

Guía de conexión:

sobre cómo trabajar de forma remota en el COI

Redacción: Amélie Bernès Periodo: Primavera, 2023

Resumen

Para trabajar en el desarrollo de modelos de IA en el COI, tienes acceso a una computadora con GPU's (que hacen que el entrenamiento de redes complejas con un gran número de datos sea posible), pero probablemente no podrás manipularla directamente. En su lugar, trabajarás con un dispositivo personal, desde el que te podrás conectar al servidor del COI de forma remota. En este escrito esbozamos unos posibles pasos a seguir para poder conectarte de forma remota a una cuenta de usuario a la que tengas permiso acceder, así como instrucciones sobre acciones básicas (como correr programas, mover archivos de tu dispositivo local al remoto y viceversa) que seguramente deberás realizar. Las palabras resaltadas en azul son links a los respectivos archivos.

Escribí estas notas para ayudarme a recordar los pasos a seguir y para aprender del proceso, pero decidí compartirlas por si son de utilidad para futuros trabajadores. Estoy segura de que estas notas contienen varios errores; si encuentras algunos, siéntete libre de enviarme tus comentarios a mi correo ammel.bernes@gmail.com.

Índice

1.	Cuenta de usuario en servidor remoto y terminal para interactuar con ella	Ι
2.	Accediendo a una cuenta de usuario remota	II
3.	Instalación de anaconda y creación de entornos virtuales 3.1. Usando sftp para intercambiar información entre servidores 3.2. Instalando Anaconda en el usuario remoto	V
4.	tmux 4.1. Qué es tmux y cómo instalarlo	
5.	Notas teóricas	IX
6.	Referencias	XII

1. Cuenta de usuario en servidor remoto y terminal para interactuar con ella

Necesitarás contar con una cuenta de usuario en el servidor del COI y una contraseña para acceder a ella. Un administrador del servidor puede crearla por ti. Para que no tengas problemas de permisos, deberán asegurarse de darte permisos de administrador.

Nota 1. Si tu dispositivo físico usa a Windows como sistema operativo, necesitarás además instalar una terminal para poder establecer contacto con el servidor del COI (cuyo sistema operativo es Linux Mint). Te sugerimos instalar MobaXterm (sólo tienes que descargar el archivo ejecutable gratis y seguir todos los pasos de instalación). Si por el contrario tienes alguna distribución de Linux, podrás establecer contacto con tu cuenta del servidor remoto a partir de una instancia de terminal cualquiera.

En lo que sigue supondremos que querremos trabajar en la cuenta de usuario creada en el servidor remoto del COI llamada **amelie**_ y que en nuestra computadora local tenemos Windows y MobaXterm.

2. Accediendo a una cuenta de usuario remota

Es usando ssh que podremos, via el comando

ssh amelie_@ratagamer.duckdns.org

escrito desde una terminal, establecer el contacto entre el dispositivo físico y el usuario amelie_.

Se te pedirá que ingreses la contraseña de la cuenta en la que pretendes entrar (en este caso, amelie_).

Nota 2. El nombre ratagamer.duckdns.org se implementó para usarlo en lugar de la dirección IP del servidor del COI, pues las direcciones IP públicas (las que usas) cambian constantemente. Más detalles.

Puedes saber que, en tu terminal, ahora estás en la cuenta remota por el cambio de @:

aquí figura

Una vez iniciada la conexión, la instancia local de terminal es en realidad una terminal abierta en tu usuario remoto; podrás realizar toda clase de acciones (por ejemplo, instalación de aplicaciones, actualización del sistema) desde tu terminal.

Te recomiendo que actualices tu cuenta remota escribiendo en la consola (una vez iniciada la conexión con ssh) la instrucción

sudo apt update sudo apt ugrade -y

Seguro necesitarás instalar aplicaciones en tu cuenta remota para trabajar en el desarrollo de tus modelos, por ejemplo, un bloc de notas o un editor de imágenes. Yo decidí instalar vim y nomacs

Cheat Sheet 1.

sudo apt install nombre: comando para instalar la aplicación llamada "nombre".

```
backup darknet examples libdarknet.so LICENSE LICENSE.gen LICENSE.meta LICENSE.weta LICENSE.weta
```

```
| The state of the
```

Figura 1: En la primera captura de pantalla, me encuentro conectada de forma remota en mi usuario amelie_ y estoy en la carpeta llamada "darknet". Puesto que necesito abrir y editar un archivo de texto, decido instalar vim. Para ello, busco primero a vim con el comando sudo apt search vim. Una vez ubicado en la lista, lo instalo con el comando sudo apt install vim. Como puedes ver por los mensajes que siguen, no parece haber ningún problema con la instalación.

```
| Case |
```

Figura 2: Una vez instalado vim, lo uso para abrir el archivo llamado "Makefile" (que, como confirmo con una aplicación de ls, está en la carpeta en la que me encuentro actualmente); hago esto escribiendo vim Makefile. Como puedes ver en la segunda captura de pantalla, en la instancia de terminal vim abre el archivo de texto, y puedo ser capaz de editarlo.

