Replication Heuristics for Efficient Workflow Execution on Grids*

J. L. Vázquez-Poletti, E. Huedo, R. S. Montero, and I. M. Llorente

Departamento de Arquitectura de Computadores y Automática. Facultad de Informática, Universidad Complutense de Madrid. 28040 Madrid, Spain.

Abstract. Among the different heuristics available for optimizing workflow execution, the replication ones have been previously used in heterogeneous environments with good results. In this work, we analyze its use for workflow scheduling on Grid infrastructures. In particular, we study its applications to an intree workflow, generated by the distribution of the CD-HIT application. The experiments were conducted on a testbed made of resources from two different grids and results show a significant reduction of the workflow execution time.

In a previous paper [1], we considered a Bioinformatics application, *CD-HIT* (Cluster Database at High Identity with Tolerance) [2], for its porting to the Grid using the Grid*W*ay metascheduler [3]. This application performs protein clustering and it can be applied in many activities such as protein family classification, domain analysis, organization of large protein databases or improving database search performance. However, the Grid version of *CD-HIT* didn't provide good performance results, even if it served to bypass memory constrains and so process large data sets. This happened because the nature of the Grid (dynamism, heterogenity and high fault rate).

In this contribution, we apply the *replication* strategy for improving the workflow's efficiency. Supplementary tasks are created for the workflow's *critical path* nodes. When one of these tasks ends, the node is taken as executed and the rest of *replicated* tasks are killed. This way, the more *replicated* tasks are created, the higher is the possibility for that node to be executed shortly by reducing the effect of job failures and queue times.

The input protein database is a compound of UniProt entries and sequence fragments of the Sargasso Sea meta-genome, all of them provided by the National Center for Biotechnology Information (NCBI)[1]. Its size is 1.7GB and it stores

* This research was supported by Consejería de Educación of Comunidad de Madrid, Fondo Europeo de Desarrollo Regional (FEDER) and Fondo Social Europeo (FSE), through BioGridNet Research Program S-0505/TIC/000101, and by Ministerio de Educación y Ciencia, through research grant TIN2006-02806. Also, this work makes use of results produced by the Enabling Grids for E-sciencE project, a project co-funded by the European Commission (under contract number INFSO-RI-031688) through the Sixth Framework Programme. Full information is available at http://www.eu-egee.org/.

[1] http://www.ncbi.nlm.nih.gov/

4,186,284 proteins. Focusing on job execution, input file size and job number depend on the number of divisions made to the starting protein database (32, 40 and 48).

For processing the proposed database, two Grid infrastructures were considered: regional and worldwide. Local and regional machines are nearer to the one where the job submission takes place so they offer less latency. On the other hand, machines pertaining to the Enabling Grids for E-siencE (EGEE) infrastructure are more in number and offer more throughput. But, even with busier machines, the EGEE infrastructure guarantees exclusiveness of CPU use. As coordinated harnessing of these infrastructures was retained necessary for the processing of such a big database, the use of Grid*W*ay was still considered, due to its interoperability capabilities [4]. Tasks were launched from Universidad Complutense de Madrid (UCM), belonging to GRIDIMadrid[2], at different times on different days of the week during April 2007. Finally, the maximum number of tasks submitted to a site was limited to 10.

Experimental results show that using the *replication* technique derived in a valuable speed-up. However, this is speed-up was limited by different factors. Firstly and due to scheduling restrictions, the number of simultaneous running jobs was 20. Then, the algorithm's shape made the level of parallelism to decrease. Finally, the Grid's nature itself derived in reschedules due to suspension timeouts and execution errors.

References

1. Vázquez-Poletti, J.L., Huedo, E., Montero, R.S., Llorente, I.M.: Workflow Management in a Protein Clustering Application. In: Proc. 5th Intl. Work. Biomedical Computations on the Grid (BioGrid'07). 7th IEEE Intl. Symp. Cluster Computing and the Grid (CCGrid 2007), IEEE CS (2007) 679–684
2. Li, W., Godzik, A.: CD-HIT: A Fast Program for Clustering and Comparing Large Sets of Protein or Nucleotide Sequences. Bioinformatics **22** (2006) 1658–1659
3. Huedo, E., Montero, R.S., Llorente, I.M.: A Framework for Adaptive Execution on Grids. Software – Practice and Experience **34** (2004) 631–651
4. Vázquez-Poletti, J.L., Huedo, E., Montero, R.S., Llorente, I.M.: Coordinated Harnessing of the IRISGrid and EGEE Testbeds with GridWay. J. Parallel and Distributed Computing **66** (2006) 763–771

[2] http://www.gridimadrid.org/