Chapter XVII

Pricing Computational Resources in Grid Economies

Kurt Vanmechelen
University of Antwerp, Belgium

Jan Broeckhove
University of Antwerp, Belgium

Wim Depoorter
University of Antwerp, Belgium

Khalid Abdelkader
University of Antwerp, Belgium

ABSTRACT

As grid computing technology moves further up the adoption curve, the issues of dealing with conflicting user requirements formulated by different users become more prevalent. In addition, the need to negotiate static sharing agreements between the different stakeholders in a grid system is time-consuming and offers limited incentive for resource owners to step into the grid’s infrastructure in a provider role. Resource management approaches that are currently adopted in grids are not able to deal with these issues in a flexible, value-maximizing way because of their system-centric approach. In this contribution, we will present a clear motivation for the use of economic forms of scheduling in grid computing environments to address these shortcomings. We also provide an introductory overview of the different forms of market mechanisms that have been adopted by researchers in the field. In addition, we present simulation results concerning the use of Vickrey auctions and commodity markets as market mechanisms for dynamic pricing in grid resource markets.

INTRODUCTION

Grid technology has reached a maturity level in which computational resources are being virtualized, shared and used on a global scale. Large scale deployments of the technology in production level academic environments are being driven by projects such as EGEE in Europe and TeraGrid in the
United States. These projects now make hundreds of teraflops in compute capacity and petabytes of storage capacity available to researchers in more than 40 countries. One of the crucial components of grid middleware, the software which enables the operation of a grid infrastructure, is the resource management system. This subsystem is responsible for mapping a user’s request for service to resources available in the grid. Often this involves a two staged process in which a grid service, called the resource broker, first determines a set of suitable local resource managers based on the request’s quality of service requirements. In the second stage, the local resource managers are contacted to fulfill the request by engaging the local resources they control (e.g. the set of CPUs in the cluster).

Currently, the resource management and scheduling approaches found in production level grid middleware have a strong focus on the efficient scheduling of jobs from a system-oriented point of view. This means that the broker schedules jobs in a way that maximizes the overall utilization of the infrastructure or that obtains the highest possible level of overall system throughput. These approaches to scheduling do not take into account the actual value grid users associate with their computations. As a result, the broker may be performing optimally from a system-oriented point of view, but suboptimally from a user-oriented point of view. The potential for value loss as a result of these suboptimal scheduling decisions increases with the load on the grid infrastructure and the heterogeneity of the valuations among users. As shown by Chun (2005), user valuations for allocations of motes on a sensor net testbed can vary by four orders of magnitude.

In addition, current resource management systems assume the existence of static bilateral sharing agreements between the different parties that contribute to the shared grid infrastructure. These agreements are often hard to negotiate, and result in limited flexibility and openness with respect to the integration of new parties in the global grid infrastructure. As a result, several grid ‘islands’ or ‘silos’ emerge and the full potential of the infrastructure is not realized. This is partly due to the specific funding contracts under which grid resources are procured and brought into operation. Indeed, a currently debated topic is how to build grid infrastructures that are sustainable over the long term, i.e. that are self-supporting and do not require funding from a source that imposes the vision of global sharing of resources for their operation.

We believe that the introduction of an electronic market place for trading grid usage rights provides a promising approach to deal with the issues of openness, sustainability, and value maximization. Such a market place is open to all parties that wish to participate in the system as a provider or consumer. The incentives for (well-behaved) participation in such a market stem from a common value model and accounting system that charges or rewards the different parties involved, for their requested or delivered service.

Firstly, this differs from the negotiation model that is currently in place in which users need to lobby for access rights which are subsequently enforced by long term, static policies. A market-based approach allows for greater flexibility in terms of accepting new parties in the infrastructure and in terms of determining usage rights for users, resulting in a more open and agile grid infrastructure. This is especially important for the potential grid user base that is formed by small research institutions or SMEs. Their relative cost for maintaining and procuring computing infrastructure is high due to limited economies of scale and their often strongly fluctuating requirements. As a consequence, these users are not able to engage in bartering during negotiations of sharing agreements as they do not own a significant amount of computing infrastructure. On the other hand, this user base has the greatest potential for benefiting from the added value a grid infrastructure can provide. The possibility to pay for usage rights enables them to tap into this added value.
Related Content

A Complementary Approach to Grid and Cloud Distributed Computing Paradigms
[www.igi-global.com/chapter/complementary-approach-grid-cloud-distributed/64574?camid=4v1a](www.igi-global.com/chapter/complementary-approach-grid-cloud-distributed/64574?camid=4v1a)

Routing in Coloured Sparse Optical Tori by Using Balanced WDM and Network Sparseness
[www.igi-global.com/article/routing-coloured-sparse-optical-tori/70769?camid=4v1a](www.igi-global.com/article/routing-coloured-sparse-optical-tori/70769?camid=4v1a)

Power Optimization Using Clock Gating and Power Gating: A Review
[www.igi-global.com/chapter/power-optimization-using-clock-gating-and-power-gating/159037?camid=4v1a](www.igi-global.com/chapter/power-optimization-using-clock-gating-and-power-gating/159037?camid=4v1a)

Grid Technology: E-Learning in Telemedicine and Organizational Collaboration
[www.igi-global.com/chapter/grid-technology-learning-telemedicine-organizational/45557?camid=4v1a](www.igi-global.com/chapter/grid-technology-learning-telemedicine-organizational/45557?camid=4v1a)