Efficient Bidding for Virtual Machine Instances in Clouds

Sharrukh Zaman
Department of Computer Science
Wayne State University
Detroit, Michigan, USA
Email: sharrukh@wayne.edu

Daniel Grosu
Department of Computer Science
Wayne State University
Detroit, Michigan, USA
Email: dgrosu@cs.wayne.edu

Poster category: Student Project

I. INTRODUCTION

Combinatorial auctions are efficient mechanisms for allocating virtual machine (VM) instances to cloud computing users. Despite the fact that, in general, these mechanisms lead to higher revenues than the currently employed fixed-price mechanisms, the current cloud computing providers do not employ them to allocate their resources. One of the main reasons is the complexity faced by the users when determining the bid (i.e., the bundle of VM instances and the bid value). We address this issue by developing efficient bidding strategies for the users requesting VM instances. We design new metrics for evaluating bundles of VM instances based on the characteristics of the computing tasks. These metrics allow us to determine the valuation of the requested bundles and to design algorithms for selecting the best bundles to bid for.

A. Problem Statement

In a combinatorial auction-based resource allocation mechanism, cloud users participate by bidding for a set of VM instances. It is a challenging task to determine an efficient bid that will maximize the valuation of the requested bundle within a user's budget. In this research, we address the problem of determining efficient bidding strategies for users participating in combinatorial auctions for allocation of cloud resources.

A user of a cloud computing platform can choose a wide range of bundles of VM instances based on the performance she desires. For example, to deploy a web application, a user may require at least two VM instances, one for the web server and another for the application/database server. The performance of the application can be increased by deploying more VM instances or replacing the VM instances with more powerful ones. However, the performance gain also depends on the degree of parallelism of the application. The more parallelism an application has, the more performance gain is possible from increasing the number of VM instances for its implementation. On the other hand, the amount of communication between the parallel components limits the number of VM instances one can use. These factors complicate determining the valuation of a bundle of VM instances for a user.

B. Contributions

In this poster, we present efficient bidding strategies for cloud users. The proposed strategies generate possible bundles of VM instances and find the best bundle to bid for in a combinatorial auction-based resource allocation mechanisms run by the cloud provider [1]. We perform simulation experiments to evaluate the proposed strategies in a simulated cloud environment.

C. Related Work

There have been a lot of research in determining efficient bidding strategies for auctions. The closest work to ours is [2] in which the authors designed efficient bidding strategies for transportation auctions. In their auction, there are only single copies of each given type of auctioned items and only one copy of each item is being allocated. Also, it is assumed that a 'synergy' value is given between pairs of items, which is a crucial element in determining the strategies. The problem we are considering is different from the above setting since we consider several items (VM instances) of few types for allocation. We design new functions to calculate the synergy values based on application characteristics such as degree of parallelism, malleability and communication overhead.

II. PROPOSED BIDDING STRATEGIES

We assume that a cloud provider is running a combinatorial auction-based resource allocation mechanism such as the one presented in [1]. In such mechanisms, k copies of each m types of VM instances are to be allocated and n users are participating. The computing power of a VM instance VM_i is denoted by w_i, which is in this case the number of processors the VM instance has. A user u_j submits a bid $B_j = (r_1^j, \ldots, r_m^j, v_j)$, where r_i^j is the number of VM_i instances requested and v_j is the valuation for the bundle.

Given the number of available resources and the users' bids, the allocation is determined using a combinatorial auction mechanism called CA-GREEDY[1]. The CA-GREEDY mechanism first sorts the bids according to
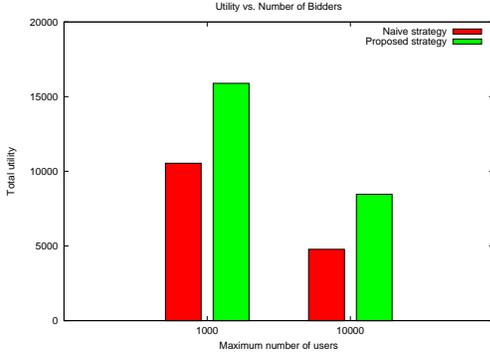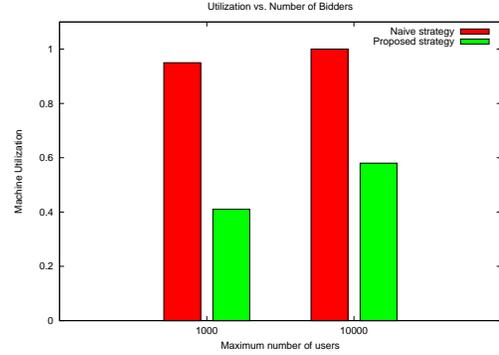
Figure 1. Effect of Bidding on Users' Utility

