Thesis Overview

Dynamic Scheduling on Grid Environments

Bertogna, Mario Leandro
Advisors: Naiouf Marcelo, De Giusti Armando
Universidad Nacional de La Plata, Facultad de Informática, August 2010
mario.l.bertogna@intel.com

The work carried out in the thesis describes efficient virtual environments management. One of the main contributions is a middleware scheduling optimization based on Grid environments [1]. Being the overall goal to get an optimal resource selection plus a coordinate task execution optimization.

In particular, the research was related with the interaction between non trivial quality service and tasks distribution issues in meta-organizations. While linking allocation and policies belonging to virtual and local organizations.

The idea was originated in the virtual and remote laboratories study for virtual spaces creation [2]. In many public and research institutes numerous computing resources are available, but these are not always accessible due to geographical distance, or do not have the ability to achieve a common goal. The concept of virtual space introduces an abstraction layer for managing these resources regardless of location and heterogeneity, thus making efficient use of available resources.

During the thesis development, we had experienced and managed virtual spaces environment generation. An infrastructure has been defined, two types of laboratories were implemented, and a scheduling optimization for parallel applications has been proposed. Today these concepts have evolved, some of the ideas has been published and were implemented in commercial functional prototypes infrastructure, although research is still in the planning of data centers to thousands of computers.

1. Multi-clusters

With regard to multi-clusters one of the first tasks was to assess how affects a Grid environment the network of dedicated clusters. As first step applications were selected and parallel execution experiments were performed to stress Grid environment adaptation [3]. Patterns were found on clusters data transfer interconnected via the Internet, both for connections types and application behavior. Finally, an optimization was performed on the equipment allocation available in the various clusters to achieve effective performance of the overall computing power.

According to experiments performed with different computer configurations and network connections, Grid middleware integration does not alter master-worker parallel applications development. But the parallel application design and deployment ought to be altered, if Grid middleware need to be incorporated. In this thesis the application logic was modified to work for multi-cluster compliant with a service-oriented model, to be more consistent with the efficient use of Grid middleware.

1.1. Meta-Scheduling

In Grid environments, many different resources are intended to work in a coordinated manner, each resource having its own features and complexity. As the number of resources grows, simplifying automation and management is among the most important issues to address. For this purpose architecture was presented with two levels of scheduling: the first one for meta-level organization, with the aim of generating clusters on demand, based on application requirements. The second for local organizations, with the primary goal of get maximum utilization by dynamically manage cluster resources.

After conducting a survey of the state of the art, the contribution was the generalization of the RFC 2753 [4] for Grid environments and the CSF meta-scheduler extension for virtual spaces generation. Through experimentation have been verified the benefits of using virtual machines for maximum clusters utilization

regarding local resource manager. The use of this kind of software components allows changing dynamically execution parameters such as memory and processors number, besides virtual machines migration. These tests led to the conclusion that this technology is feasible to apply in Grid space, where adaptation and environmental control is needed. In this sense, results improve the quality of management and better use of resources.

1.2. Virtual Space

After proposing the architecture and perform certain tests, it was decided to deepen further. A solution was implemented on the topic of scheduling management aspects, and the architecture proposed in the previous chapter was developed [5]. It uses virtual devices to achieve a homogeneous framework, while allowing a dynamic and flexible configuration, within a virtual organization on a Grid environment. The main focus is based on the interconnection of computing resources, using open standards and protocols under strict security.

This section concludes that it is possible to coordinated in a dynamically and flexible way computing resources being geographically dispersed. These tasks can be achieved because of virtual organization schemes provided by the Grid infrastructure and the various options made available by the use of virtual machines.

As contribution of this section:

- Configuration: Regarding configuration has been developed a pseudo object-oriented language for specification logic design and validation.
- Access: For access, a model of access points by the redirection of virtual terminals has been developed via the Internet to the end user.
- Resource Management: For management Grid middleware functionality has been used to allow information execution and transfer between different administrative domains without local administrator's involvement.

