

Effect of population size in heterogeneous and homogeneous machines in a distributed EA

P. García-Sánchez · J. González · C. Fernandes · M. G. Arenas · A. M. Mora · P. A. Castillo · J. J. Merelo

Received: date / Accepted: date

Abstract This paper shows a preliminary study about population size tuning in an distributed Genetic Algorithm. This adaptation is done taking into account the computational power of each node of an heterogeneous cluster. Same parameters are also tested in an homogeneous scientific cluster. Two problems with distinct characteristics have been tested: the linearly-solvable OneMax problem and the deceptive and multimodal MMDP functions. Results show that setting this parameter according computational power decreases the time to obtain the optimum in both problems in heterogeneous clusters.

Keywords Evolutionary Algorithms · Service Oriented Architecture · Service Oriented Science · Web Services · Interoperability · Distributed computing

1 Introduction

New trends such as Cloud Computing Buyya et al (2009), GRID Altunay et al (2011) or Service Oriented Science Foster (2005) are leading to heterogeneous computational devices working at the same time. Moreover, many laboratories do not count with classic clusters, but the usual workstations used by scientists can behave in group as a heterogeneous cluster. Distributed Evolutionary Algorithms (dEAs) ? have been tested successfully in these systems.

Heterogeneous dEAs can be divided in two categories: a dEA where the parameters are different in

each island (heterogeneous parameters) or the same algorithm in heterogeneous hardware. It also have been proved that this kind of algorithms are even more efficient in heterogeneous hardware configurations, than in homogeneous devices Alba et al (2002). This can be explained by different reasons, such as different memory access times, cache, or even implementation languages or compilers in each machine, leading to different exploitation rate of the search space. The heterogeneous parameters configuration also have been proved as more efficient than a fixed set of parameters for different problems Gong and Fukunaga (2011). Our motivation in this work is to combine both ideas adapting the population size of the islands to the heterogeneous hardware. To calculate the computational power, the algorithm is executed in each machine, and then, the total size of individuals is distributed according the number of generations attained in each node during the same time. Two different problems (MMDP and OneMax) have been used as a benchmark.

In this work, a distributed system has been developed to solve the following questions:

- Can a distributed EA be adapted to take the most of the capability of a heterogeneous cluster?
- Does the proposed population size adaptation to the computational power scheme any effect in both systems?
- Is there any difference in homogeneous and heterogeneous cluster?

The rest of the work is structured as follows: after the state of the art, we present the developed algorithms and experimental setting. Then, the results of the experiments are shown (Section 5), followed by conclusions and suggestions for future work.

P. García-Sánchez
Dept. of Computer Architecture and Computer Technology,
E.T.S. Ing. Informática y Telecomunicación and CITIC-UGR
University of Granada, Granada, Spain
E-mail: pgarcia@atc.ugr.es

2 State of the art

In the field of the Evolutionary Computing there are two different approaches about the algorithm parameter setting: parameter control and parameter tuning Eiben and Smit (2011). The first one refers to set a number of parameters of an EA and to change these parameter while the algorithm is running. The parameter tuning consist in establish a good set of parameters before the run (and do not change them during runtime).

Computational performance of nodes or network speeds can also be inherents parameter of an algorithm. In Alba et al (2002) authors compared a distributed GA in homogeneous and heterogeneous clusters. Super-linear performance was obtained in the heterogeneous clusters, being more efficient that the same algorithm in homogeneous ones. Some authors have expanded this idea adapting the algorithm to be executed: in Domínguez and Alba (2013) authors presented a distributed hybrid meta-heuristic that combines GAs and Simulated Annealing (SA). Their system executes the costly algorithms (GAs) in faster nodes, and simpler meta-heuristics (SA) in slower nodes, obtaining better results than other configurations. In Gong et al (2005) different configurations of heterogeneous machines for a tree topology were studied. However, the heterogeneity was simulated in an homogeneous clusters with programs to add computational load. Load-balancing was also applied taking into account the computational load of the nodes in Garamendi and Bosque (2006): a small benchmark was executed in all nodes at the beginning of the algorithm to distribute individuals of an Evolutionary Strategy (ES). However, there was not communication between the nodes. In the area of heterogeneous parameters, but homogeneous hardware, setting each island a random set of parameters can also increase the performance of a distributed Genetic Algorithm (dGA), as explained in Gong and Fukunaga (2011). That model outperformed a tuned canonical dGA with the same parameters in all islands. Finally, adapting the migration period has produced better results than homogeneous periods in homogeneous clusters, as explained by Salto and Alba (2012).

Our work presents a combination of previous ideas, where a parameter tuning given by the computational cost of the machines is performed. For our knowledge, there are not works that modify parameters of the GA depending of the node where the island is being executed.

