Available online at www.sciencedirect.com
SCIENCE DIRECT
J. Parallel Distrib. Comput. 65 (2005) 1366–1383
Journal of Parallel and Distributed Computing
www.elsevier.com/locate/jpdc

A pricing strategy for job allocation in mobile grids using a non-cooperative bargaining theory framework

Preetam Ghosh*, Nirmalya Roy, Sajal K. Das, Kalyan Basu

Center for Research in Wireless Mobility and Networking (CReWMaN), Department of Computer Science and Engineering, The University of Texas at Arlington, Arlington, TX 76019-0015, USA

Received 5 January 2004; received in revised form 5 November 2004; accepted 11 May 2005
Available online 11 July 2005

Abstract

Due to their inherent limitations in computational and battery power, storage and available bandwidth, mobile devices have not yet been widely integrated into grid computing platforms. However, millions of laptops, PDAs and other portable devices remain unused most of the time, and this huge repository of resources can be potentially utilized, leading to what is called a mobile grid environment. In this paper, we propose a game theoretic pricing strategy for efficient job allocation in mobile grids. By drawing upon the Nash bargaining solution, we show how to derive a unified framework for addressing such issues as network efficiency, fairness, utility maximization, and pricing. In particular, we characterize a two-player, non-cooperative, alternating-offer bargaining game between the Wireless Access Point Server and the mobile devices to determine a fair pricing strategy which is then used to effectively allocate jobs to the mobile devices with a goal to maximize the revenue for the grid users. Simulation results show that the proposed job allocation strategy is comparable to other task allocation schemes in terms of the overall system response time.

Keywords: Grid computing; Mobile devices; Bargaining theory; Non-cooperative games; Load balancing

1. Introduction

Grid computing provides a distributed computing infrastructure for solving large-scale advanced scientific and engineering problems through sharing of resources, usually over high-speed communication networks [8,9]. Computational grids typically have a conglomeration of various resources with different owners at geographically different sites. Several Grid systems including Globus [7] have addressed many of these issues with the exception of resource trading and quality of service (QoS)-based scheduling. The GRACE [2] architecture leverages existing technologies such as Globus, and provides new services that are essential for resource trading and aggregation, depending on their availability, capability, cost, and users' QoS requirements. An important issue of such grid computing systems is the efficient assignment of jobs and utilization of resources of unused devices, commonly referred to as the *load balancing* or *job scheduling* problem. This problem is often formulated in the context of a *system model*, an abstraction of the underlying resources, that provides information to the job allocator regarding the availability and properties of resources at any point in time. The job allocator then allocates jobs to the available resources and attempts to optimize specified performance metrics, such as time deadline or revenue maximization.

Given that millions of laptops, PDAs and other portable devices remain unused most of the time, the grid architecture is recently extended in [21] leading to what is called a *mobile grid* environment. The goal is to potentially utilize the huge repository of resources of mobile devices to provide

* Corresponding author. Fax.: +1 817 272 3784.
E-mail addresses: ghosh@cse.uta.edu (P. Ghosh), niroy@cse.uta.edu (N. Roy), das@cse.uta.edu (S.K. Das), basu@cse.uta.edu (K. Basu)
URLs: http://crewman.uta.edu/~preetam (P. Ghosh), http://crewman.uta.edu/~nirmalya, http://crewman.uta.edu/~das, http://crewman.uta.edu/~basu.

0743-7315/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jpdc.2005.05.013

a seamless source of computational power and storage capacity. However, this concept offers significant challenges mainly due to the inherent limitations in processing, memory, battery power and wireless communications capabilities of mobile devices. In a mobile grid, a more important performance metric is system *throughput* where the resources are *distributively owned*. In this environment, a resource owner has the right to define a very sophisticated *usage policy*, e.g., a job can run on a mobile device only if it generates a certain minimum revenue. Distributed ownership requires a scheduling paradigm that can operate in an environment where *resource owners* (i.e., mobile devices) and *resource users* (i.e., wireless access points servers) dynamically define their own policies and models.