Network virtualization overloading should not be underestimated while providing this kind of solutions. In the case where performance requirements are not specific, and the creation of a virtual working environment isolated from the physical one is the goal. This solution is feasible, allowing efficient use of available resources. It also worth to consider the fact that the parallel application would not suffer any change. But it should incur the cost of being adapted to the environment.

1.3. Scheduling Algorithms

Grid environments schedulers must consider aspects such as resource heterogeneity, dynamic adaptation to the tasks and high costs of communication, so that traditional schedulers are often not an option. During the study of such schedulers, Grid algorithms have presented special features, such as the ability to work in heterogeneous and dynamic environments. According to this taxonomy, the proposed solution structure type is distributed-cooperative, due to shared decision making by meta and local organization members. Resources are the key factor in the objective function, noting that they efficiently work during execution.

As a contribution for this section [6], an analytical proposal was described concerning a *cuasi-optimal* virtual space based on computing environment information. Finding the balance between effectiveness and speed was selected Hill-Climbing algorithm with *k-new* beginnings. In this case, with a substantial modification to the classical algorithm, where *k* options are not random, but a deduction based on the particular characteristics of the problem using data communication links. This optimization achieved a substantial performance increase over the total execution time of applications because the best choice of all the machines.

1.4. Simulation and Optimization

A parallel application in a multi-cluster is limited by the performance of the machines in the cluster (compute-bound) or the capacity of the network (communication bound). The maximum performance can be achieved by the application is done in a particular cluster if the limit is given by the computational power. In the scheduling algorithm development optimization techniques used to implement the use of Grid multi-clusters are described, incorporating the use of virtual machines. It also proposes a virtual space dynamic adaptation optimization, in order to improve bandwidth utilization and idle resources available.

Online migration features and dynamic scheduling are used in virtual machines management to transform limited network environments to be computed limited environments, achieving high performance in the use of

the equipment available. To create repeatable and fast configurable tests a simulator was developed, as part of thesis work.

Selection algorithms have been implemented to find groups of clusters that satisfy certain requirements in a small search space, making an effort to return a solution in a fast and optimal way. These implementations increase computing power by nearly 20%. Dynamic algorithms have been implemented to adapt cluster configuration at run time with migration of virtual machines. This implementation results in performance improvement of 10% on the total computing time, with machines processing 100% of their working time. These results allow us to conclude that this solution is feasible to be implemented on Grid environments, where automation and self-management are key to attain effective resource usage. Where clusters serve fixed applications, multi-cluster analysis could guide balance tuning between computation and communication, determining whether it is more effective to either increment/decrement the bandwidth in use, or increment/decrement engaged computing power.

1.5. References

[1] I. Foster, C. Kesselman, and S. Tuecke, "The anatomy of the grid: Enabling scalable virtual organizations," *International Jounral of Supercomputer Applications*, vol. 15, no. 3, 2001.

[2] Bertogna, Mario .L., Grosclaude, E., del Castillo, R., Lopez Luro, F. y Zanellato, C. “Arquitectura para Laboratorios Remotos Físicos y Virtual. CACIC 2005 (pp. 317328), 2005.

[3] E. Argollo, D. Rexachs and E. Luque, “Efficient Execution of Scientific Computation on Long-distance Geographically Distributed Clusters”, PARA'04 State-of-the-Art in Scientific Computing June 20-23, 2004.

[4] Yavatkar, R., Pendarakis, D. and R. Guerin, "A Framework for Policy-based Admission Control", RFC 2753, January 2000.

[5] Bertogna, Mario .L., Grosclaude, E., Francisco Lopez Luro.” Grid Virtual Laboratory Architecture”, VHPC. 2007.

[6] Bertogna, Mario .L., Grosclaude, E., Marcelo Naiouf, Armando De Giusti, Emilio Luque. “Dynamic on Demand Virtual Clusters in Grid”. VHPC. 2008.