



Hadoop : infrastructure sécurisée

CB035

Durée: 1 jour

Public :

Chefs de projet, administrateurs et toute personne souhaitant sécuriser une infrastructure hadoop.

Objectifs :

Comprendre les mécanismes de sécurité hadoop, et savoir les mettre en oeuvre.

Connaissances préalables nécessaires :

Connaissance des principes de base Hadoop et des bases de données.

Programme :

Introduction

Rappels rapides sur l'écosystème Hadoop.
Le projet et les modules : Hadoop Common, HDFS, YARN, Spark, MapReduce
Les risques et points à sécuriser dans un système distribué

Architecture sécurité hadoop

Sécurisation réseau, système d'exploitation, les rôles hadoop et stratégies

Kerberos

Principe de fonctionnement.

Atelier : kerberisation d'une grappe hadoop

Sécurité des accès

Authentification, autorisations, accounting.

Atelier : gestion des autorisations dans HDFS, YARN, HBase, ... Mise en oeuvre des ACLs dans Zookeeper



— Phirio —

Apache Sentry

Présentation du projet, architecture : sentry server, sentry plugin. Gestion de l'authentification et des droits d'accès aux données.

Atelier : intégration avec Hadoop

Sécurité des données

Cryptage des données stockées, et en transit. Mécanisme de sécurité des données en entrée et en consultation par des accès clients:
interface hadoop en ligne de commande, sqoop, oozie, HBase, webHDFS, httpFS