Nota 3. Yo elegí usar vim porque es un visualizador y editor de texto que puede usarse sólo con teclado. Aprender a usar vim puede ser frustrante y tardado, por lo que si no estás familiarizado con él pero necesitas un editor de texto, te recomiendo que instales en su lugar a gedit.

3. Instalación de anaconda y creación de entornos virtuales

3.1. Usando sftp para intercambiar información entre servidores

Puesto que requerimos usar sftp para instalar Anaconda en nuestro entorno virtual y, en general, para enviar o tomar datos (que estén ya sea en formato de imágenes, como archivos de texto, etcétera) desde nuestro dispositivo local al virtual y viceversa, explicamos cómo hacer esto en esta subsección.

Terminar.

3.2. Instalando Anaconda en el usuario remoto

Anaconda es un software usado para programar en Python; nos resultará útil pues nos permitirá crear entornos virtuales en los que podremos instalar distintas versiones de Python y otros módulos que necesites (esto es crucial si, para ejecutar cierto pedazo de código, se requiere una versión de Python distinta a la que usas por default, o si quieres usar (no simultaneamente) dos módulos de Python que interfieren uno con otro). Te recomendamos pues instalar anaconda en tu usuario del COI para poder crear entornos virtuales para cada uno de tus proyectos. Si quisieras instalar Anaconda en tu dispositivo local, lo único que tendrías que hacer es descargar el archivo ejecutable en la página oficial de Anaconda y ejecutarlo; sin embargo, ¡quieres Anaconda en tu cuenta remota! Antes de ejecutar el archivo .exe debes pues de enviarlo al usuario amelie_ usando sftp. cuál es la sintaxis exacta?

Ahora que tienes el ejecutable de Anaconda en tu usuario remoto, puedes instalarlo con el comando

bash *filename*.sh

Puesto que durante la instalación de Anaconda se instala también pip, podrás usar este gestor de paquetes de Python para instalar todas las librerías/módulos que requieras para la construcción de tus modelos.

3.3. Creando entornos virtuales

Puedes crear un entorno virtual (que llamaremos **electro**) y especificar la versión de Python que en él se cree con la siguiente orden:

conda create -n electro Python=3.9

Se ha creado ya el entorno virtual, pero para acceder a él necesitas ejecutar la orden

conda activate electro

Podrás confirmar que estás dentro de determinado entorno porque el nombre de este aparecerá entre paréntesis al inicio de las prompt en tu terminal.

Cheat Sheet 2.

pip install nombre: para instalar el módulo llamado "nombre" en el entorno de Anaconda activado actualmente.

pip list: para ver una lista con todos los módulos de Python instalados en el entorno de Anaconda activado actualmente

pip list — grep nombre: para buscar específicamente el módulo llamado "nombre" en el entorno. conda deactivate: para salir del entorno de Anaconda en el que te encuentres actualmente.

Si planeas inicializar en este entorno los **jupyter notebooks** en los que escribas los códigos de un determinado proyecto, debes instalar en este entorno (con pip) todos los módulos que ese proyecto requiera (por ejemplo, JupyterLab, notebook, pandas, numpy, tensorflow, keras), cuidando, obviamente, que sean compatibles unos con otros.

```
amelie_@ratagamer.duckdns.org's password:
Last login: Fri Jan 20 11:45:32 2023 from 189.190.210.91
(base) amelie_@COI:~$ pip list |grep tensorflow
(base) amelie_@COI:~$ pip list |grep tensorflow
tensorflow
celectro) amelie_@COI:~$ pip list |grep tensorflow
tensorflow-estimator 2.11.0
tensorflow-estimator 2.11.0
tensorflow-io-gcs-filesystem 0.29.0
(electro) amelie_@COI:~$ pip list |grep scipy
(electro) amelie_@COI:~$ pip list |grep scipy
Collecting scipy

Downloading scipy-1.10.0-cp39-cp39-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (34.4 MB)

Requirement already satisfied: numpy<1.27.0, $1.19.5 in ./anaconda3/envs/electro/lib/python3.9/site-packages (from scipy) (1.24.1)
Installing collected packages: scipy

Successfully installed scipy-1.10.0
(electro) amelie_@COI:~$ pip list |grep scipy
1.10.0
(electro) amelie_@COI:~$ pip list |grep scipy
```

