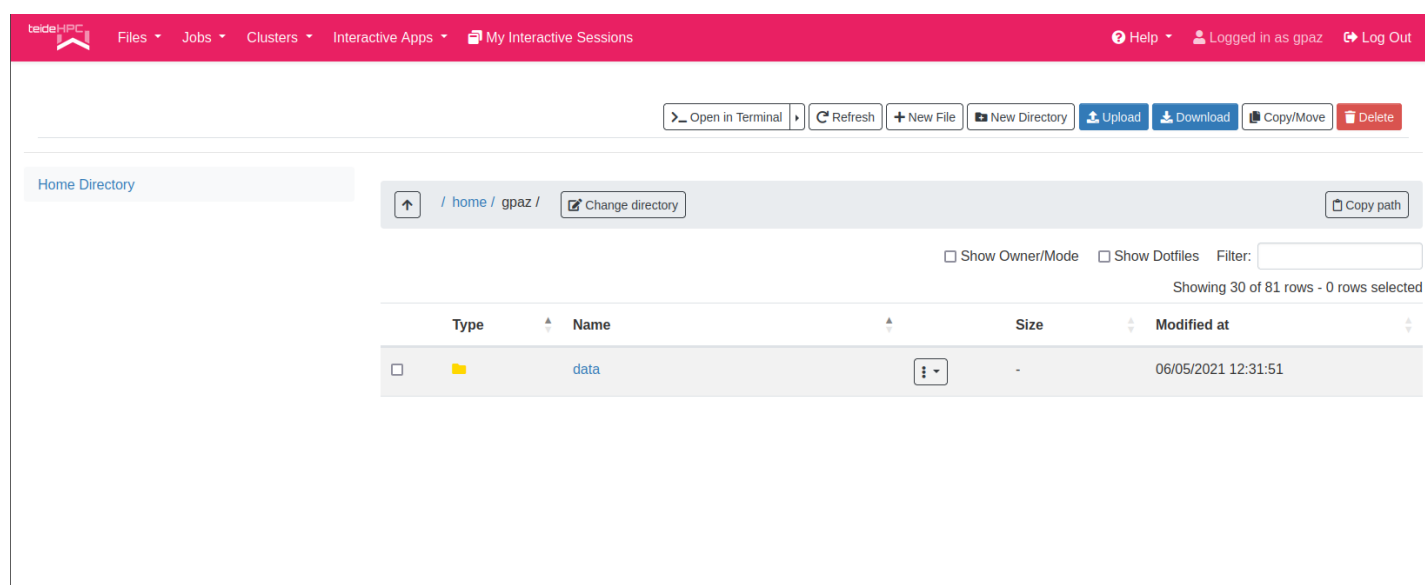
54 Trabajando con Archivos en OnDemand

Antes de trabajar en el cluster, es importante entender dónde deben almacenarse los diferentes tipos de archivos en la jerarquía de carpetas. Por lo tanto, recomendamos encarecidamente revisar nuestra página de [almacenamiento de Datos](#) para obtener esta información.

54.1 Acceso a la página de Archivos en OnDemand

1. Conéctate a OnDemand en <https://ondemand.hpc.iter.es/> con tus credenciales.
2. Haz clic en **Archivos**.
3. Selecciona **Directorio principal**.

Deberías ver lo siguiente:



The screenshot displays the OnDemand file manager interface. At the top, there is a navigation bar with the TeideHPC logo and menu items: Files, Jobs, Clusters, Interactive Apps, and My Interactive Sessions. On the right side of the navigation bar, there are links for Help, Logged in as gpaz, and Log Out. Below the navigation bar, there is a toolbar with buttons for Open in Terminal, Refresh, New File, New Directory, Upload, Download, Copy/Move, and Delete. The main content area shows the 'Home Directory' view. It includes a breadcrumb navigation path: / home / gpaz /, with a 'Change directory' button and a 'Copy path' button. There are also checkboxes for 'Show Owner/Mode' and 'Show Dotfiles', and a 'Filter' input field. Below this, a table displays the contents of the directory. The table has columns for Type, Name, Size, and Modified at. The table shows 30 of 81 rows, with 0 rows selected. A single file named 'data' is visible, with a size of '-' and a modified date of '06/05/2021 12:31:51'.

