



Università degli Studi di Salerno

Dipartimento di Scienze Aziendali – Management and Innovation Systems

Dottorato di Ricerca

Management and Information Technology

Curriculum in Informatica, Sistemi Informatici e Tecnologie del Software

XV Ciclo

Tesi di Dottorato

Parallel Genetic Algorithms in the Cloud

Tutor

Prof.ssa Filomena Ferrucci

Coordinatore

Prof. Andrea De Lucia

Candidato

Pasquale Salza

Matr. 8887600010

Anno Accademico 2015/2016

Abstract

I Genetic Algorithms (GAs) sono una metaeuristica di ricerca appartenenti alla classe degli Evolutionary Algorithms (EAs). Si sono dimostrati efficaci nel risolvere tanti problemi in svariati campi. Tuttavia, le difficoltà nello scalare spesso evitano che i GAs possano trovare una collocazione efficace per la risoluzione di problemi del mondo reale. Quindi, l'obiettivo di fornire soluzioni basate altamente scalabili, assieme alla riduzione dei costi di architetture parallele, motivano la ricerca sui Parallel Genetic Algorithms (PGAs).

Il cloud computing potrebbe essere una valida opzione per la parallelizzazione, dato che non c'è necessità di possedere hardware fisico che può, invece, essere acquistato dai cloud provider, per il tempo desiderato, quantità e qualità. Esistono differenti tecnologie e approcci cloud impiegabili a tal proposito ma, tutti, introducono overhead di computazione. Quindi, ci si può chiedere se, e possibilmente quando, approcci specifici, ambienti e modelli mostrino migliori performance rispetto alle versioni sequenziali, in termini di tempo di esecuzione e uso di risorse.

Questa tesi indaga se, e quando, i GAs possono scalare nel cloud utilizzando approcci specifici. Prima di tutto, Hadoop MapReduce è sfruttato per modellare e sviluppare un framework open source, i.e., elephant56, che riduce l'effort nello sviluppo e velocizza i GAs usando tre diversi modelli paralleli. Le performance del framework sono poi valutate attraverso uno studio empirico. Successivamente, i software container e le message queue sono impiegati per sviluppare, distribuire e eseguire PGAs e il sistema ideato valutato, attraverso uno studio empirico, su un cloud provider commerciale. Infine, le tecnologie cloud sono esplorate per la parallelizzazione di altri EAs, ideando e sviluppando cCube, un'architettura a microservizi collaborativa per risolvere problemi di machine learning.