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

1. Environnement de travail : Installation des logiciels & tests

procedure d’installation : doc/f2\_installation.pdf

1. Le versioning des codes et documents avec Git
   1. Git: commandes de base

TP : git\_tuto/git\_intro.pdf

* 1. Git: La gestion des conflits

TP : git\_tuto/git\_conflit.pdf

1. Les containers Docker

cours : docker-intro.pdf

* 1. La technologie des containers
  2. Docker : commandes de base

1. R studio

TP : dossier R-studio

1. Jupyter notebook & lab
   1. Jupyter pour Bash

TP : dossier jupyter-bash

* 1. Jupyter pour R

TP: dossier jupyter-R

* 1. Jupyter pour python

TP : dossier jupyter-python

* 1. le framework spark en python (pyspark)

TP : dossier jupyter-spark

1. 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

1. cluster et reproductibilité

dossier slurm-singularity

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