



Proposta di Tesi

TITOLO:

Performance optimization of a cloud-based data analysis software package

DESCRIZIONE:

Partendo dall'analisi di un software esistente, che si occupa di analizzare i dati presenti su cloud provenienti da diverse linee di imbottigliamento, lo studente si occuperà dell'ottimizzazione di: manutenibilità, performance e possibilità di espansione del software in previsione di un aumento esponenziale del numero di macchine collegate al cloud (e quindi del volume di dati da analizzare).

In particolare lo studente si occuperà di:

- Identificare ed implementare la migliore architettura software Python
- Definire l'architettura cloud per ottimizzare tempi e costi (VM singola, Container, serverless, etc.)
- Ottimizzare l'architettura delle macchine virtuali da utilizzare (RAM, CPU, GPU, etc.)

La tesi è rivolta ai percorsi di Laurea Magistrale in Ingegneria Informatica.

La tesi sarà svolta presso l'azienda con una durata prevista di 6 mesi.

Requisiti preferenziali:

- Buona conoscenza dei linguaggi (Python, SQL)
- Conoscenza di tecniche di Machine Learning, Big Data analysis
- Propensione al lavoro in team
- Conoscenza dell'ambiente GCP - Google Cloud Platform (opzionale)

Opportunità

- Collaborazione con esperti di cloud e machine learning
- Accesso a corsi e documentazione

Referente universitario:

prof. **Michele Amoretti** (michele.amoretti@unipr.it, tel. 0521 906390)

Distributed Systems Group, Dipartimento di Ingegneria e Architettura