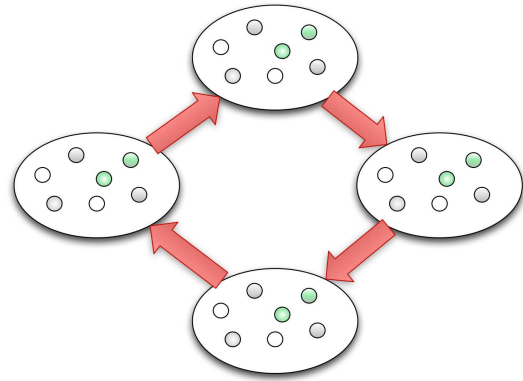


Fig. 1 Island model using a ring topology.

3 Service Oriented Evolutionary Algorithms

The EAs are a general technique for solving optimization and search problems based in the evolution of species and natural selection. These algorithms are formed by a population of possible solutions (called *individuals*) that competes using their *fitness* (quality of adaptation) with the rest of solutions. In each iteration of the algorithm (or *generation*) the individuals are evolved by means of selection and recombination/mutation to create a new set of candidates, until a *stop criterion* (i.e. number of generations) is met. Fitness function is a quality function that gives the grade of adaptation of an individual respect the others. This function is usually the problem to solve.

There are different ways to parallelize the EAs:

- *Farming model (centralized)*: A central node coordinate several slave nodes. The central node executes the EA in a sequential way, but distributes the individuals of the population to the slaves just for being evaluated.
- *Island model (distributed)*: A number of nodes executes simultaneously the EA, working with different sub-populations at the same time. Each certain number of generations is interchanged (migrated) between populations.
- *Cellular EAs (fine grain)*: Each node has one individual of the population, and selection and reproduction is limited with the individuals of the neighbourhood of the node. Usually a bi-dimensional grid is used for topology.

The distributed EAs are very popular because the implementation is not complicated and they explote a coarse grain parallelism with sporadical communications, being appropriated to be executed in distributed architectures such as clusters or GRIDs .

As discussed in ? the evolutionary algorithms research area is a propitious environment to migrate to

```

population ← initializePopulation()
while stop criterion not met do
  parents ← selection(population)
  offspring ← recombination(parents)
  offspring ← mutation(offspring)
  population ← population + offspring
  if time to migrate then
    migrants ← selectMigrants(population)
    remoteBuffer.send(migrants)
  end if
  if localBuffer.size DISTINTO zero then
    population ← population + localBuffer.read()
  end if
  population ← removeWorst(population)
end while

```

Fig. 2 Pseudo-code of the used dEA.

SOA for several reasons: SOA fits with the genericity advantages in the development of software for EAs (?) and adds new features, like language independence and distribution mechanisms. Moreover, there are a large number of frameworks for EAs mostly incompatible due to different programming languages, operating systems or communication protocols (see ? for a survey). Also, new research trends, like self-adaptation (?), require many changes and modifications in the algorithms behavior in real-time. And finally, the increase of technologies such as GRID and Cloud Computing (Buyya et al, 2009), where the computation elements are distributed in different machines, with many operating systems and programming languages.

4 Experimental setup

This section presents the parameters and systems to conduce the experiments.

The algorithm to improve is a dGA. Parameters are described in Table 1. The algorithm is steady-state: the offspring is mixed with the parents and the worst individuals are removed. A ring topology is used, and the best individual is sent after a fixed number of generations of each node (64). Two different parameter configurations have been used: 64 individuals per node (homogeneous size) and a different number of individuals proportional to the number of generations attained in this first homogeneous size execution (heterogeneous size).

Figure 2 shows the pseudo-code of the used algorithm.

The problems to evaluate are the Massively Multimodal Deceptive Problem (MMDP) Goldberg et al (1992) and the OneMax problem Schaffer and Eshelman (1991). Each one requires different actions/abilities by the GA at the level of population sizing, individual selection and building-blocks mixing. The MMDP

Table 1 Parameters used.

Name	Value
Total individuals	256
Population size in HoSi	64
Population size in HeSi	98, 84, 66, and 8
Crossover type	Uniform crossover
Crossover rate	0.5
Mutation rate	1/genome size
Selection	2-tournament
Replacement	Steady-state
Generations to migrate	64
Genome size for MMDP	150
Genome size for OneMax	5000

is designed to be difficult for an EA, due to its multimodality and deceptiveness. Deceptive problems are functions where low-order building-blocks do not combine to form higher order building-blocks. Instead, low-order building-blocks may mislead the search towards local optima, thus challenging search mechanisms. MMDP it is composed of k subproblems of 6 bits each one (s_i). Depending of the number of ones (unitation) s_i takes the values shown in Table 4.

Table 2 Basic deceptive bipolar function (s_i) for MMDP.

