CLDR-2 - CLOUDERA DATA DEVELOPER

Categoria: **Cloudera**

INFORMAZIONI SUL CORSO

Durata:	Categoria:	Qualifica Istruttore:	Dedicato a:	Produttore:
5 Giorni	Cloudera	Docente Senior (min. 5 anni)	Sviluppatore	PCSNET

OBIETTIVI

Al termine del corso gli allievi avranno acquisito tutti gli elementi necessari all'implementazione di piattaforme Big Data per l'analisi ed il processamento dei dati.

PREREQUISITI

- Buona conoscenza della programmazione strutturata derivata dall'uso di un qualsiasi linguaggio di programmazione
- Buona conoscenza dei database e del linguaggio SQL
- Buona conoscenza del sistema operativo Linux

CONTENUTI

Data Ingestion

- Usare Apache Airflow
- Usare Nifi per connettere vari sistemi (HDFS e altri)
- Usare Apache Spark e Nifi per caricare dati dentro Apache Kudu

Costruire e gestire dati in ambiente Big Data e Datawarehouse

- Usare HDFS per memorizzare i dati
- Usare Impala e Hive per interrogare i dati

Costruire e manipolare data pipeline

- Usare Apache AirFlow per schedulare pipeline ETL
- Usare Apache Wifi per schedulare e trasformare dati

Lavorare i dati per renderli disponibili agli utenti finali

- Capire Apache Spark
- Trasformare e caricare i dati con Spark
- Usare vari formati di storage con Spark

Backup

- Usare Nifi per creare backup

-Fare Backup utilizzando Cloudera Data Engineering

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)