Formation Reproductibilité en science de données

Programme de la formation

- 1. Introduction
- 2. Environnement de travail : Installation des logiciels & tests
- 3. Le versioning des codes et documents avec Git
 - 1. Git: commandes de base
 - 2. Git: La gestion des conflits
- 4. Les containers Docker
 - 1. La technologie des containers
 - 2. Docker: commandes de base
- 5. R studio
- 6. Jupyter notebook & lab
 - 1. Jupyter pour Bash
 - 2. Jupyter pour R
 - 3. Jupyter pour python
 - 4. le framework spark en python (pyspark)
- 7. Le deploiement dans le cloud avec Openstack
 - 1. Pourquoi le cloud
 - 2. Horizon, l'interface d'openstack
 - 3. Déploiement
 - 1. configuration et lancement de VM
 - 2. Acces en SSH
 - 3. Installation de Docker sur VM
 - 4. Deploiement Jupyter et interface
- 8. cluster et reproductibilité
 - 1. Job scheduler avec Slurm
 - 2. Container avec Singularity

Liste des documents associés

1. Introduction

diaporama: doc/f1-1_introduction.pdf

plan : :doc/f1-plan.pdf

- 2. Environnement de travail : Installation des logiciels & tests procedure d'installation : doc/f2_installation.pdf
- 3. Le versioning des codes et documents avec Git
 - 1. Git: commandes de base

TP: git_tuto/git_intro.pdf

2. Git: La gestion des conflits TP: git_tuto/git_conflit.pdf

4. Les containers Docker

cours: docker-intro.pdf

- 1. La technologie des containers
- 2. Docker : commandes de base
- 5. R studio

TP: dossier R-studio

- 6. Jupyter notebook & lab
 - 1. Jupyter pour Bash

TP: dossier jupyter-bash

2. Jupyter pour R

TP: dossier jupyter-R

3. Jupyter pour python

TP: dossier jupyter-python

4. le framework spark en python (pyspark)

TP: dossier jupyter-spark

7. Le deploiement dans le cloud avec Openstack

cours et TP : dossier slurm-openstack

- 1. Pourquoi le cloud
- 2. Horizon, l'interface d'openstack
- 3. Déploiement
 - 1. configuration et lancement de VM
 - 2. Acces en SSH
 - 3. Installation de Docker sur VM
 - 4. Deploiement Jupyter et interface
- 8. cluster et reproductibilité

dossier slurm-singularity

- 1. Job scheduler avec Slurm
- 2. Container avec Singularity