54.2 Crear una carpeta

1. Conéctate a OnDemand en <https://ondemand.hpc.iter.es/> con tus credenciales.
2. Haz clic en **Archivos**.
3. Selecciona **Directorio principal**
4. Haz clic en **Nuevo Directorio**.
5. Ingresa el nombre del directorio.
6. Haz clic en **OK**.

¡El directorio ha sido creado!. Puedes navegar a él seleccionándolo en la interfaz de archivos, y ahora puedes añadir archivos allí.

54.3 Subir un archivo

1. Conéctate a OnDemand en <https://ondemand.hpc.iter.es/> con tus credenciales.
2. Haz clic en **Archivos**.
3. Selecciona **Directorio principal**
4. Haz clic en **Subir**.
5. Selecciona **Examinar archivos** para subir un archivo.
6. Navega al archivo que te gustaría subir. Ten en cuenta que puedes seleccionar varios.
7. Haz clic en **Subir archivo**.

Tu archivo ha sido subido al directorio que especificaste en Adroit. Puedes ver el archivo haciendo clic en su nombre de archivo en la interfaz de Archivos.

NOTA:

Cuando subes archivos a través del portal OnDemand, cada subida está limitada a un tamaño de 2 GB. Si necesitas subir archivos más grandes, recomendamos usar comandos de Linux, como se describe en nuestra página de [transferencia de datos](#). Puedes introducir estos comandos de Linux dentro del Terminal, el cual puedes abrir como se describe en nuestra sección de [ejecutar un terminal](#) desde OnDemand más adelante en esta página. Contacta a support@hpc.iter.es si necesitas ayuda con esto.

54.4 Crear un archivo

1. Conéctate a OnDemand en <https://ondemand.hpc.iter.es/> con tus credenciales.
2. Haz clic en **Archivos**.
3. Selecciona **Directorio principal**
4. Haz clic en **Nuevo Archivo**.
5. Ingresa el nuevo nombre del archivo.
6. Haz clic en **OK**.

¡Tu archivo ha sido creado!

54.5 Editar un archivo

1. Conéctate a OnDemand en <https://ondemand.hpc.iter.es/> con tus credenciales.
2. Haz clic en **Archivos**.
3. Selecciona **Directorio principal**
4. Navega a tu archivo.
5. Selecciona la opción del menú con 3 puntos verticales y una flecha hacia abajo al lado de tu archivo.

6. Selecciona **Editar** del menú desplegable que aparece.

Ahora estás en la interfaz de edición para tu archivo. Rellénalo con cualquier contenido que necesites. Una vez que hayas terminado, haz clic en **Guardar** en la esquina superior izquierda de la interfaz del Editor.

54.6 Renombrar un archivo

1. Conéctate a OnDemand en <https://ondemand.hpc.iter.es/> con tus credenciales.
2. Haz clic en **Archivos**.
3. Selecciona **Directorio principal**
4. Navega a tu archivo.
5. Selecciona la opción del menú con 3 puntos verticales y una flecha hacia abajo al lado de tu archivo.
6. Selecciona **Renombrar** del menú desplegable que aparece.
7. Ingresa tu nuevo nombre de archivo.
8. Haz clic en **OK**.

¡Tu archivo ha sido renombrado!

54.7 Copiar/Mover un archivo

1. Conéctate a OnDemand en <https://ondemand.hpc.iter.es/> con tus credenciales.
2. Haz clic en **Archivos**.
3. Selecciona **Directorio principal**
4. Navega a tu archivo.
5. Selecciona la casilla a la izquierda del(los) archivo(s) que te gustaría copiar/mover.
6. Haz clic en el botón **Copiar/Mover** en la parte superior derecha de la interfaz de Archivos.
7. Navega al directorio al que te gustaría copiar/mover tu(s) archivo(s).
8. Haz clic en el botón **Copiar** o **Mover** en el diálogo en la parte izquierda de tu pantalla dependiendo de la operación que desees realizar.

¡Tu(s) archivo(s) han sido copiados/movidos!

54.8 Eliminar un archivo

1. Conéctate a OnDemand en <https://ondemand.hpc.iter.es/> con tus credenciales.
2. Haz clic en **Archivos**.
3. Selecciona **Directorio principal**
4. Navega a tu archivo.
5. Selecciona la opción del menú con 3 puntos verticales y una flecha hacia abajo al lado de tu archivo.

