Université Paris 8 U.F.R MITSIC – Département Informatique

Master 2 – Big Data et Fouille de Données Dockérisation pour le Big Data 2020-2021



Travail à réaliser

Le but de cet exercice est de mettre en pratique l'ensemble des enseignements Docker vues pendant les séances de cours et TPs pour réaliser un mini projet sur l'architecture micro-service - Docker.

Merci de respecter les consignes suivantes, pour un travail bien accompli :

- L'ensemble de fichiers de votre projet devraient être compressés dans un seul fichier zip sous la forme TAR_Noms_Prénoms.zip, et ensuite envoyé avec comme objet : [Projet_Docker_BDFD] à l'adresse mail : otman.manad02@univ-paris8.fr;
- Le travail est effectué soit en trinôme, binôme ou individuellement;
- Le nombre de pages du rapport ne doit pas dépasser **5 pages** ;
- Le travail devrait être envoyé avant **12/01/2021 à 23:59** (dernier délai, aucun projet ne sera accepté après ce délai).
- **NB**: la note finale est calculée sur la logique de réponse à chaque exercice + son exécution, l'état de fonctionnement de l'ensemble des exercices et le rapport.

Exercice 1:

Créer un service qui contient un container nommé « algo_container » dans le but est de faire exécuter un algorithme de machine/deep learning (supervisé/non supervisé) de votre choix sur un dataset. Vous pouvez réutiliser les algorithmes réalisés pendant vos cours (notamment avec M. Mariage, ...), avec un langage de programmation de votre choix (de préférence python).*

Exercice 2:

Créer un service contenant un container nommé «ui_container » son but est de récupérer les résultats du «algo_container» et les afficher dans une interface web avec le langage/serveur web de votre choix (HTML, JavaScript, Flask, ...).*

Exercice 3:

Une fois l'exercice 1 et 2 réalisés, configurer l'environnement de développement virtuel **Vagrant** dans le but d'inclure les deux containers. Le but étant d'isoler l'environnement de développement, et d'automatiser les compilations.

Exercice 4 (bonus): créer un nouveau répertoire nommé « **bonus** », contenant le fichier docker-compose.yml dont les services sont de type **images** (pas de build) des deux conteneurs crées en Exercice 1 et 2 (vous pouvez utiliser docker-hub, Github ou autre pour stocker vos images).

Notes pour la rédaction du rapport: Expliquez, avec des schémas/captures/commandes, le fonctionnement de chaque container, ainsi, la technique de machine learning adoptée (apprentissage profond, non supervisé ou supervisé, ...), autrement dit, expliquez l'intérêt de chaque container : comment fonctionne-t-il ? quelle est la structure des données d'entrée/sortie ? l'objectif de l'apprentissage ? ...

Otman MANAD Dockérisation pour le Big Data

^{*} le nombre de conteneurs n'est pas limité à deux, vous pouvez avoir 2 ou plus.