

Hortonworks

5 j (35 heures)

Ref : HORTON

Public

Administrateurs système, consultants SI

Pré-requis

Maîtriser les commandes Shell

Moyens pédagogiques

Formation réalisée en présentiel ou à distance selon la formule retenue
Exposés, cas pratiques, synthèse, assistance post-formation pendant trois mois
Un poste par stagiaire, vidéoprojecteur, support de cours fourni à chaque stagiaire

Modalités de suivi et d'évaluation

Feuille de présence émargée par demi-journée par les stagiaires et le formateur
Exercices de mise en pratique ou quiz de connaissances tout au long de la formation permettant de mesurer la progression des stagiaires
Questionnaire d'évaluation de la satisfaction en fin de stage
Auto-évaluation des acquis de la formation par les stagiaires
Attestation de fin de formation

Objectifs

- Maîtriser l'administration des clusters Hortonworks
- Apprendre à élaborer une architecture Hadoop
- Apprendre à mettre en oeuvre un cluster Hadoop
- Configurer et monitorer le cluster avec Ambari
- Savoir gérer la maintenance des noeuds dans un cluster

Programme détaillé

PRESENTATION

Big Data et Hadoop

MISE EN OEUVRE

Configuration de l'environnement

Installation de HDP

GESTION DES DROITS UTILISATEURS

Gestion des utilisateurs et des groupes d'Ambari

LA COUCHE SERVICES

Gestion des services Hadoop

LE SYSTEME HDFS (HADOOP DISTRIBUTED FILE SYSTEM)

Utilisation du stockage HDFS

Utilisation de Hadoop Storage

Utilisation de WebHDFS

Utilisation des listes de contrôle d'accès HDFS

Gestion du stockage HDFS

Gestion du stockage Hadoop

Gestion des quotas HDFS

Ajout, remplacement et suppression de noeuds de travail

GESTION DES RACKS

Configuration de la prise en compte des racks (rack awareness)

GESTION DES RESSOURCES AVEC YARN

Présentation et architecture de YARN

Gestion du service YARN à l'aide de l'interface utilisateur Web Ambari

Gestion du service YARN à l'aide de l'interface ligne de commande CLI

Applications YARN

Exécution d'exemples d'applications YARN

Planificateur de capacité YARN

Configuration de planificateur de capacité (Capacity Scheduler)

Gestion des conteneurs et files d'attente de YARN

Gestion des listes de contrôle d'accès (ACL) de YARN et des droits utilisateurs

HAUTE DISPONIBILITE

Configuration de HDFS et YARN en haute disponibilité (High Availability)

Configuration d'un NameNode en HA

Configuration de ResourceManager en HA

MONITORING

Hortonworks

Suivi d'un cluster

Gestion des alertes d'Ambari

SYSTEME DE BACKUP

Protéger un cluster avec des sauvegardes

Gestion des photos instantanées HDFS

Utilisation de DistCp