Figura 3: En la primera linea, con ssh establecemos conexión con amelie. Observa que, al inicio de cada linea, aparece entre paréntesis la palabra "base"; este es el entorno de Anacando activado por default; observa que el módulo tensorflow no está instalado. Para entrar al entorno "electro" (creado previamente), ejecutamos conda activate electro; como puedes ver, ahí sí está instalado tensorflow, pero no scipy; lo instalamos pues con la orden pip install scipy

4. tmux

Para poder iniciar procesos que no se detengan aún cuando tu dispositivo personal se apague o desconecte del servidor del COI, puedes crearlos en sesiones de tmux.

4.1. Qué es tmux y cómo instalarlo

Asegúrate de haber establecido conexión con tu usuario remoto via ssh. imagen que muestre la diferencia entre estar en tu local y tu remoto. Para instalar tmux debes ejecutar la instrucción

sudo apt install tmux

en una terminal en la que estés conectando con tu usuario en el servidor del COI.

Para iniciar una sesión en tmux, tienes que elegir un nombre para ella (pongamos **nuevaSesion**); ejecuta la siguiente instrucción:

tmux new -s nuevaSesion

Nota 4. Si en su lugar hubíesemos ejecutado

tmux new-session

entonces se habría iniciado una nueva sesión cuyo nombre es un número elegido por default. Lo mejor es que tú le des un nombre distintivo a tu sesión de tmux.

Nota 5. Recuerda que quieres ejecutar tus jupyter notebook en entornos virtual. Piensa que ya creaste uno y que ya tienes tmux instalado. Debes de activar tal entorno virtual antes de crear la

sesión de tmux en la que vas a inicializar tu jupyter notebook. Una vez que lo crees, siempre que entres a esa sesión de tmux, lo harás desde el entorno virtual especificado durante su creación.

A menos que ejecutes una instrucción para cerrar esta sesión de tmux o que el dispositivo del COI se apague (cosa que, en teoría, no debería ocurrir), esta y todo lo que en ella hayas ejecutado estará siempre corriendo.

Cheat Sheet 3.

tmux ls: Para desplegar una lista con todas las sesiones de tmux que tengas iniciadas.

tmux attach -t nombre: Para anclar la sesión de tmux llamada "nombre". tmux kill-session -t nombre: Para eliminar la sesión de tmux llamada "nombre".

Nota 6. ¿Por qué no correr los programas en tu dispositivo físico, sino en tu usuario remoto? Dos razones:

- tal vez tu dispositivo físico no tenga GPU's. La computadora del COI sí tiene. Algunas desventajas de, por ejemplo, extraer datos del remoto pero ejecutar código en el físico son
- If something breaks the connection between your workstation and the remote server you have to re-establish the connection.
- If you have full access to the remote server you may be able to run your entire workload there, and any analysis you start on the remote server will continue to run even if your local workstation gets interrupted.

4.2. Ejecutando procesos en sesiones de tmux y accediendo a estos localmente

Primero, debes asegurarte de conectarte a tu usuario remoto con ssh. Creemos una sesión de tmux, y llamémosla jupyterAmelie. Creo que en mi ejemplo me faltó inicializar mi entorno anaconda!!. Al ejecutar esta orden se crea la sesión de tmux; en ella, escribimos jupyter notebook para iniciar ahí un cuaderno de jupyter. Ejecutada la instrucción aparecerá un texto como el que sigue:

```
| Columbia Geoli- | Columbia G
```

Figura 4: En la primera linea, con ssh establecemos conexión con amelie. Observa que, al inicio de cada linea, aparece entre paréntesis la palabra "base"; este es el entorno de Anacando activado por default; observa que el módulo tensorflow no está instalado. Para entrar al entorno "electro" (creado previamente), ejecutamos conda activate electro; como puedes ver, ahí sí está instalado tensorflow, pero no scipy; lo instalamos pues con la orden pip install scipy