Unitation	Subfunction value
0	1.000000
1	0.000000
2	0.360384
3	0.640576
4	0.360384
5	0.000000
6	1.000000

The fitness value is defined as the sum of the s_i subproblems with an optimum of k (equation 1). The search space is composed of 2^{6k} combinations from which there are only 2^k global solutions with 22^k deceptive attractors. Hence, a search method will have to find a global solution out of 2^{5k} additionally to deceptiveness. In this work $k = 25$.

$$f_{MMDP}(\mathbf{s}) = \sum_{i=1}^k fitness_{s_i} \quad (1)$$

OneMax is a simple linear problem that consists in maximising the number of ones in a binary string. That is, maximize the expression:

$$f_{OneMax}(\mathbf{x}) = \sum_{i=1}^N x_i \quad (2)$$

To test the algorithm two different computational systems have been used: an *heterogeneous cluster* and an *homogeneous cluster*. The first one is formed by 4 different computers of our lab with different processors,

operating systems and memory size. The latter is a dedicated scientific cluster formed by homogeneous nodes. Table 4 shows the features of each system.

Because the operating system and architecture heterogeneity the ANONYMOUS framework Nonymous (1), based in Java, has been used. This is a service-oriented evolutionary framework that automatically configures services to use and be used in a local network. In this case, each node offers a migration buffer to accept foreign individuals. Also, to avoid bottlenecks in distributed executions, asynchronous communication has been provided to avoid idle time using reception buffers (that is, the algorithm does not wait until new individuals arrive). This kind of communication offers excellent performance when working with different nodes and operating systems, as demonstrated by Alba et al (2002). The transmission mechanism is based in ECF Generic server (over TCP)¹. The source code of the algorithms used in this work is available in <http://anonymous> under a GPL V3 License.

Each different configuration has been tested 30 times. Acronyms for each configuration are HoSi (homogeneous population size), HeSi (heterogeneous population size), HoHa (homogeneous hardware) and HeHa (heterogeneous hardware).

The number of individuals in each node of the HeSi configuration is proportional to the average number of generations obtained in the nodes of the HoSi/HeHa configuration for the MMDP problem. Thus, the HeSi configuration uses 98, 84, 66, and 8 individuals (from N1 to N4). Note that, having two nodes with the same processors and memory (N1 and N2), they have different computational power.

5 Results

The objectives of the parallel programming are to tackle large computational problems, increase the performance of algorithms in a finite time, or reduce time. In this work we focus in the last objective. As claimed by Alba and Luque (2006) the number of evaluations can be misleading in the parallel algorithms area. In our case, for example, the evaluation time is different in each node of the heterogeneous cluster, and the real algorithm speed could not be reflected correctly. However, the number of evaluations has been added in this section to better understand the results. Also, the total number of generations, and the maximum number of generations of the slower of the four nodes are shown. It is difficult to compare between the HoHa and HeHa for the same

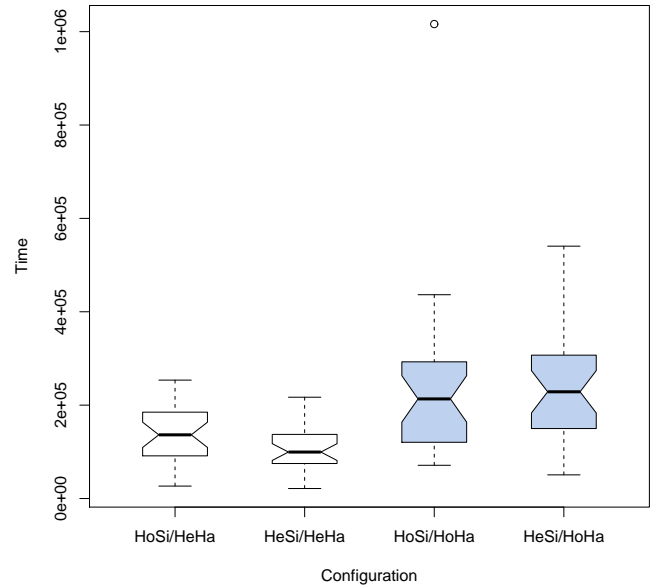


Fig. 3 Time to obtain the optimum in the MMDP problem (milliseconds). White is the heterogeneous cluster and gray the homogeneous one.

reasons: the evaluation time is different in each system (and also in each node).

5.1 MMDP Problem

Table 4 shows the results for the MMDP problem. These results are also shown in the boxplots of the Figure 4 (evaluations) and Figure 3 (time). Table 6 shows the statistical significance of the results. First, a Kolmogorov-Smirnov test is performed to asset the normality of the distributions. If the results fit a normal distribution, then a Student's T-Test is calculated. Otherwise, the non-parametric test Wilcoxon signed rank is applied (see Derrac et al (2011) for a tutorial for comparing EAs).

