Performance Oriented Development and Tuning of GRID Applications

E.Mancini[2], M. Rak[1], R. Torella[1], and U. Villano[2]

[1]Università del Sannio, Facoltà di Ingegneria, C.so Garibaldi 107, 82100 Benevento, Italy
villano@unisannio.it
[2]DII, Seconda Università di Napoli, via Roma 29, 81031 Aversa (CE), Italy
{massimiliano.rak, roberto.torella}@unina2.it

Abstract

The presence of distributed software systems is pervasive in current computing applications. In commercial and business environments, the majority of time-critical applications has moved from mainframe platforms to distributed systems. In the academic and research field, the advances in high-speed networks and improved microprocessor performance have made clusters or networks of workstations and Computational GRIDS an appealing vehicle for cost-effective parallel computing.

However, the systematic use of distributed programming can be frustrating, especially if the final application performance is more than an issue. Even if great effort has been made in developing methodologies and tools that could help the final programmer to develop applications independently of the underlying architecture, as in GRID environments, very few results have been obtained to support prediction and evaluation of prototypal applications.

In the last few years, our research group has been active in the performance analysis and prediction field, developing HeSSE [6-7], a simulator of distributed applications executed in heterogeneous systems, and MetaPL, a prototype-based language based on XML, able to support many different programming paradigm.

This paper presents a simulation-based methodology, based on HeSSE and MetaPL, which makes it possible to predict GRID application and system performance, even when the execution environment is not available and the application is not completely developed. This methodology can be used as the basis for performance-driven GRID application development, or for GRID system performance tuning and design. The paper will show how to model a GRID application and a GRID system using this approach, and how to predict its performance. Finally, the prediction results will be compared to the actual results obtained in the real (i.e., non-simulated) GRID environment, discussing the accuracy of the model used and the effectiveness of the proposed approach.

References

1. I. Foster, C. Kesselman, J. Nick, and S. Tuecke: “The physiology of the grid: An open grid services architecture for distributed systems integration”. Technical report, Open Grid Service Infrastructure WG, Global Grid Forum, June 2002.
2. I. Foster, C. Kesselman, and S. Tuecke. The anatomy of the grid: Enabling scalable virtual organizations. International J. Supercomputer Applications, 15(3), 2001.
3. Buyya, R. , Murshed, M.: "GridSim: A Toolkit for the Modeling and Simulation of Distributed Resource Management and Scheduling for Grid Computing", The Journal of Concurrency and Computation: Practice and Experience (CCPE), Wiley Press, May 2002.
4. Aversa, R., Mazzeo, A., Mazzocca, N., Villano, U.: Developing Applications for Heterogeneous Computing Environments using Simulation: a Case Study. Parallel Computing **24** (1998) 741-761
5. Mazzocca N., Rak M., Villano U. 2000. “The Transition from a PVM Program Simulator to a Heterogeneous System Simulator: The HeSSE Project”. *Recent Advances in Parallel Virtual Machine and Message Passing Interface*, in J. Dongarra et al. (eds.) Lecture Notes in Computer Science, Vol. 1908, Springer-Verlag, Berlin 2000, (pp. 266-273).
6. Mazzocca N., Rak M., Villano U., 2001. “MetaPL: a Notation System for Parallel Program Description and Performance Analysis” *Parallel Computing Technologies*, in Malyshkin. V. (ed.), Lecture Notes in Computer Science, Vol. 2127, Springer-Verlag, Berlin 2001, (pp. 80-93)
7. Labarta, J., Girona, S., Pillet, V., Cortes T., Gregoris, L.: DiP: a Parallel Program Development Environment. Proc. Euro-Par ’96, Lyon, France (Aug. 1996) Vol. II 665-674