Présentiel/distanciel

Apache Hadoop Développer des applications pour le Big Data

Objectifs pédagogiques

- Développer une application de traitements distribués sur de gros volumes de données ;
- Configurer et optimiser Hadoop pour chaque besoin ;
- Organiser et construire un cluster Hadoop ;
- Prototyper et intégrer rapidement de nouveaux besoins grâce à Hadoop Streaming et Python.

Prérequis

Maîtriser impérativement un langage de programmation orientée objet tel que Java ou Python.

Public concerné

Développeurs backend et architectes SI.

Présentiel/distanciel

Programme détaillé

MapReduce en pratique

- Le » Hello World » BigData : WordCount
- Concepts
- Jobs, tasks, slots, splits ...
- WebUI
- Input et Ouput Format
- Partitionner, Grouper, Trier
- Application, Jobs et Ordonnancement
- Configuration, tunning, optimisations
- Uses cases
- Bonnes pratiques et patterns
- Travaux pratiques :
-
  - Développer le WordCount avec MapReduce

HDFS

- Architecture
- Patterns et usages
- Opérations et commandes
- WebHDFS, interfaces
- API Java
- Travaux pratiques :
-
  - Manipuler HDFS avec les commandes de base
  - Utiliser l'API Java pour lire/écrire des fichiers dans HDFS

YARN

- MRv1 vs YARN
- Architecture et composants
- Abstraction Tez : Pig, Hive, Cascading
- Travaux pratiques :
-
  - Recréer un moteur de recherche avec Hadoop Partie 1

Hadoop Streaming

- Python avec Hadoop Streaming
- Fonctionnement et configuration d'Hadoop Streaming
- Travaux pratiques :
-
  - Transformer le WordCount avec Hadoop Streaming
  - Recréer un moteur de recherche avec Hadoop Partie 2

Délais d'accès à la formation

Les inscriptions sont possibles jusqu'à 48 heures avant le début de la formation.
Dans le cas d'une formation financée par le CPF, ENI Service est tenu de respecter un délai minimum obligatoire de 11 jours ouvrés entre la date d'envoi de sa proposition et la date de début de la formation.

Modalités et moyens pédagogiques, techniques et d'encadrement

Formation avec un formateur, qui peut être suivie selon l'une des 3 modalités ci-dessous :
1 Dans la salle de cours en présence du formateur.
2 Dans l'une de nos salles de cours immersives, avec le formateur présent physiquement à distance. Les salles immersives sont équipées d'un système de visio-conférence HD et complétées par des outils pédagogiques qui garantissent le même niveau de qualité.
3 Depuis votre domicile ou votre entreprise. Vous rejoignez un environnement de formation en ligne, à l'aide de votre ordinateur, tout en étant éloigné physiquement du formateur et des autres participants. Vous êtes en totale immersion avec le groupe et participez à la formation dans les mêmes conditions que le présentiel. Pour plus d'informations : Le téléprésentiel notre solution de formation à distance.

Le nombre de stagiaires peut varier de 1 à 12 personnes (5 à 6 personnes en moyenne), ce qui facilite le suivi permanent et la proximité avec chaque stagiaire.
Chaque stagiaire dispose d'un poste de travail adapté aux besoins de la formation, d'un support de cours et/ou un manuel de référence au format numérique ou papier.
Pour une meilleure assimilation, le formateur alterne tout au long de la journée les exposés théoriques, les démonstrations et la mise en pratique au travers d'exercices et de cas concrets réalisés seul ou en groupe.

Modalités d'évaluation des acquis

En début et en fin de formation, les stagiaires réalisent une auto-évaluation de leurs connaissances et compétences en lien avec les objectifs de la formation. L'écart entre les deux évaluations permet ainsi de mesurer leurs acquis.
En complément, le formateur évalue chaque stagiaire sur l'atteinte des objectifs pédagogiques de la formation selon quatre niveaux (non évalué, non acquis, en cours d'acquisition, acquis). Cette évaluation repose sur une modalité choisie par le formateur en cohérence avec la formation : QCM, exercices pratiques réalisés pendant la formation, évaluation finale de synthèse, quiz interactif de validation, étude de cas, mise en situation, analyse de l'auto-évaluation, autres modalités adaptées.
Pour les stagiaires qui le souhaitent, certaines formations peuvent être validées officiellement par un examen de certification. Les candidats à la certification doivent produire un travail personnel important en vue de se présenter au passage de l'examen, le seul suivi de la formation ne constitue pas un élément suffisant pour garantir un bon résultat et/ou l'obtention de la certification.
Pour certaines formations certifiantes, le passage de l'examen de certification est inclus et réalisé en fin de formation. Les candidats sont alors préparés par le formateur au passage de l'examen tout au long de la formation.

Moyens de suivi d'exécution et appréciation des résultats

Feuille de présence, émargée par demi-journée par chaque stagiaire et le formateur.
Evaluation qualitative de fin de formation, qui est ensuite analysée par l'équipe pédagogique ENI.
Attestation de fin de formation, remise au stagiaire en main propre ou par courrier électronique.

Qualification du formateur

La formation est animée par un professionnel de l'informatique et de la pédagogie, dont les compétences techniques, professionnelles et pédagogiques ont été validées par des certifications et/ou testées et approuvées par les éditeurs et/ou notre équipe pédagogique.
Il est en veille technologique permanente et possède plusieurs années d'expérience sur les produits, technologies et méthodes enseignés.
Il est présent auprès des stagiaires pendant toute la durée de la formation.

Accessibilité de la formation