In the HeHa system, adapting the population to the computational power of each node makes the algorithm to end significantly faster, but with the same number of evaluations (with no statistical significance). This can be explained because the evaluation time is different in all nodes. On the other hand, in the HoHa system, setting the same population sizes makes no difference in time and evaluations, that is, changing this parameter does not influence the performance of the algorithm.

To see the difference of how the evolution is being performed, the average fitness in each node of HeHa is shown in Figures 5 and 6. As can be seen, with the HeSi (Figure 6), the local optima are overtaken in less time

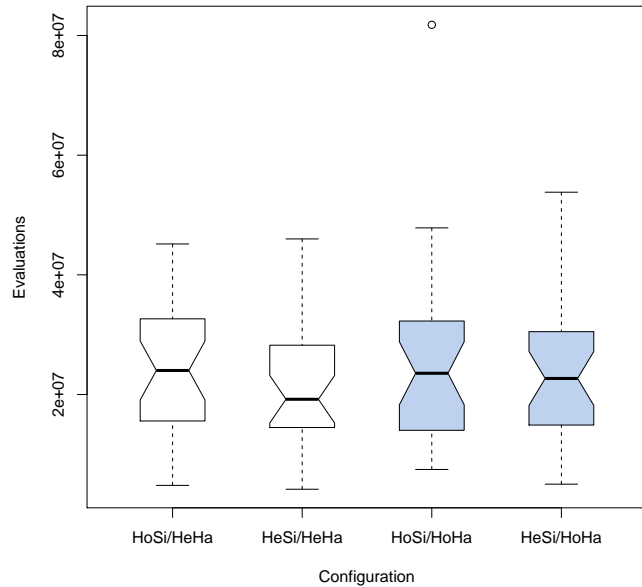
¹ <http://www.eclipse.org/ecf/>

Table 3 Details of the clusters used.

Name	Processor	Memory	Operating System	Network
Homogeneous cluster				
Cluster node	Intel(R) Xeon(R) CPU E5320 @ 1.86GHz	4GB	CentOS 6.7	Gigabit Ethernet
Heterogeneous cluster				
N1	Intel(R) Core(TM)2 Quad CPU Q6600 @ 2.40GHz	4GB	Ubuntu 11.10 (64 bits)	Gigabit Ethernet
N2	Intel(R) Core(TM)2 Quad CPU Q6600 @ 2.40GHz	4GB	Ubuntu 11.04 (64 bits)	Gigabit Ethernet
N3	AMD Phenom(tm) 9950 Quad-Core Processor @ 1.30Ghz	3GB	Ubuntu 10.10 (32 bits)	100MB Ethernet
N4	Intel (R) Pentium 3 @ 800MHz	768 MB	Ubuntu 10.10 (32 bits)	10MB Ethernet

Table 4 Results for the MMDP problem.

Configuration	Max. generations	Total generations	Total evaluations	Time (ms)
HoSi/HeHa	146401,48 \pm 65699,69	380967,25 \pm 168568,84	24382416,51 \pm 10788405,87	136914,03 \pm 60028,48
HeSi/HeHa	96051,5 \pm 45110,90	289282,3 \pm 135038,10	21784528,66 \pm 10161989,38	109875,76 \pm 49185,51
HoSi/HoHa	107334,46 \pm 78167,19	393119,86 \pm 241835,27	25273201,06 \pm 15386663,12	237759,43 \pm 178709,86
HeSi/HoHa	149732,6 \pm 81983,74	438171,16 \pm 240169,19	24430043,46 \pm 13395037,34	245776,93 \pm 134715,52

**Fig. 4** Number of evaluations for MMDP problem. White is the heterogeneous cluster and gray the homogeneous one.

than HoSi (Figure 5). This can be explained because in HeSi, the migration from N4 to N1 is performed faster, adding more heterogeneity to the whole system. White gaps in the figures are the time where the nodes are sending the individual to other nodes (not while they are receiving them). In the HoHa systems, the populations are evolved at the same time, being the average fitness similar in all nodes during all run.

5.2 OneMax Problem

Results for this problem are shown in Table 5 and Figures 7 and 8. In this problem, adapting the popula-

tion sizes decreases significantly the time for solving in the heterogeneous cluster, and, as before, the number of evaluations remains the same (see statistical significance in Table 6). In the homogeneous system, the effect of changing the population sizes is more evident, and this time the evaluations (and therefore, the time) are reduced (both significantly).

