

Référence	4-PP-BIG
Durée	2 jours (14 heures)
Éligible CPF	NON
Mise à jour	27/11/2023

BigData, architecture et technologies



OBJECTIFS PÉDAGOGIQUES

- Comprendre les concepts du BigData et connaître les technologies associées.



PUBLIC CONCERNÉ

Chefs de projets, architectes, et toute personne souhaitant connaître les outils et solutions pour mettre en place une architecture BigData.



PRÉREQUIS

Bonne culture générale sur les systèmes d'information.



MOYENS PÉDAGOGIQUES

- Réflexion de groupe et apports théoriques du formateur
- Travail d'échange avec les participants sous forme de réunion-discussion
- Utilisation de cas concrets issus de l'expérience professionnelle
- Validation des acquis par des questionnaires, des tests d'évaluation, des mises en situation et des jeux pédagogiques.
- Remise d'un support de cours.



MODALITÉS D'ÉVALUATION

- Feuille de présence signée en demi-journée,
- Evaluation des acquis tout au long de la formation,
- Questionnaire de satisfaction,
- Attestation de stage à chaque apprenant,
- Positionnement préalable oral ou écrit,
- Evaluation formative tout au long de la formation,
- Evaluation sommative faite par le formateur ou à l'aide des certifications disponibles



MOYENS TECHNIQUES EN PRÉSENTIEL

Accueil des stagiaires dans une salle dédiée à la formation équipée à minima d'un vidéo projecteur et d'un tableau blanc et/ou paperboard.

Pour les formations nécessitant un ordinateur, un PC est mis à disposition de chaque participant.



MOYENS TECHNIQUES EN DISTANCIEL

A l'aide d'un logiciel (Teams, Zoom...), d'un micro et éventuellement d'une caméra les apprenants interagissent et communiquent entre eux et avec le formateur.

Sessions organisées en inter comme en intra entreprise.

L'accès à l'environnement d'apprentissage ainsi qu'aux preuves de suivi et d'assiduité (émargement, évaluation) est assuré.

Pour toute question avant et pendant le parcours, assistance technique à disposition au 04 67 13 45 45.



ORGANISATION

Délai d'accès : 5 jours ouvrés
(délai variable en fonction du financeur)

Les cours ont lieu de 9h à 12h30 et de 13h30 à 17h



ACCESSIBILITÉ

Les personnes en situation de handicap sont invitées à nous contacter directement, afin d'étudier ensemble les possibilités de suivre la formation.

Pour tout renseignement, notre référent handicap reste à votre disposition : mteyssedou@ait.fr



PROFIL FORMATEUR

Nos formateurs sont des experts dans leurs domaines d'intervention

Leur expérience de terrain et leurs qualités pédagogiques constituent un gage de qualité.



CERTIFICATION POSSIBLE

Aucune

BigData, architecture et technologies

INTRODUCTION

- Le besoin : volumes importants de données, traitements optimisés de flux de données au fil de l'eau (nouvelles technologies et nouveaux usages)
- Domaines concernés : recherche scientifique, médical, e-commerce, sécurité...
- Développement des techniques sur différents aspects : stockage, indexation/recherche, calcul
- Définition de ETL : Extract Transform Load
- Les acteurs aujourd'hui

STOCKAGE

- Caractéristiques NoSQL :
- Structure de données proches des utilisateurs, développeurs
- Données structurées et non structurées, documents, images
- Fichiers XML, JSON, CSV...
- Les différents modes et formats de stockage
- Stockage réparti : réplication, sharding, gossip protocol, hachage
- Systèmes de fichiers distribués : GFS, HDFS, BigTable...
- Les bases de données
- Quelques exemples de produits et leurs caractéristiques : cassandra, MongoDB, CouchDB, DynamoDB...

INDEXATION ET RECHERCHE

- Moteurs de recherche
- Principe de fonctionnement
- Méthodes d'indexation
- Exemple de Lucene, et mise en oeuvre avec solr
- Recherche dans les bases de volumes importants
- Exemples de produits et comparaison : dremel, drill, elasticsearch, MapReduce...

CALCUL ET RESTITUTION, INTÉGRATION

- Différentes solutions : calculs en mode batch, ou en temps réel, sur des flux de données ou des données statiques
- Les produits : langage de calculs statistiques, R Statistics Language
- Outils de calcul sur des volumes importants : storm en temps réel, hadoop en mode batch
- Zoom sur Hadoop : complémentarité de HDFS et MapReduce

EVOLUTIONS

- Les offres SaaS BigData, type Google BigQuery
- Les limites actuelles. Les avancées annoncées