6. Selecciona **Eliminar** del menú desplegable que aparece. O:
7. Selecciona la casilla a la izquierda del(los) archivo(s) que te gustaría eliminar.
8. Haz clic en **Eliminar** en la esquina superior derecha de la interfaz de Archivos.

¡Tu(s) archivo(s) han sido eliminados!

55 Ejecución de trabajos en OnDemand

Para ejecutar trabajos utilizando la interfaz OnDemand, debe crear un **script slurm** para su código. Para obtener más información sobre cómo enviar trabajos a los **clústeres**, le recomendamos que consulte nuestra Guía de **mi primer trabajo con Slurm**.

55.1 Acceder a la Pestaña de Trabajos

Una vez que hayas iniciado sesión en OnDemand en <https://ondemand.hpc.iter.es/>, sigue estos pasos para acceder a la pestaña de trabajos:

1. En la barra de navegación superior, encuentra y haz clic en el menú "Trabajos".
2. Verás varias opciones, incluyendo "Trabajos activos", "Trabajos completados", y "Enviar trabajo". Selecciona la opción que desees para proceder.

55.2 Enviar un Nuevo Trabajo

Para enviar un nuevo trabajo a través de la interfaz de OOD:

1. Dentro de la pestaña de "Trabajos", selecciona "Enviar trabajo".
2. Serás dirigido a una página donde puedes seleccionar la aplicación o el entorno de ejecución para tu trabajo. Elige la que se ajuste a tus necesidades.
3. Completa el formulario con los detalles de tu trabajo, incluyendo recursos requeridos como número de núcleos, memoria y tiempo de ejecución.
4. Opcional: Adjunta archivos o especifica comandos de ejecución según sea necesario.
5. Haz clic en "Enviar" para encolar tu trabajo. Recibirás una confirmación con el ID de tu trabajo.

55.3 Monitorear y Gestionar Trabajos Activos

Para ver y gestionar trabajos que has enviado:

1. Ve a "Trabajos activos" desde la pestaña de "Trabajos".
2. Aquí, puedes ver una lista de tus trabajos en ejecución o en cola, junto con su estado actual.
3. Para acciones específicas como cancelar un trabajo, selecciona el trabajo deseado y utiliza las opciones disponibles.

Deberías ver lo siguiente:

The screenshot shows the 'Active Jobs' section of the TeideHPC dashboard. The top navigation bar includes 'Files', 'Jobs', 'Clusters', 'Interactive Apps', and 'My Interactive Sessions'. The user is logged in as 'gpaz'. The page displays a table with one job entry:

ID	Name	User	Account	Time Used	Queue	Status	Cluster	Actions
10883	sys/dashboard/sys/bc_teidehpc_jupyter_GPU	gpaz	hpc	00:06:39	gpu	Running	Anaga	[Delete]

Below the table, it indicates 'Showing 1 to 1 of 1 entries' and provides navigation buttons for 'Previous', '1', and 'Next'.

Además puedes consultar todo los detalles de los trabajos desplegándolo con la pestaña de la izquierda. Podrás ver información como el id del trabajo, el nº de nodos usados, el estado, ...:

This screenshot shows the same 'Active Jobs' page, but with the details for job 10883 expanded in a modal window. The job is running on the 'Anaga' cluster. The details include:

- Cluster:** Anaga
- Job Id:** 10883
- Job Name:** sys/dashboard/sys/bc_teidehpc_jupyter_GPU
- User:** gpaz
- Account:** hpc
- Partition:** gpu
- State:** RUNNING
- Reason:** None
- Total Nodes:** 2
- Node List:** node17101-1, node17105-1
- Total CPUs:** 50
- Time Limit:** 1-00:00:00
- Time Used:** 2:06:48
- Start Time:** 2024-03-21 10:32:13
- End Time:** 2024-03-22 10:32:13
- Memory:** 76800M
- GRES:** gpu:a100:1

The 'Output Location' is shown as `/home/gpaz/ondemand/data/sys/dashboard/batch_connect/sys/bc_teidehpc_jupyter_GPU/output/0a69954c-5afd-4d07-80df-6003e2a9e208`. At the bottom, there are buttons for 'Open in File Manager', 'Open in Terminal', and 'Delete'.

55.4 Consultar Trabajos Completados

Para revisar trabajos que han finalizado:

1. Selecciona "Trabajos completados" desde la pestaña de "Trabajos".
2. Podrás ver una lista de trabajos con información sobre su ejecución y resultados.
3. Utiliza esta sección para descargar resultados o analizar el rendimiento de tus trabajos.