Job scheduling in mobile grid computing thus demands for a decentralized algorithm with a robust system model. We also need to consider an *economic pricing model* that will govern the cost benefits of mobile device owners to allow complex computational jobs to be performed at those devices. Due to the conflict of interest between the players, namely the mobile device and the wireless access point server (WAPS), this pricing model can be more realistically formulated using a *non-cooperative bargaining theory* [20] framework.

Although Game Theoretic approaches have been proposed to develop economic models for resource management and scheduling in grid computing [3], they suffer from precise lack of formulation in the sense that the actual mapping of the problem into a game between two players has not been shown, nor are stated analytical modeling and results. Also, mobile grid computing is a completely new paradigm for which only a very crude economic model has been specified in [21]. We envision that potentially there are many mobile devices distributed in the network, which will be competing to share the jobs originated by the grid community. There arise several challenging issues such as:

(1) efficient job allocation to different mobile devices taking into account various performance requirements;
(2) handling fairness in pricing the job allocation;
(3) the ability to implement the allocation scheme in a distributed manner with minimum communication overheads;
(4) maximizing the network efficiency, i.e., minimizing the response time.

1.1. Related works

In mobile grid environments, the integration of wireless mobile devices to exploit the available processing power introduces new challenges. A proxy-based clustered architecture for mobile grids is proposed in [21]. The pricing and job scheduling policies in mobile grids need to manage resources and application execution depending on the requirements of resource consumers (i.e., WAP Servers) and resource owners (i.e., mobile devices). They also need to continuously adapt to changes in the availability of resources. This introduces a number of challenging issues that need to be addressed; namely, site autonomy, resource allocation or co-allocation, online control and so on. Several grid systems including Globus [7] have addressed many of these issues with the exception of resource trading and QoS-based scheduling. The GRACE framework [2] particularly addresses these two later issues by leveraging existing technologies such as Globus and providing new services that are essential for resource trading and aggregation, depending on their availability, capability, cost, and user QoS requirements. It develops a generic distributed game theoretic architectural framework and strategies for resource trading using different economic models.

Some scheduling mechanisms based on game theoretic negotiation, deployed in existing grid computing systems, is shown in Table 1. However, none of these games attempts to capture the competitiveness among the mobile devices, nor do they aim at maximizing the grid community's revenue. Cooperative game theory has been used to obtain a Nash bargaining framework to address issues like network efficiency, fairness and revenue maximization for bandwidth allocation and pricing in broadband networks [28]. Direct application of a cooperative bargaining theory solution [11] and an optimal scheme based on the overall system response time for load balancing [25] do not consider the pricing constraints of a mobile device. In this paper, we first propose a pricing model and subsequently address the issue of dynamic job allocation such that the grid community's revenue is maximized and also the overall expected job execution time is minimized.

1.2. Our contributions

The main contributions of this paper are two-fold. First, we propose a game theoretic framework to implement the pricing model. The two players, namely the WAP Server (acting on behalf of the grid community) and the mobile device, play an incomplete information alternating-offer, non-cooperative bargaining game [4,6,17] to decide upon the *price per unit resource* charged by that mobile device. The dynamics of interaction is shown in Fig. 1. The concept of incomplete information ensures that the two players have no idea of each other's reserved valuations, i.e., the maximum offered price for WAP Server (acting as the *buyer* of resources) and minimum expected price for mobile device (acting as the *seller* of resources). Assuming there are n mobile devices under a single WAP Server, the WAP Server has to play n such games with the corresponding devices to form the price per unit resource vector, p_i. In particular, by drawing upon the Nash bargaining framework from non-cooperative game theory, the pricing strategy is guaranteed to be fair. Furthermore, we make this pricing scheme stable, so that there would be no incentives for the grid community or the mobile devices to deviate from the mutual