



ENI Service

référence
T265-020

14h

Apache Hadoop Présentation de l'écosystème

Mise à jour
17 juillet 2023

Formation
intra-entreprise
sur devis

 (Télé-)présentiel

Apache Hadoop Présentation de l'écosystème



Objectifs

- ✓ Décrire l'architecture et identifier les composants d'Apache Hadoop
- ✓ Configurer et superviser les services
- ✓ Identifier et manipuler les outils de développement et de supervision Hadoop
- ✓ Exploiter et orchestrer les développements

Pré-requis

Connaissances générales des systèmes d'information.

Public

- DSI
- Directeur Technique
- Chef de Projet
- Architecte SI
- Administrateur de bases de données



 02 40 92 45 50

 formation@eni.fr

www.eni-service.fr

ENI Service - Centre de Formation

adresse postale : BP 80009 44801 Saint-Herblain CEDEX

SIRET : 403 303 423 00020 B403 303 423 RCS Nantes, SAS au capital de 864 880

1 / 3



ENI Service

référence
T265-020

14h

Apache Hadoop Présentation de l'écosystème

Mise à jour
17 juillet 2023

Formation
intra-entreprise
sur devis

 (Télé-)présentiel

Programme de la formation

Le projet et les modules

- Hadoop Common, HDFS, YARN, MapReduce Oozie, Pig, Hive, HBase
- Les fonctionnalités du framework Hadoop
- Les différentes versions
- Les principales distributions et leurs spécificités : Apache, Cloudera, Hortonworks, EMR, MapR
- Architecture et principe de fonctionnement
- Terminologie : NameNode, DataNode, ResourceManager, NodeManager
- Rôle des différents composants

Administration Hadoop

- Etude des fichiers de configuration core-site.xml, hdfs-site.xml, mapred-site.xml, yarn-site.xml et capacity-scheduler.xml
 - > Gestion des utilisateurs pour les daemons hdfs et yarn
 - > Droits d'accès sur les exécutables et répertoires
 - > Architecture et gestion des services généraux Hadoop des utilisateurs pour les daemons hdfs et yarn
- HDFS
- YARN
- MapReduce
- HBase
- Monitoring du cluster Hadoop
 - > Suivi de charges
 - > Journaux (jConsole)
 - > Gestion des noeuds
 - > Accès JMX
 - > Mise en oeuvre d'un client JMX
 - > Administration HDFS
 - > Stockage des fichiers : fsck, dfsadmin
 - > Gestion centralisée de caches avec Cacheadmin
- Sécurité
 - > Activation de la sécurité avec Kerberos dans core-site.xml, et hdfs-site.xml pour les NameNode et DataNode
 - > Gestion de la sécurité avec Apache Sentry
- Exploitation
 - > Supervision des éléments par le NodeManager
 - > Monitoring graphique avec Ambari, Kibana, Cloudera Manager
 - > Visualisation des alertes en cas d'indisponibilité d'un noeud
 - > Configuration des logs avec log4j

Développement Hadoop

- HDFS
 - > Architecture
 - > Commandes SHELL
- MapReduce
 - > Architecture de MapReduce
 - > Exécuter un code MapReduce

- Hue
 - > Introduction
 - > Fonctionnalités et utilisation
- HBase
 - > Architecture
 - > Commandes SHELL de création et d'interrogation de données
- Hive
 - > Architecture
 - > Méthodes d'accès à Hive
 - > HiveQL : création de bases de données/tables/vues, interrogation de données avec HiveQL, manipulation de fonctions définies par l'utilisateur (UDF), partitionner vos données, archivage de vos données
- Pig
 - > Introduction
 - > Méthodes d'exécution
 - > Pig Latin
 - > Communication entre Pig et Hive
- Sqoop
 - > Introduction
 - > Pour quels usages ?
 - > Méthodes d'utilisation
 - > Import et export des données
- Oozie
 - > Introduction
 - > Planification de workflows avec paramètres



02 40 92 45 50

formation@eni.fr

www.eni-service.fr

ENI Service - Centre de Formation

adresse postale : BP 80009 44801 Saint-Herblain CEDEX

SIRET : 403 303 423 00020 B403 303 423 RCS Nantes, SAS au capital de 864 880

2 / 3



ENI Service

référence
T265-020

14h

Apache Hadoop Présentation de l'écosystème

Mise à jour
17 juillet 2023

Formation
intra-entreprise
sur devis

 (Télé-)présentiel



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, 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 (ex : ITIL, DPO, ...), 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.

Évaluation 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.



02 40 92 45 50

formation@eni.fr

www.eni-service.fr

ENI Service - Centre de Formation

adresse postale : BP 80009 44801 Saint-Herblain CEDEX

SIRET : 403 303 423 00020 B403 303 423 RCS Nantes, SAS au capital de 864 880

3 / 3