Nota que se marca que el puerto 8888 ya está ocupado; casi siempre un proceso de jupyter notebook buscará iniciarse en el puerto 8888 (c.f. [1]), pero bien puede ocurrir (como aquí) que ese puerto ya esté ocupado con otro proceso. En ese caso, se busca un puerto más arriba que esté libre. Como puedes apreciar en la imagen ??, el puerto remoto en el que se ha iniciado jupyter es el número 8889. Elijamos un puerto local (pongamos, el 9900) para establecer la conexión con el proceso que corre en el 8889 remoto. Escribimos para esto

ssh -L 9900:localhost:8889 amelie_@ratagamer.duckdns.org

Si todo salió bien, cuando escribas

localhost:9900

en tu navegador de preferencia, podrás abrir el jupyter notebook que corre en el puerto 8889 de tu cuenta remota en una instancia del navegador que abriste localmente. A veces, para terminar este

paso, se te requerirá introducir el token del proceso.

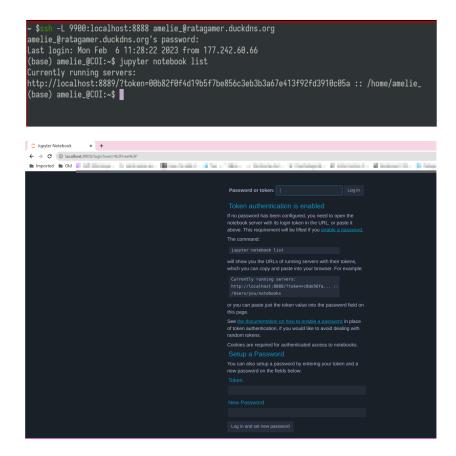


Figura 5: A veces tendrás que introducir el token del proceso. Como puedes ver en la imagen, escribiendo jupyter notebook list puedes ver el token.

Ahora, no importa si pierdes la conexión o si apagas tu equipo local; siempre y cuando el equipo del COI permanezca prendido, tus sesiones de tmux y los procesos que en ellos se ejecuten seguirán corriendo. Si el servidor del COI se apaga accidentalmente, perderás todas tus sesiones de tmux, por lo que tendrás que realizar de nuevo todos los pasos de esta sección.

5. Notas teóricas

Sobre tmux

• el nombre tmux es una abreviación de "terminal multiplexer"; es común usar tmux para abrir más de un panel (ventana) a partir de una sola terminal. Cada ventana será independiente

una de otra. Esto es particularmente útil cuando entras en contacto con otro dispositivo con ssh, pues podrás tener abiertos múltiples paneles al mismo tiempo, each with their own shell running, but using the same, single SSH connection.

- All tmux commands consist of a prefix followed by a command. The most common prefix you'll use with tmux is Ctrl+b. In all cases, you can press Ctrl+b instead of typing tmux within a tmux pane.
- Sessions are basic to tmux. Sessions can consist of multiple windows, any of which may have multiple panes. The active window is called a notion.

Creando una sesión de tmux

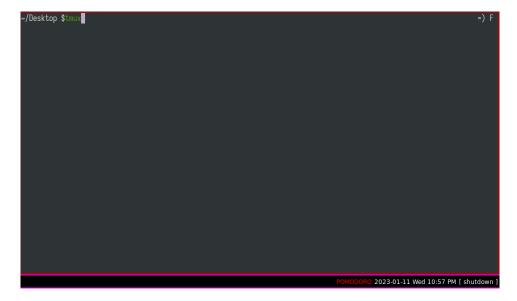


Figura 6



Figura 7

Cambiando el nombre de esa sesión



Figura 8

Lo que pasa si en esa sesión de tmux ejecutas exit

```
~/Desktop $tmux
[exited]
~/Desktop $tmux ls
no server running on /tmp/tmux-1000/default
~/Desktop $█
```

Figura 9

Lo que pasa si, en su lugar, ejecutas tmux detach

```
~/Desktop $tmux
[detached (from session sesionAme)]
~/Desktop $tmux Is
sesionAme: 1 windows (created Wed Jan 11 23:18:34 2023)
~/Desktop $
```

Figura 10

Matando instancias (?) de tmux

```
-/Desktop $tmux ls
sesionAme: 1 windows (created Wed Jan 11 23:18:34 2023)
-/Desktop $tmux kill-session -t sesionAme
-/Desktop $tmux ls
no server running on /tmp/tmux-1000/default
-/Desktop $
```

Figura 11

Bibliografía

https://towardsdatascience.com/setting-up-and-using-jupyter-notebooks-on-aws-61a9648db6c5

6. Referencias

Fuente sobre tmux

sobre dotfiles

Remote Pair Programming Made Easy with SSH and tmux