The efficiency on OneMax problems depends more on the ability to mix the building-blocks, and less on the genetic diversity and size of the population (as with MMDP). No genetic diversity is particularly required. When properly tuned, a simple Genetic Algorithm is able to solve OneMax in linear time. Sometimes, problems like OneMax are used as control functions, in order to check if very efficient algorithms on hard functions fail on easier functions. As can be seen in Figure 9, the HoSi/HeHa, the average fitness of all populations are increasing in linear way. However, the lower processor evaluates extremely less times. On the other side, in Figure 10, the adaptation of the population size makes that lower processors increase the number of evaluations, but the average fitness is also maintained in linear way (and in smaller increase rate) between migrations. However, the other processors are still spending more number of evaluations. That is the reason why the number of evaluations is higher in HeHa, and lower in HoHa. Computational time is more efficiently used in faster nodes, having more chance to mix the individuals. Also, because the larger size of the individuals in the OneMax problem (5000 bits vs. 150), the transmission time is larger, (white gaps in the figures). That also implies for N4 send their best individual to N1 in a extremely large time when using HoSi (each 64 generations).

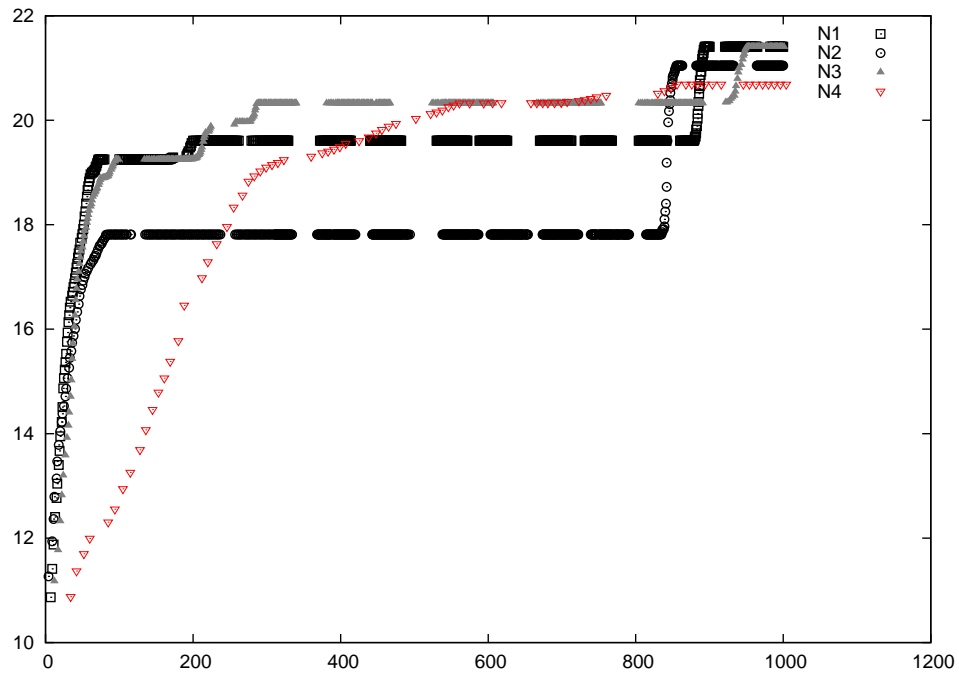


Fig. 5 Average fitness in the first 1000 milliseconds of execution of the four nodes of the heterogeneous cluster with the same population sizes (HoSi/HeHa) for the MMDP problem.

