



Big Data intégration SQL : Hive, SparkDataframes

VOUS ÊTES

Experts en bases de données, chefs de projet et toute personne souhaitant comprendre le fonctionnement et les apports des bases NoSQL.

VOS OBJECTIFS

Connaître les caractéristiques techniques des bases de données NoSQL, les différentes solutions disponibles.

Identifier les critères de choix.

VOTRE FORMATION



DURÉE : 2 jours
14 heures



PROCHAINES SESSIONS :
14 et 15 mai 2020
29 et 30 juin 2020



LIEU : En distanciel



PRIX : 1 470 €
net de taxes

PRÉ-REQUIS

Connaissance générale des systèmes d'informations et des bases de données.

MODALITÉS

La formation est accessible à distance en classe virtuelle : accès à l'infrastructure de travaux pratiques, machines physiques, outils pédagogiques, échanges avec le formateur

De 4 à 12 participants

Financement éligible au FNE Formation pour tout salarié d'entreprise en activité partielle

VOTRE PROGRAMME

Présentation

Besoin. Comment concilier le quoi et le comment?

Faciliter la manipulation de gros volumes de données en conservant une approche utilisateurs.

Rappels sur le stockage : HDFS, Cassandra, HBase et les formats de données : parquet, orc, raw, clés/valeurs

Les outils : Hive, Impala, Tez, Presto, Drill, Pig, Spark/QL

Hive et Pig

Présentation. Mode de fonctionnement. Rappel sur map/reduce.

Hive : le langage HiveQL. Exemples.

Pig : le langage pig/latin. Exemples.

Impala

Présentation. Cadre d'utilisation. Contraintes. Liaison avec le métastore Hive.

TP : mise en évidence des performances.

Presto

Cadre d'utilisation. Sources de données utilisables.

TP : mise en oeuvre d'une requête s'appuyant sur Cassandra, HDFS et PostgreSQL

Spark DataFrame

Les différentes approches. Syntaxe Spark/SQL. APIs QL.

Compilation catalyst. Syntaxe, opérateurs. Comparaison avec Presto.

TP : mise en oeuvre d'une requête s'appuyant sur HBase et HDFS.

Drill/sqoop

Utilisation d'APIs JDBC, ODBC. Indépendance Hadoop. Contraintes d'utilisation.

Performances.

Comparatifs

Compatibilité ANSI/SQL. Approches des différents produits.

Critères de choix.

VOTRE CONTACT :



Andrea FALLOURD

Conseillère en formation

06 74 51 44 97

afalourd@itescia.fr

ITESCIA - Campus de Pontoise

8 rue Pierre de Coubertin

95300 PONTOISE

www.itescia.fr



Rejoignez nos réseaux sociaux

