



Unione Europea



REGIONE CAMPANIA



REPUBBLICA ITALIANA



ISTITUTO SUPERIORE
PALMIERI RAMPONE POLO

Ministero dell'Istruzione, dell'Università e della Ricerca

ISTITUTO SUPERIORE PALMIERI - RAMPONE - POLO

Istituto Professionale Industria e Artigianato "Luigi Palmieri" via Traiano Boccalini, 23-25 Benevento Tel. 0824.24806 - Fax 0824.21094

Settori: Elettrotecnico ed Elettronico - Termoidraulico - Abbigliamento e Moda - Meccanico Automobilistico - Produzioni Audiovisive

Istituto Tecnico Commerciale e Industriale "Salvatore Rampone" via Luigi Stasi, 6 Benevento Tel. 0824.25984 - Fax 0824.22331

Settori: Economico - Informatico - Grafico

Cod Mecc. BNIS027006 - Cod. Fisc. 92057600626 ✉ bnis027006@istruzione.it ✉ bnis027006@pec.istruzione.it

www.palmieriramponepolo.gov.it

TITOLO

Progettazione di un'infrastruttura per l'analisi dei dati in (near) real-time
provenienti da una rete IoT
(..verso la certificazione Cloudera)

ABSTRACT

L'Internet Of Things, permette la comunicazione, e quindi lo scambio, di grandi moli di dati, quasi in tempo reale, utilizzando la rete globale al fine di ottimizzare i processi produttivi. La diffusione di questo fenomeno, insieme ad altri fattori come l'utilizzo sempre più crescente dei social-media, ha posto le basi al "web pervasivo", il quale ha permesso la comunicazione tra qualsiasi utente, e tra utente e macchina o qualsiasi "smart objects", in qualsiasi momento, dovunque ci si trovi:

the right content, at the right time, in the right place on the right device

La quantità di dati circolanti in rete, ha messo in crisi le infrastrutture di calcolo, ed archiviazione("storage") tradizionali, evidenziandone la difficoltà nel soddisfare, in modo adeguato, la richiesta crescente di capacità computazionale, soddisfacendo il requisito di bassa latenza, soprattutto nel recupero("retrieving") dei dati.

Si concretizza, quindi, la necessità di sviluppare nuovi sistemi, realizzati per mezzo degli strumenti e delle tecnologie più moderne, che permettano di raccogliere e gestire questi "big dati", al fine di produrre i risultati migliori nel minor tempo possibile, generalmente nell'ordine dei secondi, tempo minimo impiegato dalle macchine per processarli(near real-time).

I ragazzi della quinta B informatico del Rampone, guidati dalle prof.sse Genni Rotondi e Marianonietta De Pascale, progettano e realizzano un'infrastruttura Big Data per l'acquisizione, l'analisi, e l'estrapolazione di informazioni, da dati in streaming provenienti da reti IOT, lavorando quindi in tempo reale.

I frameworks utilizzati sono:

Kafka : piattaforma in continua crescita, sviluppata per velocizzare e per massimizzare la sicurezza circa i flussi di dati prodotti dalla messaggistica. Il programma è passato dalla gestione di un miliardo di messaggi agli oltre tre trilioni di oggi, dimostrandosi sempre più desiderato dalle grandi aziende, tra cui Linkedin(creatore del tool), Twitter, Netflix e Mozilla. E' di sicuro la scelta giusta per coloro che si occupano della gestione di grandi volumi di dati.

Spark: che permette di caricare ed elaborare dati per ottenere informazioni utili, può effettuare operazioni direttamente in memoria centrale riuscendo ad offrire prestazioni

anche 100 volte superiori su applicazioni specifiche. Il tutto contestualizzato in ambiente *Linux* nuovo per i ragazzi del Rampone, utilizzando una macchina virtuale *hypervisor di tipo 2 in ambiente Microsoft*. Partendo da un esempio dettagliato di implementazione ed utilizzo, i ragazzi conducono analisi descrittive dei fenomeni di interesse, con l'obiettivo di interpretare le rilevazioni che si registreranno nel corso del tempo, capendo scostamento dal valore centrale, e conducendo misurazioni statistiche (correlazioni) al fine di prevedere tali discrepanze. quantificare il bias dei dati, ovvero lo scostamento tra l'output previsto e quello verificatosi.

Spark e kafka rappresentano valide alternative per la gestione di moli di dati in tempo quasi reale, oltre, ad essere dei must-to-know nel settore Big Data, in continua espansione e destinato a cambiare il mondo in cui viviamo: sanità, automotive, difesa, aereospazio in maniera sempre più rilevante.

Queste tecnologie soprattutto se certificate ,a livello internazionale contribuiscono ad aumentare la possibilità di trovare lavoro e di ottenere retribuzioni gratificanti, come dimostrano i dati relativi agli stipendi percepiti dai professionisti del settore, quelle piu richieste sono quelle CLAUDERA.

L'analisi di dati live consente alle aziende di porre in essere strategie di business migliori in brevissimo tempo, tenendo conto del quadro economico, e sensibilizzando il modello ad eventuali cambiamenti futuri del contesto di operatività, con lo scopo di incrementare i profitti ottenuti.