56 Clústeres en TeideHPC

En TeideHPC disponemos de dos clústeres: el clúster TeideHPC y el clúster Anaga.

La pestaña "Clusters" en **OnDemand** ofrece a los usuarios una interfaz gráfica para interactuar con los recursos de los clusters de supercomputación disponibles. A través de esta interfaz puedes acceder a terminales de los nodos de login de cada clúster.

56.1 Cómo Acceder a la Pestaña "Clusters"

1. Conéctate a OnDemand en <https://ondemand.hpc.iter.es/> usando un navegador web.
2. En la barra de navegación principal, encuentra y haz clic en la pestaña "Clusters".
3. Verás una lista de los clusters disponibles a los cuales tienes acceso.

Al acceder deberías ver lo siguiente:



```
Host: 10.5.22.100
teideHPC
Last login: Wed Mar 20 08:22:00 2024 from 10.5.22.18
(base) [gpaz@login-cpu01 ~]$
```

57 Aplicaciones Interactivas en OnDemand

57.1 Acceso a las Aplicaciones Interactivas

Para comenzar a usar las aplicaciones interactivas:

1. Conéctate a OnDemand en <https://ondemand.hpc.iter.es/> con tus credenciales.
2. Dirígete a la pestaña "Interactive Apps" en la barra de navegación principal.
3. Aquí encontrarás un listado de las aplicaciones disponibles. Selecciona la aplicación deseada: Code Server, JupyterLab, o RStudio, y elige entre la versión para CPU o para GPU según tus necesidades.

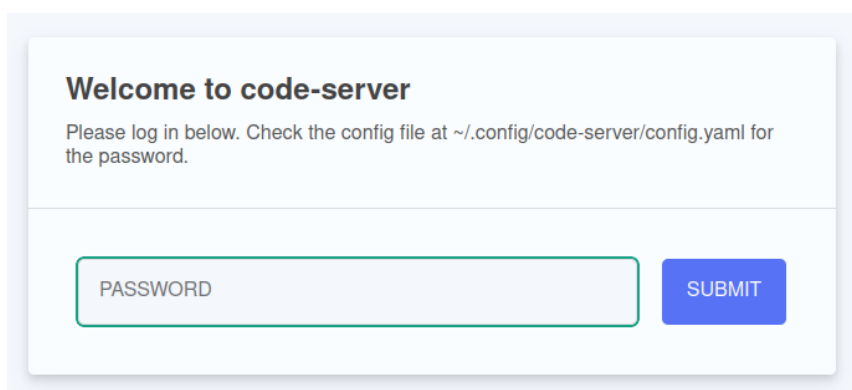
57.2 Aplicaciones Disponibles

57.2.1 Code Server

Code Server proporciona una versión de Visual Studio Code que se ejecuta en el servidor y es accesible a través del navegador. Ideal para desarrollo de software y edición de código.

- **Para CPU/GPU:** Selecciona la versión que necesites según el tipo de procesamiento requerido para tu proyecto.
- **Configuración:** Especifica los recursos necesarios (como número de cores, memoria y tiempo de ejecución). Para la versión GPU, asegúrate de seleccionar también el tipo de GPU deseado.
- **Uso:** Una vez lanzada, la aplicación abrirá Visual Studio Code en tu navegador, permitiéndote trabajar con tus archivos y proyectos almacenados en el servidor.