Figure 2. Effect of Bidding on Utilization

$\sqrt{v_j / \sum_{i=1}^{m} r_i^j w_j}$, which is called the 'density' of bids. Resources are allocated in the order of density, from highest to lowest, until all VM instances are allocated. A user u_j pays 0 if she is not selected for allocation. Otherwise, a user u_l is found, who would get her requested bundle if user u_j were not participating. u_j's payment is calculated by multiplying the density of u_l with $\sqrt{\sum_{i=1}^{m} r_i^j w_j}$.

We create a bidding strategy for a user u_j that has a budget V_j and needs to run an application a_j with workload ω_j. We devise a valuation function $\gamma_j(\omega_j, \mathbf{r^j})$ that computes the speedup and the communication overhead of running the application on the set of VM instances. Using this valuation function, we devise algorithms to compute an efficient bundle of VM instances that will maximize user's utility within the given budget. We consider three types of application depending on their malleability: perfectly malleable, partially malleable, and non-malleable. Malleability refers to the ability of an application to execute on any number of processors.

The proposed algorithms compute the optimal number of processors required to achieve the maximum utility. This number is adjusted if the corresponding valuation exceeds the user's budget. Then, a bundle of VM instances is selected to allocate the number of processors. The valuation function is evaluated to reflect the communication overhead of the application. This bundle is then adjusted if the valuation exceeds the budget. Initially, the bundle includes the largest possible VM instances. If this bundle exceeds the budget in terms of valuation function, we replace one largest VM instance from the bundle with smaller VM instances having an equivalent number of processors.

We prove that the algorithm runs in polynomial time and that it creates efficient bundles. We verify our claim using simulation experiments. Because of the space limitation, we do not give here a formal description of the proposed bidding strategy algorithms.

III. EXPERIMENTAL RESULTS

We evaluate the algorithms through simulation experiments. The goal is to determine the effect of these bidding strategies on the cloud provider's revenue and the utilities of the bidders. In Figures 1 and 2 we present the results of our experiments. These are some primary results of our experiments. In these simulations, 1000 and 10000 users are participating in CA-GREEDY auctions run by the cloud provider, in a period of five simulation days. There are 50 instances of four types of VM available for allocation and users generate only perfectly malleable applications. Each user creates an application with a random workload and a random budget. Then, the users participate in two CA-GREEDY auctions simultaneously, creating bundles employing a naive strategy and the proposed strategy. The results show that employing our proposed strategy generates higher utility for the user compared to the naive strategy. On the other hand, the system utilization deteriorates because of higher competition for the larger VM instances. However, the utilization increases with increased participation. We are currently performing more experiments to investigate the effect of the proposed bidding strategy on system utilization.

IV. POSTER ORGANIZATION

The poster will be organized as follows. We will describe the problem and the contribution. We will present the algorithms for determining the bidding strategies for malleable and partially malleable applications. We will present at least 3-4 graphs summarizing the results of our experiments performed on a simulated cloud computing environment.

REFERENCES

[1] S. Zaman and D. Grosu, "Combinatorial auction-based allocation of virtual machine instances in clouds," in *Proc. of the 2nd IEEE International Conference on Cloud Computing Technology and Science*, 2010.

[2] N. An, W. Elmaghraby, and P. Keskinocak, "Bidding strategies and their impact on revenues in combinatorial auctions," *Journal of Revenue and Pricing Management*, vol. 3, no. 4, pp. 337–357, 2005.