

Hadoop: Progattazione pratica di applicazioni

di

Gabriele Pellegrinetti

Tecnet Dati s.r.l.
Corso Svizzera 185
10149 – Torino
Tel. 011-7718090

<http://www.tecnetdati.com>
mail: info@tecnetdati.it



Hadoop: progettazione pratica di applicazioni

Tra gli strumenti per la memorizzazione e l'elaborazione dei Big Data, la piattaforma Hadoop si sta proponendo come standard di mercato. Ma come funziona Hadoop? Quali sono gli strumenti principali per il suo utilizzo? In quale modo è possibile combinare questi strumenti per ottenere applicazioni realmente utilizzabili? Quando scegliere uno strumento piuttosto che un altro? Questo corso, di carattere tecnico-pratico, dopo aver descritto le componenti della piattaforma e l'architettura delle applicazioni Hadoop-Based, vuole fornire le basi per la progettazione delle applicazioni tramite la proposizione di esempi pratici, risolti insieme ai partecipanti. Gli esempi sono realizzati utilizzando le tecnologie Hive, Pig, Spark e map/reduce.

Destinatari:

- progettisti
- sviluppatori

Prerequisiti:

- Conoscenze base sui Big Data, sul linguaggio SQL, sulla programmazione in Java e in altri linguaggi.

Requisiti per l'aula:

- videoproiettore con risoluzione nativa minima di 1024x768 (meglio se superiore) funzionante.
- Connessione ad internet non filtrata per il portatile del docente (preferibilmente via cavo)
- lavagna a fogli mobili con pennarelli di diversi colori

Contenuti:

- Breve ripasso sui Big Data:
 - cosa sono i Big Data;
 - strumenti disponibili;
 - architettura di riferimento.
- La piattaforma Hadoop:
 - file system distribuito HDFS;
 - memorizzazione e recupero delle informazioni;
 - YARN, Map/Reduce e TEZ;
 - l'ecosistema Hadoop (Hive, PIG, Spark, Hbase,...);
 - le componenti per la gestione della sicurezza (Knox, Ranger,...);
 - altre tecnologie correlate.

- Progettare applicazioni per hadoop:
 - architettura delle applicazioni Hadoop;
 - pattern di riferimento per l'ingestion, l'elaborazione e la fruizione dei dati;
 - come scegliere le giuste tecnologie.
- Esercitazioni pratiche:
 - presentazione del caso di studio;
 - progettazione delle componenti principali dell'applicazione;
 - progettazione e implementazione delle principali componenti (es. query hive, script PIG,...).

Durata:

3 giorni