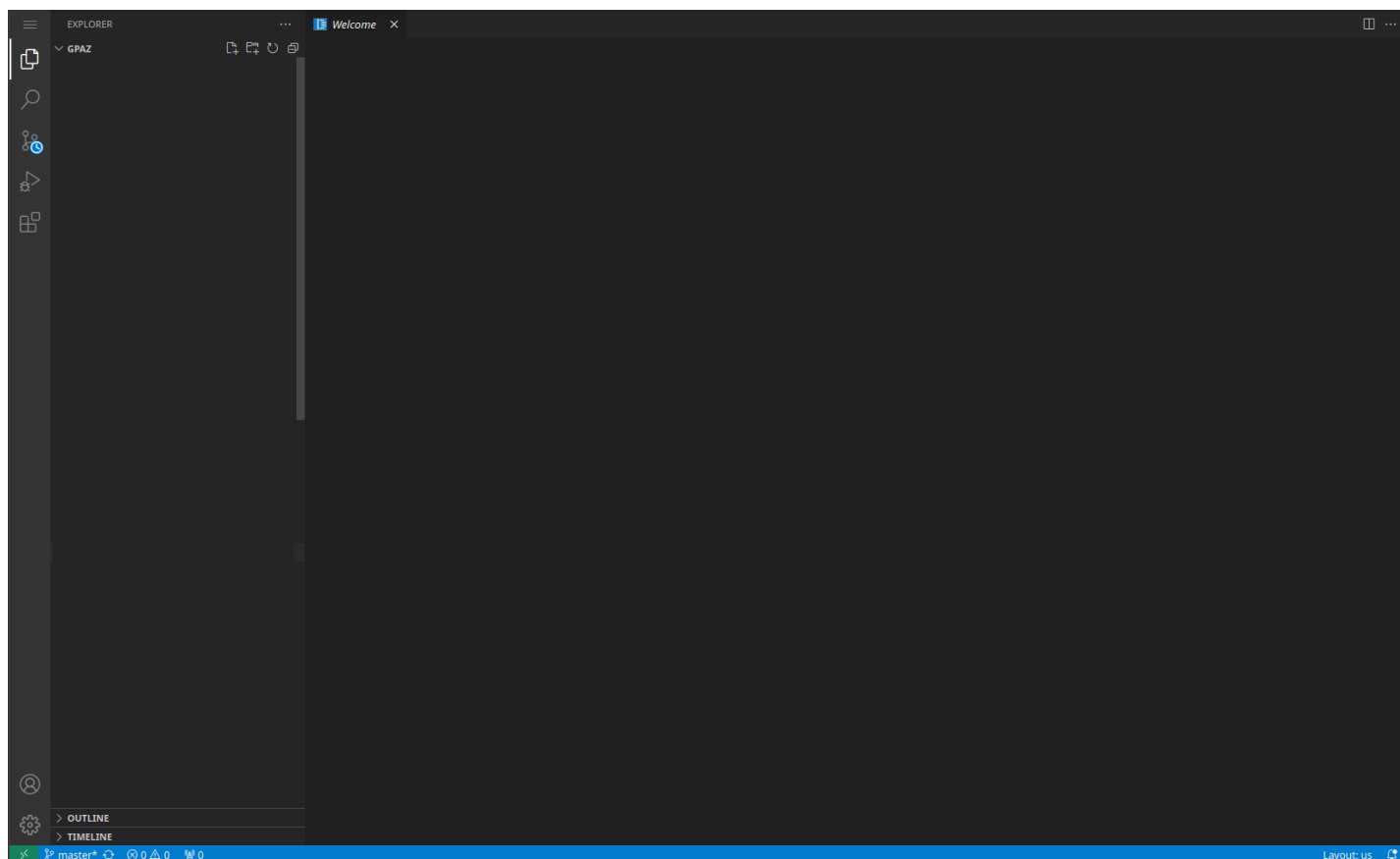
La aplicación te pedirá una contraseña para poder acceder:



Welcome to code-server

Please log in below. Check the config file at `~/.config/code-server/config.yaml` for the password.

Una vez dentro tendrás el editor de código referencia del marco actual:

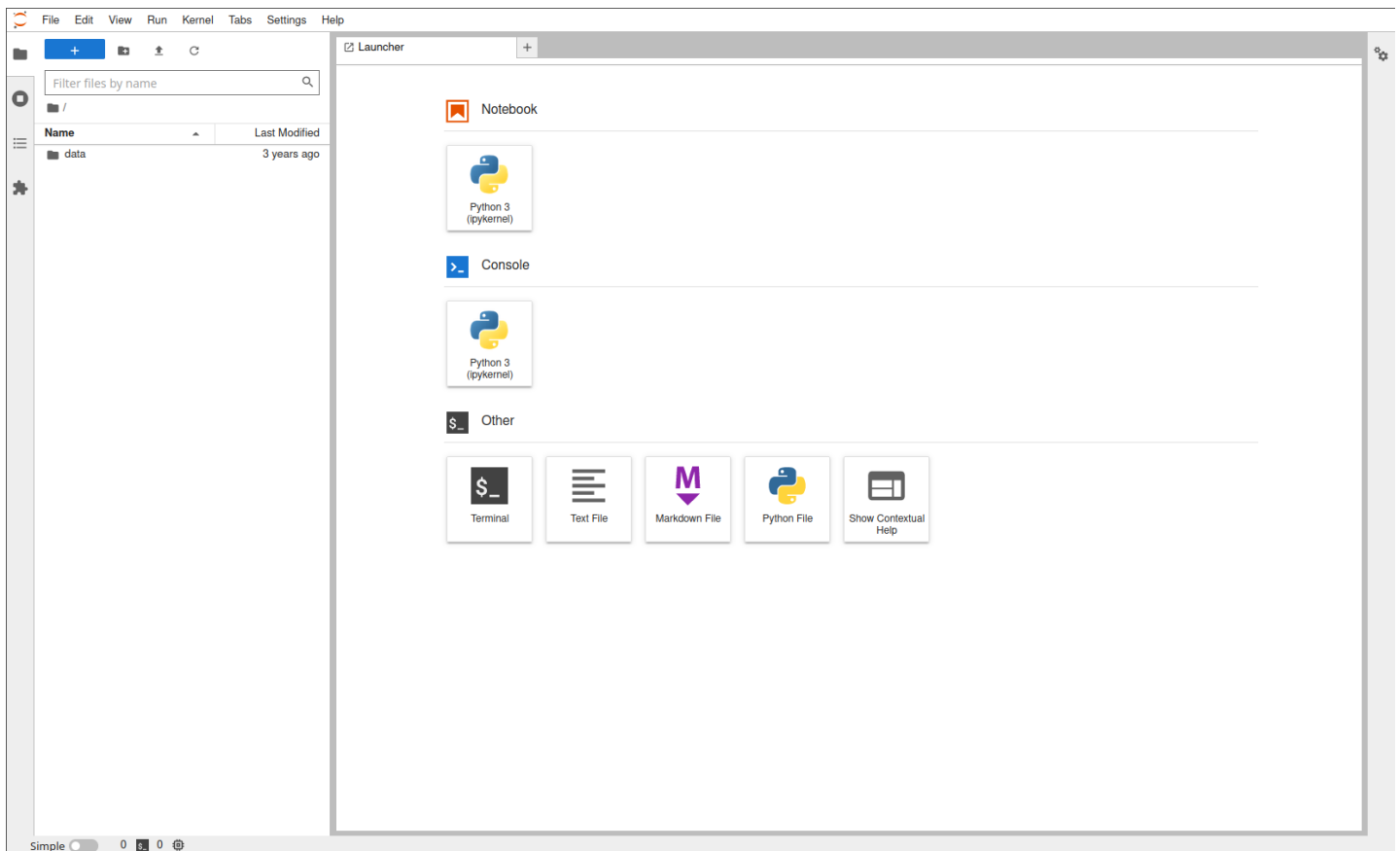


57.2.2 JupyterLab

JupyterLab ofrece un entorno interactivo de ciencia de datos que soporta lenguajes de programación como Python, R, y Julia.

- **Para CPU/GPU:** Elige según las demandas computacionales de tus notebooks de Jupyter.
- **Configuración:** Indica los recursos computacionales necesarios. Para la versión GPU, selecciona el tipo de GPU que prefieras.
- **Uso:** Al lanzar JupyterLab, se abrirá una nueva ventana en tu navegador para trabajar con notebooks, código, y datos.

Una vez dentro verás lo siguiente:

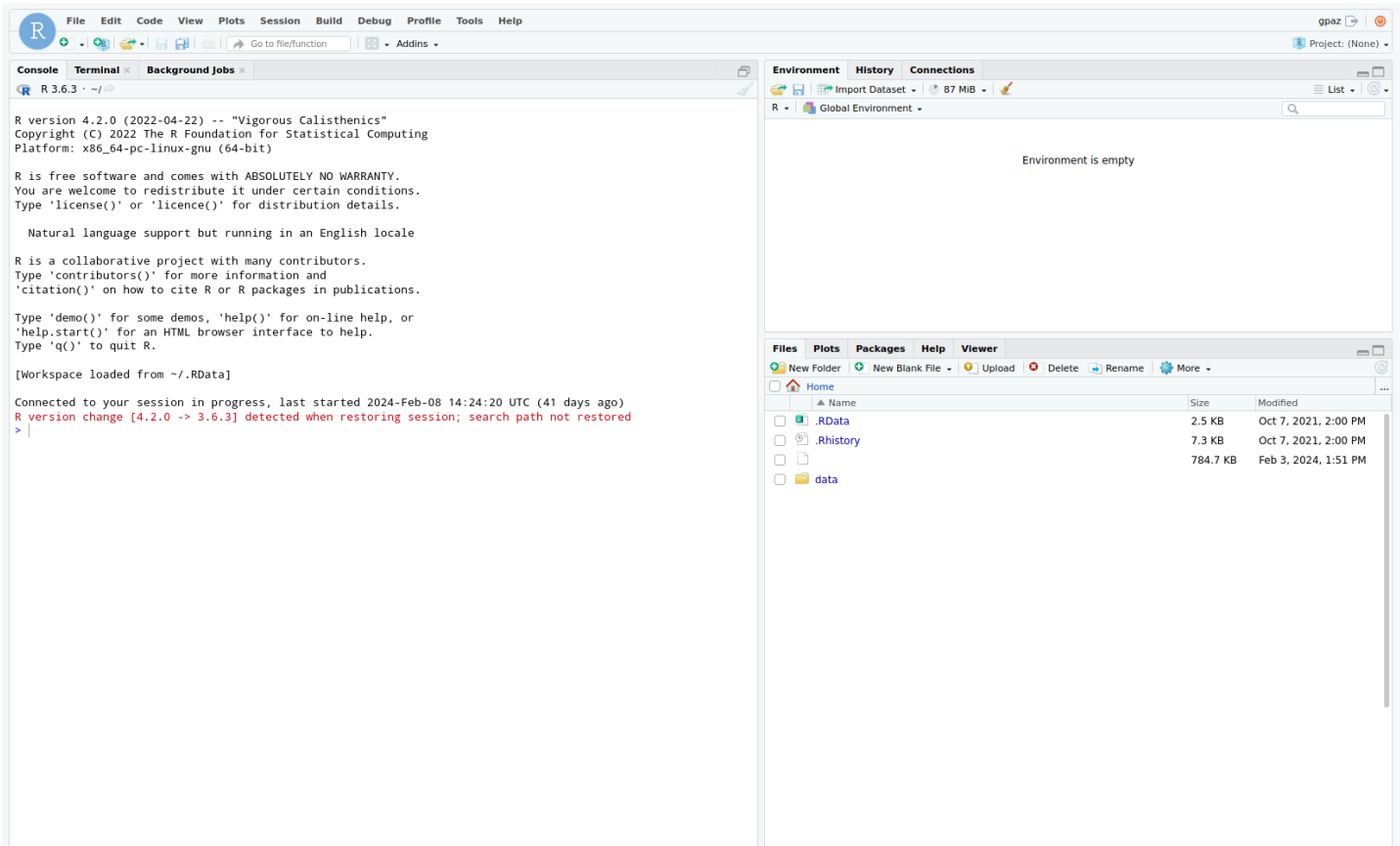


57.2.3 RStudio

RStudio brinda un entorno de desarrollo integrado para R, facilitando la programación, visualización de datos, y más.

- **Para CPU/GPU:** Dependiendo del procesamiento necesario para tus análisis de datos, elige la versión adecuada.
- **Configuración:** Asigna los recursos que requieres, incluyendo el tipo de GPU si optas por esa versión.
- **Uso:** RStudio se abrirá en el navegador, proporcionando acceso a las herramientas de desarrollo de R en un entorno familiar.

Una vez dentro verás lo siguiente:



57.3 Consejos Generales

- **Guarda tu Trabajo:** Asegúrate de guardar tus archivos y proyectos regularmente para no perder progreso.
- **Gestiona tus Recursos:** Cierra las aplicaciones interactivas cuando no las estés utilizando para liberar recursos en el cluster.
- **Consulta la Documentación:** Cada aplicación tiene características y configuraciones específicas. Consulta la documentación oficial para aprovecharlas al máximo.

57.4 Soporte

Si encuentras problemas o tienes preguntas acerca de las aplicaciones interactivas en Open OnDemand, no dudes en ponerte en contacto con nosotros a través del correo electrónico support@hpc.iter.es. Estamos aquí para ayudarte.

58 Preguntas frecuentes.

Si tienes alguna preguntanos en support@hpc.iter.es

58.1 HPC

58.2 Opennebula

Error VNC -> VNC Server disconnected (code: 1015)