

Lancer un Jupyter Notebook sur PlaFRIM

Objectif

Dans le cadre d'AI4INDUSTRY et principalement pour les sujets orientés IA, nous souhaitons permettre aux étudiants de lancer du code sur PlaFRIM (voir section suivante) afin de traiter des jeux de données (plus ou moins importants).

Pour ce faire, ils pourront utiliser le langage Python ainsi que des bibliothèques spécifiques.

Nous leur suggérons aussi, d'utiliser Jupyter Notebook pour pouvoir exécuter leur code et visualiser leurs résultats facilement.

INFOS IMPORTANTES: - Si vous êtes sur Windows, vous pouvez télécharger WSL qui permet d'installer linux sur Windows. Ainsi vous pourrez copier les lignes de code suivantes dedans. Si vous avez linux sur votre ordinateur, copiez seulement les lignes dans un terminal.

- Les étapes de l'installation sont résumées dans la partie: **Résumé des étapes d'installation.**

Présentation de PlaFRIM

PlaFRIM est une plateforme scientifique conçue pour soutenir la recherche expérimentale dans tous les domaines des mathématiques appliquées liés à la modélisation et au calcul haute performance.

Les étudiants exécuteront leur script Python (resp. cellules de leur Jupyter Notebook) sur cet outil, ce qui leur permettra par exemple d'avoir accès à des GPUs.

Comment utiliser PlaFRIM

Vous allez avoir besoin d'ouvrir deux terminaux. Nous les nommerons pour la suite: terminal 1 et terminal 2.

Connexion à PlaFRIM (dans terminal 1)

Pour vous connecter à PlaFRIM, ouvrez un premier terminal et lancez le script bash `script_terminal1.sh`, pour cela copiez cette commande dans le terminal 1 :

```
bash script_terminal1.sh
```

Une fois la commande lancée, vous êtes désormais connecté à votre compte PlaFRIM.

Ce script permet de vous authentifier par mot de passe à PlaFRIM formation puis effectue une seconde authentification par clé avec la paire de clé qui se trouve sur la machine intermédiaire présente uniquement sur la plateforme PlaFRIM formation.

Demande d'utilisation de GPUs (dans terminal 1)

Toujours dans ce terminal, lancez le script présent dans votre répertoire PlaFRIM avec la commande suivante:

```
bash script_terminal1_plafrim_part2.sh
```

Par défaut, ce script permet de choisir une machine (parmi les sirocco) dont l'état est *idle*, par exemple ici la machine sirocco06, d'obtenir une allocation de noeuds pour 4 heures d'utilisation (voir paramètre: `--time=240`) et de s'y connecter.

Ne fermez pas ce terminal!

Création d'un tunnel ssh, activation de l'environnement virtuel et ouverture d'un jupyter notebook (dans terminal 2)

Ouvrez maintenant un second terminal et lancez le dernier script: `script_terminal2.sh` en précisant un port. Dans cet exemple, nous prenons "8888".

```
bash script_terminal2.sh 8888
```

Ce script effectue un ajout de la clé distante à son agent pour 120s le temps de créer un tunnel puis crée un tunnel ssh. C'est à dire: nous allons faire correspondre un port local de notre machine à un port distant de la machine sur laquelle on se connecte (la machine sirocco06 de PlaFRIM), le tout au travers de notre connexion SSH. Voir principe du tunneling :

<https://www.it-connect.fr/chapitres/tunneling-ssh/> (<https://www.it-connect.fr/chapitres/tunneling-ssh/>).

Pour finir, restez dans ce second terminal et lancez le dernier script en spécifiant à nouveau le même port choisi:

```
bash script_terminal2_plafrim_part2.sh 8888
```

Ce script active un environnement conda avec Python 3.7 et les bibliothèques suivantes:

- torch 2.1.1
- ipykernel 6.27.1
- tensorflow 2.15.0
- pandas 2.1.4
- sklearn 1.3.2
- jupyter 1.0.0

Il lance ensuite un Jupyter notebook sur PlaFRIM. Vous pouvez désormais utiliser des GPUS présents sur la plateforme PlaFRIM (via la bibliothèque torch).

ATTENTION: Vous pouvez vous connecter à plusieurs sur un compte PlaFRIM mais n'oubliez pas de choisir un port différent par utilisateur. Ex: 7777, 8888, etc.

Résumé des étapes d'installation

1. ETAPE 1 Dans le terminal 1:

```
bash script_terminal1.sh
```

2. ETAPE 2 Puis dans le terminal 1:

```
bash script_terminal1_plafrim_part2.sh
```

Ne fermez pas ce terminal.

3. ETAPE 3 Ouvrez un second terminal. Lancez le script en précisant un port unique par utilisateur d'un même compte. Il y a un compte par groupe d'une vingtaine d'étudiants, chacun des étudiants doit choisir un port différent.

```
bash script_terminal2.sh 8888
```

4. ETAPE 4

Toujours dans le second terminal. Lancez le dernier script en précisant le même port que précédemment:

```
bash script_terminal2_plafrim_part2.sh 8888
```

Vous pouvez désormais utiliser jupyter notebook sur PlaFRIM dans votre navigateur.

Informations complémentaires

Pour vérifier si des GPUs sont accessibles, tapez dans votre jupyter notebook:

```
import torch
print(torch.cuda.is_available())
```

Cette cellule devrait renvoyer *True* dans ce cas-là. *False* sinon. Pour plus d'informations sur le fonctionnement de Jupyter Notebook, [cliquez ici \(https://docs.jupyter.org/en/latest/index.html\)](https://docs.jupyter.org/en/latest/index.html).

Problèmes éventuellement rencontrés

Débordement quota d'espace disque

Si vous avez un problème de débordement de quota d'espace disque. Effectuez les lignes suivantes:

```
pip cache purge  
conda clean all
```

Suppression environnement virtuel

Vous pouvez aussi supprimer votre environnement pour en créer un avec une version différente de Python :

```
conda env remove --name ai4industry
```

Copier les images du pc local vers remote

Dans un terminal effectuez:

```
scp -p -r <file> <user>@plafrim.fr:/home/<user>
```

Quota espace disque

Par défaut, un compte PlaFRIM a une capacité de 10 Giga, vous pouvez nous envoyer un mail pour augmenter cette capacité.