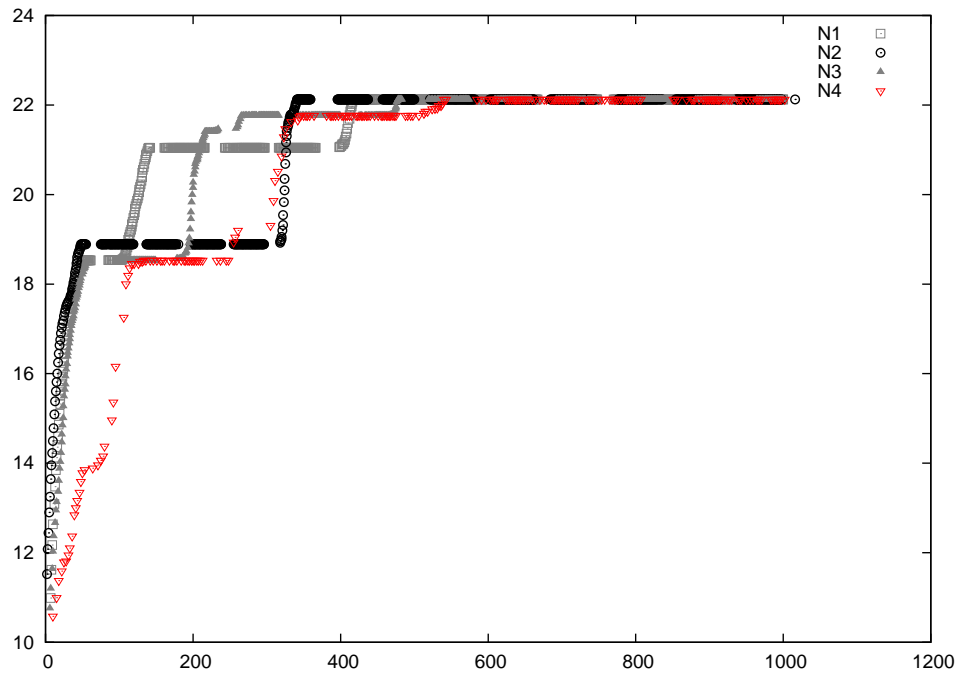


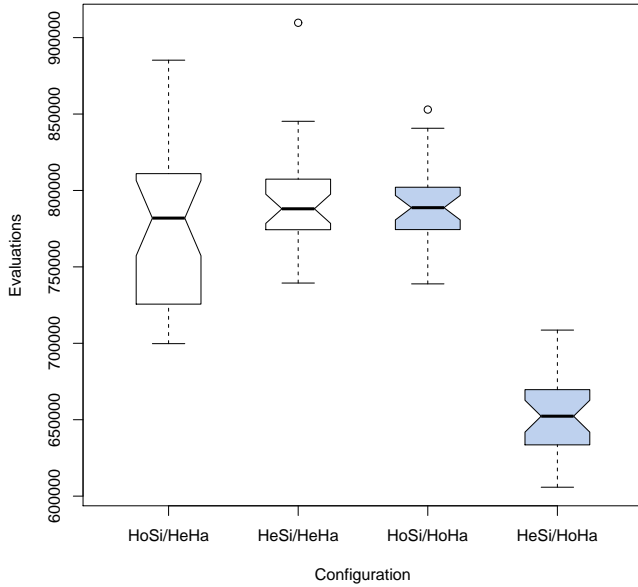
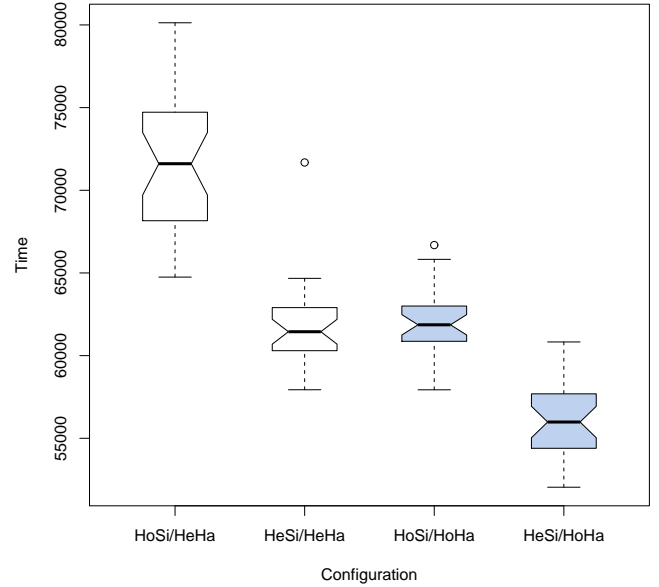
Fig. 6 Average fitness in the first 1000 milliseconds of execution of the four nodes of the heterogeneous cluster with different population sizes (HeSi/HeHa) for the MMDP problem.

Table 5 Results for the OneMax problem.

Configuration	Max. generations	Total generations	Total evaluations	Time (ms)
HoSi/HeHa	4739,41 ± 305,32	12081,51 ± 776,35	773729,03 ± 49686,72	72152,32 ± 4994,71
HeSi/HeHa	3438,03 ± 149,47	11277,33 ± 471,77	794157,73 ± 31843,10	61870,2 ± 2518,74
HoSi/HoHa	3133,36 ± 101,70	12347,83 ± 394,99	790773,33 ± 25279,52	62105,03 ± 1964,75
HeSi/HoHa	13897,86 ± 625,27	20725,63 ± 929,43	651952,8 ± 29114,54	56120,53 ± 2491,92

Table 6 Statistical significance of the results.

Configuration	Normal	Test applied	P-value	Significant difference?
Time for MMDP				
HoSi/HeHa vs HeSi/HeHa	Yes	T-Test	0.032	Yes
HoSi/HoHa vs HeSi/HoHa	No	Wilcoxon	0.567	No
Evaluations for MMDP				
HoSi/HeHa vs HeSi/HeHa	Yes	T-Test	0.231	No
HoSi/HoHa vs HeSi/HoHa	No	Wilcoxon	0.958	No
Time for OneMax				
HoSi/HeHa vs HeSi/HeHa	Yes	T-Test	9×10^{-15}	Yes
HoSi/HoHa vs HeSi/HoHa	No	Wilcoxon	1×10^{-6}	Yes
Evaluations for OneMax				
HoSi/HeHa vs HeSi/HeHa	No	Wilcoxon	0.14	No
HoSi/HoHa vs HeSi/HoHa	Yes	T-Test	2×10^{-27}	Yes

**Fig. 7** Number of evaluations for OneMax problem. White is the heterogeneous cluster and gray the homogeneous one.**Fig. 8** Time to obtain the optimum in the OneMax problem (milliseconds). White is the heterogeneous cluster and gray the homogeneous one.

