

1. Hadoop est un framework qui fonctionne avec une variété d'outils connexes. Les cohortes communes incluent :

- A. MapReduce, Hive and HBase**
- B. MapReduce, MySQL and Google Apps
- C. MapReduce, Hummer and Iguana
- D. MapReduce, Heron and Trumpet

2. Quel type de données Hadoop peut traiter

- A. Données structurées
- B. Données semi-structurées
- C. Données non structurées
- D. Toutes les réponses**

3. L'unique changement de Hadoop v1 à Hadoop v2 était la séparation de la gestion des ressources et des traitements des jobs. Vrai ou faux ?

- A. Vrai
- B. Faux**

4. Lequel de ces composants n'existe pas dans la distribution cloudera 4.7

- A. java
- B. pig
- C. spark**
- D. hue

5. Comment fonctionne la distribution de fichiers sur HDFS ?

- A. Répartition en fonction de la taille des fichiers sur chaque nœud du cluster.
- B. Répartition en blocs répliqués sur les nœuds du cluster.**
- C. Répartition en nœuds répliqués sur les blocs du cluster.
- D. Répartition en fonction des choix de l'utilisateur au moment de l'upload.

6. Le besoin de réplication de données peut se produire dans divers scénarios comme:

- A. DataNode ne fonctionne plus
- B. Les blocs de données sont corrompus
- C. Les deux réponses précédentes**
- D. Aucune réponse

7. Job Tracker fonctionne sur le :

- A. NameNode**
- B. DataNode
- C. NameNode secondaire
- D. DataNode secondaire

8. Comment fonctionne la distribution de fichiers sur HDFS ?

- A. Répartition en fonction de la taille des fichiers sur chaque nœuds du cluster.
- B. Répartition en blocs répliqués sur les nœuds du cluster.**
- C. Répartition en nœuds répliqués sur les blocs du cluster.
- D. Répartition en fonction des choix de l'utilisateur au moment de l'upload.

9. Le modèle de programmation MapReduce a été initialement développé par...

- A. Apache Software Foundation
- B. Google**
- C. Microsoft Research
- D. Aucune réponse

10. L'entrée du _____ est la sortie triée des Mappers.

- A. Reducer**
- B. Mapper
- C. Shuffle
- D. Aucune des réponses

11. Lequel des fichiers contient le paramètre de configuration pour NodeManager et ResourceManager

- A. yarn-site.xml**
- B. hdfs-site.xml
- C. mapred-site.xml
- D. Aucune de ces réponses

12. Quelles sont les améliorations du YARN par rapport au paradigme mapreduce v1 (MPR1)?

- A. Il divise JobTracker en deux processus : ResourceManager et ApplicationManager
- B. Il divise le TaskTracker en deux processus : ResourceManager et ApplicationManager

13. Chaque base de données créée dans Hive est stockée sous :

- A. Un fichier
- B. Un répertoire
- C. Un bloc HDFS
- D. Un fichier jar

14. Si la structure de la table ne correspond pas aux types de données présents dans le fichier contenant la table alors Hive

- A. Supprime automatiquement le fichier
- B. Corrige automatiquement les données
- C. Envoyer des valeurs nulles pour les données incompatibles
- D. N'autorise aucune requête à s'exécuter sur la table

15. Les données d'une table Hive sont stockées dans:

- A. Un script SQL
- B. Un répertoire HDFS
- C. Un job Mapreduce
- D. Aucune de ces réponses

16. Pig était initialement développé

- A. Yahoo!
- B. Facebook
- C. Apache
- D. Aucune de ces réponses

17. Les colonnes dans HBase sont organisées

- A. Groupe de colonnes
- B. Familles de colonnes
- C. Liste des colonnes
- D. Base de colonne

18. HBase est une base de données non relationnelle calquée sur le _____ de Google

- A. BigTop
- B. Bigtable**
- C. FoundationDB
- D. Aucune de ces réponses

19. La quelle parmi ces commandes est opérationnelles dans Hbase ?

- A. GET
- B. Put
- C. Delete
- D. Tous les réponses**