

FORMATION CERTIFIANTE HADOOP ADMINISTRATEURS

FORMATEUR : CLUDERA / DUREE : 3 JOURS (21 HEURES)

Cette formation de 3 jours s'adresse aux administrateurs systèmes et aux personnes qui gèrent et maintiennent les clusters Apache Hadoop en production ou qui évoluent dans un environnement de développement.

Xebia Training réitère pour 2013 son partenariat avec Cloudera en tant que centre de formation officiel pour les certifications Hadoop. Nous proposons des formations certifiantes Hadoop Apache pour administrateurs et développeurs. Ces formations sont dispensées à Paris en langue française. Pour la session de mars, une copie gratuite du livre "Hadoop Operations" par Eric Sammer vous est offerte (format livre électronique).

PRÉSENTATION

Cette formation de 3 jours s'adresse aux administrateurs systèmes et aux personnes qui gèrent et maintiennent les clusters Apache Hadoop en production ou qui évoluent dans un environnement de développement. Il est utile d'avoir une expérience dans l'administration système Linux. Aucune expérience avec Hadoop n'est nécessaire pour suivre cette formation et obtenir la certification.

VOUS APPRENDREZ:

- Comment fonctionne « Hadoop Distributed File System » et « MapReduce »
- La configuration de matériel la mieux adaptée pour les clusters Hadoop
- Comment gérer le réseau dans la mise en place du cluster ?
- Comment configurer les options Hadoop pour assurer une meilleure performance des clusters ?
- Comment configurer FairScheduler afin d'assurer une charte de qualité pour les différents utilisateurs d'un cluster ?
- Comment maintenir son cluster ?
- Comment charger les données dans un cluster avec Flume et Sqoop ?
- Quelles problématiques d'administration système peuvent survenir avec d'autres projets Hadoop tel que Hive, Pig ou Hbase ?

PROGRAMME

UNE INTRODUCTION A HADOOP ET HDFS

- Pourquoi choisir Hadoop
- Hadoop: les concepts de base
- Rédiger un programme « MapReduce »
- L'écosystème Hadoop
- Intégrer Hadoop dans un « Workflow »
- Les algorithmes « MapReduce » communs
- Travailler avec Apache Hive et Pig
- Débugger des programmes « MapReduce »
- Le développement « MapReduce » avancé
- L'utilisation de Graph avec Hadoop
- Le nouvel API Hadoop
- L'examen « Cloudera Certified Developer »

PLANNING YOUR HADOOP CLUSTER

FORMATION CERTIFIANTE HADOOP DÉVELOPPEURS

FORMATEUR : JULIEN BURET / DUREE : 4 JOURS (28 HEURES)

- Choisir le bon matériel
- Réseau
- “Configuring Nodes”

CONFIGURER ET DEPLOYER UN CLUSTER

- Types de déploiement
- Installer Hadoop
- Utiliser “Cloudera Manager for Easy Installation”
- Les paramètres de configuration typiques
- Utiliser “Configuration Management Tools”
- Exercices pratiques

GERER LES TACHES

- Gérer les tâches en cours
- Exercices pratiques
- Le “FIFO Scheduler”
- Le “FairScheduler”
- Exercices pratiques

GESTION DE CLUSTERS

- HDFS
- Exercices pratiques : copier des données entre clusters
- Rajouter et retirer des « Cluster Nodes »
- Rééquilibrer un Cluster
- Exercices pratiques : Optimiser un Cluster

CLUSTER MONITORING ET TROUBLESHOOTING

- Gérer les “Log files” d’Hadoop
- Utiliser le “NameNode” et “Job Tracker Web UI”
- Exercices pratiques
- Gestion de Cluster avec Ganglia
- Troubleshooting
- Benchmarking un Cluster

CONCLUSION

CLOUDERA CERTIFIED ADMINISTRATOR FOR APACHE HADOOP EXAM