6 Conclusions

New trends, such as Cloud Computing or Service Oriented Architecture are providing a massively amount of heterogeneous computational devices. This work shows a preliminary study about adapting the population size of an EA to computational power of different nodes in

an heterogeneous cluster. Results show that adapting the population size decrease the execution time significantly in heterogeneous clusters, while changing this parameter in homogeneous clusters not always performs better. This is a promising start for adapting EAs to the computational power of each execution node.

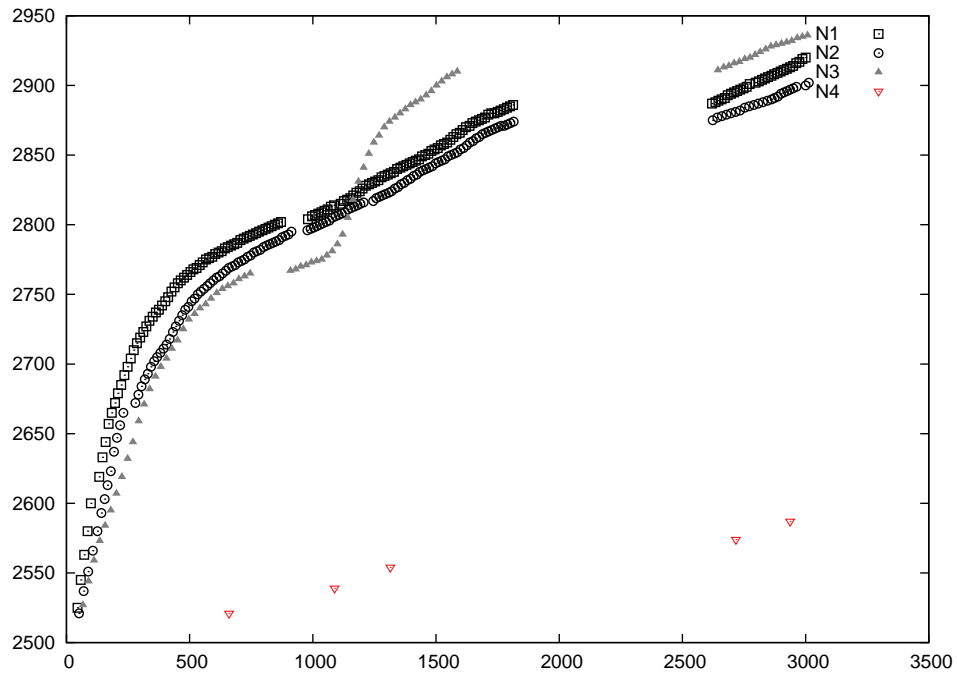


Fig. 9 Average fitness in the first 3000 milliseconds of execution of the four nodes of the heterogeneous cluster with the same population sizes (HoSi/HeHa) for the OneMax problem.

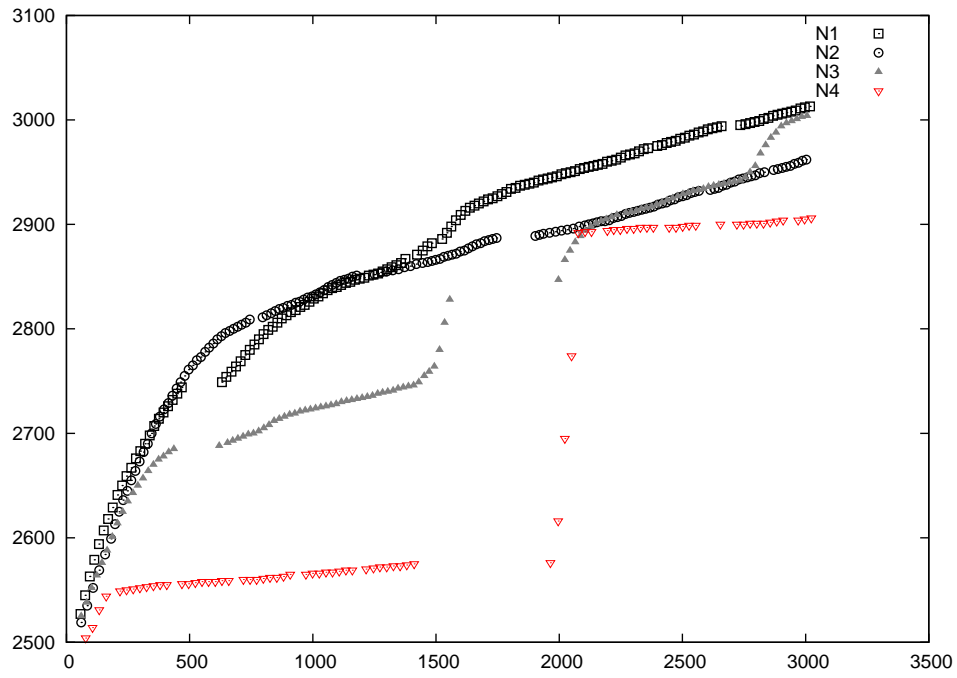


Fig. 10 Average fitness in the first 3000 milliseconds of execution of the four nodes of the heterogeneous cluster with different population sizes (HeSi/HeHa) for the OneMax problem.

