

Giornata di Studio

Big Data: Applicazioni e tecnologie abilitanti

Mercoledì, 29 ottobre 2014

Sala Conferenze del Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB)
Politecnico di Milano, Via Ponzio 34/5

Analizzare grandi quantitativi di dati è oggi una priorità per molte aziende, sia che questi siano prodotti internamente, sia che provengano da fonti esterne (social network, sensori, etc.) L'aumento dei volumi informativi prodotti e manipolati molto spesso eccede la capacità delle tecnologie tradizionali (data base). Inoltre, le prestazioni dei sistemi di storage non sono ancora al passo con i miglioramenti ottenuti a livello computazionale o di rete. A ciò si aggiunge la necessità di elaborare i dati in tempo reale, all'atto della loro produzione, senza doverli preliminarmente conservare in un data base, come previsto dalle tecniche di data management tradizionali. Tutti questi fattori stanno portando l'interesse degli esperti IT su tecnologie e paradigmi alternativi, che possano soddisfare tali requisiti. MapReduce e la sua implementazione open-source Hadoop consentono una soluzione efficiente per l'analisi offline di grandi volumi di dati. Hadoop, inoltre, costituisce l'infrastruttura di base al di sopra della quale sono state realizzati framework di più alto livello come Pig e Hive. In maniera analoga, sono stati sviluppati diversi sistemi di Data Streaming, che consentono l'elaborazione in tempo reale di grandi volumi di dati prodotti da sorgenti eterogenee.

La giornata si concentrerà sui principali aspetti legati ai Big Data, illustrando le problematiche inerenti e le tecnologie proposte per risolverle, da Apache Hadoop ai principali sistemi di Data Streaming. Verranno presentati inoltre alcuni esempi di applicazioni di tali tecnologie in ambienti di produzione.

PROGRAMMA

- 09.30** **Apertura e saluto**
Andrea Penza, *Presidente AICT*
- 10.00** **Introduzione**
Emanuele Della Valle, *Politecnico di Milano*
- 10.30** **Data Streaming: When Big Data Flies**
Gianpaolo Cugola, *Politecnico di Milano*
- 11.00** **MapReduce e Hadoop Ecosystem**
Daniilo Ardagna, *Politecnico di Milano*
- 11.30** *Coffee break*
- 12.00** **HDInsight: Applicazioni e soluzioni di Business**
Davide Luzzu, *Microsoft*
- 12.45** *Pranzo libero*
- 13.45** **IBM Infosphere Streams e BigInsights - Benefici di una piattaforma Big Data multipurpose**
Filippo Mariani, *IBM*
- 14.30** **Genomic Applications**
Stefano Ceri, *Politecnico di Milano*
- 15.00** **Tavola Rotonda e chiusura dei lavori**