Big Data – Architecture et infrastructure Hadoop On-Premise et sur le Cloud

INFORMATIONS GÉNÉRALES

Type de formation : Formation continue

Éligible au CPF : Non

Domaine : IA, Big Data et Bases de données

Action collective : Non

Filière : Oracle

Rubrique : Conception - développement

Code de formation : RDBN401

PRÉSENTATION

Objectifs & compétences

A l'issue de la formation, le stagiaire sera capable :

- Exploiter les architectures Big Data Hadoop On-Premise et Cloud
- Mettre en place des socles techniques complets pour des projets Big Data.

Public visé

Administrateurs systèmes, développeurs, architectes et chefs de projets techniques,consultants en Business Intelligence.

Pré-requis

Avoir une connaissance pratique du Shell Linux, des bases du réseau, des bases de données relationnelles.

PROGRAMME

Jour 1
Présentation du Big Data et d'Hadoop
– Introduction au Big Data : de quoi s'agit-il ?
– Perspectives offertes par le Big Data
– Les acteurs du Big Data
– Exemples pratiques
– Démystification du Big Data
– Les technologies concernées
– Les outils Hadoop
– Les différents métiers du Big Data
– Les distributions Big Data
– Démonstration de traitements avec Hadoop

Jour 2
Solutions Big Data avec Hadoop
– Projets, applications, plateformes
– Solutions de stockage des données
– Considérations Hardware
– Stockage
– CPU
– Mémoire
– Réseau
– Systèmes distribués
– La science et l'art de l'analyse prédictive
– Aspects économiques (OPEX, CAPEX, TRI)

Jour 3

€ Tarifs

Prix public : 3800 €

Tarif & financement :
Nous vous accompagnons pour trouver la meilleure solution de financement parmi les suivantes :
Le plan de développement des compétences de votre entreprise : rapprochez-vous de votre service RH.
Le dispositif FNE-Formation.
L'OPCO (opérateurs de compétences) de votre entreprise.
France Travail: sous réserve de l'acceptation de votre dossier par votre conseiller Pôle Emploi.
CPF -MonCompteFormation
Contactez nous pour plus d'information : contact@aston-institut.com

Lieux & Horaires

Durée : 35 heures

Délai d'accès : Jusqu'a 8 jours avant le début de la formation, sous condition d'un dossier d'insciption complet

Prochaines sessions

Consultez-nous pour les prochaines sessions.

Informations non contractuelles - PDF généré le 7/11/2026

Architecture cluster applicative et technique avec Hadoop
– Les différentes étapes de gestion des données
– Les outils du marché
– L'ingestion
– Le stockage
– La transformation
– La restitution et la visualisation

Jour 4
Installation et configuration d'un cluster Hadoop On-Premise et sur le Cloud Hadoop
– Installation d'une plateforme Hadoop
– Définir l'architecture
– Installation d'Hadoop / HDFS et d'autres composants
– Installation des principaux composants (Hive, Pig, Impala...)
– Configuration de l'environnement
– Etude des fichiers de configuration
– Modes de configuration (ligne de commandes, API Rest, serveur HTTP intégré, API natives)
– Configuration des services (HDFS, YARN, logs...)
– Configuration des droits d'accès

Configuration d'un cluster hybride On-Premise / Cloud
– Configuration d'Hadoop
– Configuration de l'infrastructure Cloud

Jour 5
Administration d'un cluster Hadoop
– Gestion des ressources (noeuds, disques, CPU, mémoire, réseau...)
– Administration de HDFS
– Administration de MapReduce
– Infrastructure physique des centres de données, serveurs, racks, réseaux
– Optimisation au niveau du cluster (HDFS, MapReduce), noeuds / OS et réseau (binding,
jumbo frames)
– Planification des stratégies de sécurité et de récupération des données et de reprise sur panne

Exemples de travaux pratiques (à titre indicatif)
– Installation, congiguration et administration pas à pas d'un cluster Hadoop On-Premise et de l'infrastructure Cloud associée
– Data Warehousing avec Sqoop, Hue, Hive, Impala
– Conception et optimisation d'architectures cluster matérielles
(racks / noeuds / CPU / RAM / disques...)
– Comparatif technico-économique entre une solution On-Premise et une solution Cloud

MODALITÉS

Modalités

Modalités : en présentiel, distanciel ou mixte . Toutes les formations sont en présentiel par défaut mais les salles sont équipées pour faire de l'hybride. – Horaires de 9H à 12H30 et de 14H à 17H30 soit 7H – Intra et Inter entreprise.
Pédagogie : essentiellement participative et ludique, centrée sur l'expérience, l'immersion et la mise en pratique. Alternance d'apports théoriques et d'outils pratiques.
Ressources techniques et pédagogiques : Support de formation au format PDF ou PPT Ordinateur, vidéoprojecteur, Tableau blanc, Visioconférence : Cisco Webex / Teams / Zoom.
Pendant la formation : mises en situation, autodiagnostics, travail individuel ou en sous-groupe sur des cas réels.

Méthode

Fin de formation : entretien individuel.
Satisfaction des participants : questionnaire de satisfaction réalisé en fin de formation.
Assiduité : certificat de réalisation.
Validations des acquis : grille d'evalution des acquis établie par le formateur en fin de formation.

Informations non contractuelles - PDF généré le 7/11/2026