Developing Genetic Algorithms using Different MapReduce Frameworks: MPI vs. Hadoop*

*Note: The full contents of this paper have been published in the volume Lecture Notes in Artificial Intelligence 11160 (LNAI 11160)

Carolina Salto  
Facultad de Ingeniería  
Universidad Nacional de La Pampa  
Argentina  
minettig@ing.unlpam.edu.ar

Gabriela Minetti  
CONICET  
Argentina  
saltoc@ing.unlpam.edu.ar

Enrique Alba, Gabriel Luque  
Departamento de Lenguajes  
Universidad de Málaga  
Málaga, Spain  
{eat,gabriel}@lcc.uma.es

Abstract—MapReduce is a quite popular paradigm, which allows to no specialized users to use large parallel computational platforms in a transparent way. Hadoop is the most used implementation of this paradigm, and in fact, for a large amount of users the word Hadoop and MapReduce are interchangeable. But, there are other frameworks that implement this programming paradigm, such as MapReduce-MPI. Since, optimization techniques can be greatly beneficiary of this kind of data-intensive computing modeling, in this paper, we analyze the performance effect of developing genetic algorithms (GA) using different frameworks of MapReduce (MRGA). In particular, we implement MRGA using Hadoop and MR-MPI frameworks. We analyze and compare both implementations considering relevant aspects such as efficiency and scalability to solve a large dimension problem. The results show a similar efficiency level between the algorithms but Hadoop presents a better scalability.

Index Terms—