

Formation Big Data, Sécurité des données

Présentation

La sécurisation des clusters Hadoop est une problématique de haute importance pour chaque organisation qui l'utilise. Permettant à une organisation de conserver des données en toute sécurité, la sécurisation des clusters Hadoop permet également de s'assurer que l'organisation est conforme à diverses normes comme PCI ou HIPAA.

Cette formation vous apprendra à mettre en œuvre des clusters Hadoop sécurisés en utilisant l'authentification, l'autorisation et le chiffrement.

Durée : 14,00 heures (2 jours)

Tarif INTRA : [Nous consulter](#)

Objectifs de la formation

- Comprendre la qualification complexe des données
- Identifier les principaux risques touchant les solutions de traitement des données massives
- Maîtriser le cadre juridique (CNIL et PLA (Privacy Level Agreement))
- Connaître les principales solutions techniques de base pour se protéger des risques
- Savoir mettre en œuvre une politique de sécurité pour traiter les risques, les menaces, les attaques

Prérequis

- Avoir de bonnes connaissances dans la sécurité réseau et système,
- Connaître les plateformes Hadoop.

Public

- Ingénieurs de données,
- Développeurs logiciel,
- Consultants sécurité et SI,
- Administrateurs système.



Programme de la formation

Aperçu de la sécurité Hadoop

- Qu'est-ce que Hadoop Security ?
- Pourquoi est-ce important ?
- Aspects de sécurité : Éléments clés à prendre en compte
- Sécuriser l'écosystème Hadoop
- Mettre en place un cluster multi-noeuds de Cloudera

Authentification

- Approches, avantages et inconvénients de l'authentification
- Introduction à Kerberos
- Comment travailler avec Kerberos ?
- Comment activer Kerberos
- Authentification dans Cloudera Manager (CM)
- Exécution des commandes HDFS
- Travailler avec les applications YARN
- Effectuer une analyse ad hoc en utilisant Hive
- Intégration de Kerberos avec Windows Active Directory (AD)
- Avantages de travailler avec l'AD
- Installation de Windows Active Directory Server
- Configuration du serveur AD
- Intégration de Hue avec Active Directory
- Préparer un cluster avec Kerberos
- Authentification
- Intégration de Kerberos avec Active Directory
- Activation de l'identification unique (Single Sign On)

Autorisation

- Principaux frameworks d'autorisation
- Qu'est-ce qu'Apache Sentry ?
- Travailler avec l'autorisation Sentry
- Intégration de Sentry avec HUE
- Interroger Hive
- Listes de contrôle (ACLs) étendues HDFS
- Limitations de Sentry

- Qu'est-ce que Cloudera Record Service ? Son utilité.
- Mise en œuvre du service d'enregistrement (Record Service)

Chiffrement

- Types de chiffrement
- Chiffrement au niveau du système d'exploitation
- Chiffrement HDFS
- Mise en place d'une zone de chiffrement HDFS
- Travailler avec les zones de chiffrement HDFS
- Chiffrement des données en mouvement
- Introduction aux outils SSL
- Utilisation d'un certificat auto-signé avec une autorité de certification racine (Self-Signed Root CA)
- Activation et validation du SSL pour Hadoop Core
- SASL Hive et HiveServer2
- SSL avec Hue

Autres sujets liés à la sécurité (si le temps le permet)

- Audit avec Cloudera Navigator
- Les ACLs HBase
- Usurpation d'identité
- Masquage des données sensibles
- Connexion à l'aide d'une clé
- Fichier à onglets
- Fondamentaux UserGroupInformation

Organisation

Formateur

Les formateurs de Docaposte Institute sont des experts de leur domaine, disposant d'une expérience terrain qu'ils enrichissent continuellement. Leurs connaissances techniques et pédagogiques sont rigoureusement validées en amont par nos référents internes.

Moyens pédagogiques et techniques

- Alternance d'apports théoriques et de mises en pratique contextualisées.
- Progression pédagogique structurée, de la compréhension à la mise en œuvre opérationnelle.
- Pédagogie immersive favorisant une mise en application immédiate des compétences.
- Formation interactive et participative s'appuyant sur des ateliers pratiques, études de cas et échanges collectifs (+ 60 % du temps de formation.)
- Séquences pédagogiques rythmées et évaluations formatives régulières pour renforcer l'ancrage des apprentissages.
- Supports de cours, ressources documentaires et références mis à disposition par le formateur

Dispositif de suivi de l'exécution et de l'évaluation des résultats de la formation

En amont de la formation *(évaluation en ligne dans l'extranet apprenant)*

- Recueil des besoins et des attentes des participants.
- Auto-positionnement permettant d'évaluer le niveau initial des apprenants.

Tout au long de la formation

- Évaluations continues réalisées par le formateur à travers des questions orales, exercices, QCM, études de cas et mises en situation pratiques.

À l'issue de la formation *(évaluations en ligne dans l'extranet apprenant)*

- Auto-positionnement afin de mesurer l'acquisition des compétences visées.
- Validation de l'acquisition des compétences par le formateur
- Questionnaire de satisfaction « à chaud » à l'issue de la formation
- Questionnaire de satisfaction "à froid", 45 jours après la formation afin d'évaluer l'ancrage des acquis et leur mise en application en situation professionnelle.

Accessibilité

Nos formations peuvent être adaptées à certaines conditions de handicap. Nous contacter pour toute information et demande spécifique.