
Examen : Session de Contrôle

Traitement du Big data avancé

Filière : Mastère Pro. Data Science

Niveau : 1 ère année

Enseignant: M. SAKKARI

A.U : 2022-2023

Durée : 01H30

Documents : non autorisés

Exercice 1 :(9 pts)

Une table HBase est un tableau multidimensionnel de données distribué et persisté sur le HDFS sous forme de fichiers spécifiques appelés HFiles.

Question 1: Donner les dimensions qui forment une table HBase ? (4 pts)

Question 2: Quelle est la structure (l'architecture de base) d'un cluster HBase ? Donner un schéma explicatif (5 pts)

Exercice 2 : (11 pts)

Question 1 : Comment fonctionne HDFS ? (2 pt)

Question 2 : Qu'est-ce que le mode "high availability" (2 pts)

Question 3 : Qu'est-ce que Hadoop YARN (2 pt)

Question 4 : Donner (avec explications) l'architecture Apache Hadoop YARN? (5 pt)

Bon travail