OSDA-21 - CLOUDERA HADOOP ADMINISTRATOR

Categoria: **Data & Artificial Intelligence**

INFORMAZIONI SUL CORSO

Durata:	Categoria:	Qualifica Istruttore:	Dedicato a:	Produttore:
4 Giorni	Data & Artificial Intelligence	Docente Senior (min. 5 anni)	Professionista IT	PCSNET

OBIETTIVI

Il corso si pone l'obiettivo di introdurre alla suite Cloudera Hadoop e alle sue componenti open source, dettagliando le modalità di archiviazione, elaborazione e analisi dei dati.

PREREQUISITI

-Conoscenza di Linux.
-Capacità di utilizzo della interfaccia da linea di comando.

CONTENUTI

Apache Hadoop

-Introduzione ad Apache Hadoop e concetti fondamentali
-Struttura e architettura di Hadoop

Hadoop Cluster Installation

-Cluster Management Solution
-Cloudera Manager: caratteristiche e istallazione
-Hadoop (CDH) Installation

Hadoop Distributed File System (HDFS)

-Caratteristiche principali di HDFS
-Leggere e scrivere dai files
-Web UIs per HDFS
-Hadoop File Shell

MapReduce

-Scopi
-MapReduce: concetti fondamentali
-Apache Spark
-Computational Frameworks
-Application Logs

Hadoop Configuration and Daemon Logs

- -Cloudera Manager Constructs per Managing Configurations
- -Gestire le Role Instances e i servizi
- -Configurare il servizio HDFS
- -Configurare i Logs

Importare i dati nel HDFS

- -Importare dati da fonti esterne
- -Importare dati da un DB relazionale
- -Interfacce REST
- -Best Practices

Gestire un Hadoop Cluster

- -Considerazioni generali
- -Considerazioni sull'Hadware
- -Considerazioni sulla rete
- -Configurare i nodi

Installing and Configuring Hive

Hadoop Clients

- -Scopi
- -Installare e configurare Hadoop Clients
- -Installare e configurare Hue

Cluster Maintenance

- -HDFS Status
- -Copiare i dati tra i clusters
- -Aggiungere e togliere Nodi
- -Ribilanciare iCluster

Cluster Monitoring and Troubleshooting

- -Cloudera Manager Monitoring
- -Monitoring Hadoop Clusters
- -Troubleshooting

INFO

Materiale didattico: Materiale didattico e relativo prezzo da concordare
Costo materiale didattico: NON incluso nel prezzo del corso
Natura del corso: Operativo (previsti lab su PC)