In future work a scalability study will be performed, with more computational nodes and larger problem instances. Also, other parameters such as migration rate or crossover probability will be adapted to the execution nodes. This studies will lead to automatic adaptation during runtime, with different nodes entering or exiting in the topology during the algorithm execution or adapting to the current load of the system.

Acknowledgements This work has been supported in part by FPU research grant AP2009-2942 and projects AmIVital (CENIT2007-1010), EvOrq (TIC-3903), and TIN2011-28627-C04-02.

References

- Alba E, Luque G (2006) Evaluation of parallel meta-heuristics. In: Springer (ed) *Parallel Problem Solving from Nature (PPSN)*, LNCS, vol 4193, pp 9–14
- Alba E, Nebro AJ, Troya JM (2002) Heterogeneous computing and parallel genetic algorithms. *Journal of Parallel and Distributed Computing* 62(9):1362 – 1385
- Altunay M, Avery P, Blackburn K, Bockelman B, Ernst M, Fraser D, Quick R, Gardner R, Goasguen S, Levshina T, Livny M, McGee J, Olson D, Pordes R, Potekhin M, Rana A, Roy A, Sehgal C, Sfiligoi I, Wuerthwein F, Open Sci Grid Executive Board (2011) A Science Driven Production Cyberinfrastructure-the Open Science Grid. *Journal of GRID Computing* 9(2, Sp. Iss. SI):201–218, DOI {10.1007/s10723-010-9176-6}
- Buyya R, Yeo CS, Venugopal S, Broberg J, Brandic I (2009) Cloud computing and emerging it platforms: Vision, hype, and reality for delivering computing as the 5th utility. *Future Gener Comput Syst* 25:599–616, DOI 10.1016/j.future.2008.12.001, URL <http://portal.acm.org/citation.cfm?id=1528937.1529211>
- Derrac J, García S, Molina D, Herrera F (2011) A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. *Swarm and Evolutionary Computation* 1(1):3–18
- Domínguez J, Alba E (2013) HydroCM: A hybrid parallel search model for heterogeneous platforms. In: Talbi EG (ed) *Hybrid Metaheuristics*, Studies in Computational Intelligence, vol 434, Springer Berlin Heidelberg, pp 219–235
- Eiben AE, Smit SK (2011) Parameter tuning for configuring and analyzing evolutionary algorithms. *Swarm and Evolutionary Computation* 1(1):19–31
- Foster I (2005) Globus Toolkit version 4: Software for service-oriented systems. In: Jin, H and Reed, D and Jiang, W (ed) *Network and Parallel Computing Proceedings*, Lecture Notes in Computer Science, vol 3779, pp 2–13
- Garamendi J, Bosque J (2006) Parallel implementation of evolutionary strategies on heterogeneous clusters with load balancing. In: *Parallel and Distributed Processing Symposium, 2006. IPDPS 2006. 20th International*, p 8 pp.
- Goldberg DE, Deb K, Horn J (1992) Massive multimodality, deception, and genetic algorithms. In: R Männer, Manderick B (eds) *Parallel Problem Solving from Nature*, 2, Elsevier Science Publishers, B. V., Amsterdam, pp 37–48, URL citeseer.ist.psu.edu/goldberg92massive.html
- Gong Y, Fukunaga A (2011) Distributed island-model genetic algorithms using heterogeneous parameter settings. In: *Proceedings of the IEEE Congress on Evolutionary Computation, CEC 2011, New Orleans, LA, USA, 5-8 June, 2011*, IEEE, pp 820–827
- Gong Y, Nakamura M, Tamaki S (2005) Parallel genetic algorithms on line topology of heterogeneous computing resources. In: *Proceedings of the 2005 conference on Genetic and evolutionary computation*, ACM, New York, NY, USA, GECCO '05, pp 1447–1454
- Nonymous A (1) Anonymous framework. *Secret Journal* 1(1):1–1
- Salto C, Alba E (2012) Designing heterogeneous distributed gas by efficiently self-adapting the migration period. *Applied Intelligence* 36:800–808
- Schaffer J, Eshelman L (1991) On Crossover as an Evolutionary Viable Strategy. In: Belew R, Booker L (eds) *Proceedings of the 4th International Conference on Genetic Algorithms*, Morgan Kaufmann, pp 61–68