Dynamic Resource Pricing on Federated Clouds

Marian Mihailescu and Yong Meng Teo
Department of Computer Science
National University of Singapore
Computing 1, 13 Computing Drive, Singapore 117417
Email: [marianmi, teoym]@comp.nus.edu.sg

Abstract

Current large distributed systems allow users to share and trade resources. In cloud computing, users purchase different types of resources from one or more resource providers using a fixed pricing scheme. Federated clouds, a topic of recent interest, allows different cloud providers to share resources for increased scalability and reliability. However, users and providers of cloud resources are rational and maximize their own interest when consuming and contributing shared resources. In this paper, we present a dyanmic pricing scheme suitable for rational users requests containing multiple resource types. Using simulations, we compare the efficiency of our proposed strategy-proof dynamic scheme with fixed pricing, and show that user welfare and the percentage of successful requests is increased by using dynamic pricing.

1. Introduction

Modern large distributed systems such as peer-to-peer networks, grids, and more recently clouds, make the sharing and usage of computing resources over the Internet ubiquous. In cloud computing [5], resources are provided over the Internet on-demand, as a service, without the user having knowledge of the underlying infrastructure. Based on the resource type, cloud providers deliver users Infrastructure as a Service (IaaS), Communication as a Service (CaaS), Platform as a Service (PaaS), Software as a Service (SaaS) etc. In this context, the *cloud* refers to both the hardware and the software that enable the service. *Public clouds* are available to all users, while *private clouds* use similar infrastructure to provide services for users within an organization. Currently, several companies including Amazon [1] provide computing and storage services using pay-per-use fixed pricing. However, cloud computing usage is increasing both in breadth, such as the number of resource types and

services offered, and in depth, such as the number of resource providers. Thus, with an increasing number of cloud users, it is expected that more providers will offer similar services. Furthermore, with inter-operability between different providers [3], users will able to use the same service across clouds to improve scalability and reliability. In this context, the aim of *federated clouds*, a topic of recent interest, is to integrate resources from different providers such that access to the resources is transparent to the user while scalability and reliability are increased.

Cloud providers and users are rational, self-interested parties, which exercise their partial or complete autonomy to maximize their benefit [9]. Accordingly, both resource providers and users are not trusted to reveal their truthful resource valuations when allocation resources. Although this is not an issue when using fixed pricing, the performance of dynamic pricing schemes that are not strategy-proof is significantly affected by rational users.

In this paper we discuss a strategic-proof dynamic pricing scheme suitable for allocating resources on federated clouds, where pricing is used to manage rational users. A rational user may represent either an individual user, a group, or an organization, depending on the application context. In federated clouds, users request more than one type of resources from different providers. In contrast to fixed pricing, where users have to manually aggregate resources from different providers, our pricing scheme is designed to allocate a request for multiple resource types. Moreover, in a federated cloud, resource demand and supply fluctuate as users join and leave the system. We show using simulations that using the proposed *dynamic* scheme, the user welfare, the percentage of successful requests, and the percentage of allocated resources increases when compared to fixed pricing.

This paper is structured as follows. Section 2 presents background information and related works from grid computing and distributed systems. We

discuss dynamic pricing for cloud computing and federated clouds together with the proposed pricing scheme in Section 3, while Section 4 evaluates the economic efficiency, the individual user welfare, and the impact of multiple resource types of our dynamic scheme. Finally, Section 5 contains our conclusions and discusses our future work.

2. Background

There are several approaches to allocate shared resources using resource markets, such as bartering, pricing, or bargaining [14]. A *resource market* consists of the environment, rules and mechanisms where resources are exchanged. In this paper, we focus on resource pricing, which enables financial incentives for rational users.

Pricing is the process of computing the exchange value of resources relative to a common form of currency. Economic models for the allocation of shared resources may use fixed or dynamic pricing. When using fixed pricing, each resource type has a predefined price, set by the seller. For example, Amazon provides disk space for $0.15/GB. In contrast, when using dynamic pricing, the resource price is computed for each request according to the pricing mechanism used. Pricing schemes use financial incentives in addition to payments to motivate rational users to be truthful.

Market-based resource allocation mechanisms based on pricing introduce several economic and computational challenges. From a computational perspective, a mechanism must compute in polynomial time the allocation of multiple resource types while maximizing the number of allocated resources and satisfied requests. However, an optimal allocation mechanism for multiple resource types such as combinatorial auctions requires a NP-complete algorithm [10]. Accordingly, many systems share only one resource type, such as CPU cycles in volunteer computing, and file blocks in file-sharing networks.

From an economic perspective, the desirable properties for resource allocation are: *individual rationality*, *incentive compatibility*, *budget balance* and *Pareto efficiency* [8]. In an individual rational allocation mechanism, rational participants gain higher utility by participating in resource sharing than from avoiding it. Incentive compatibility ensures that the dominant strategy for each participant is truth-telling. Budget-balance verifies that the sum of all payments made by buyers equal the total payments received by the sellers. *Pareto efficiency*, the highest economic efficiency, is achieved when, given an allocation, no improvement can be made that makes at least one participant better off, without making any other participant worse off. However, according to the Myerson-Sattherwithe

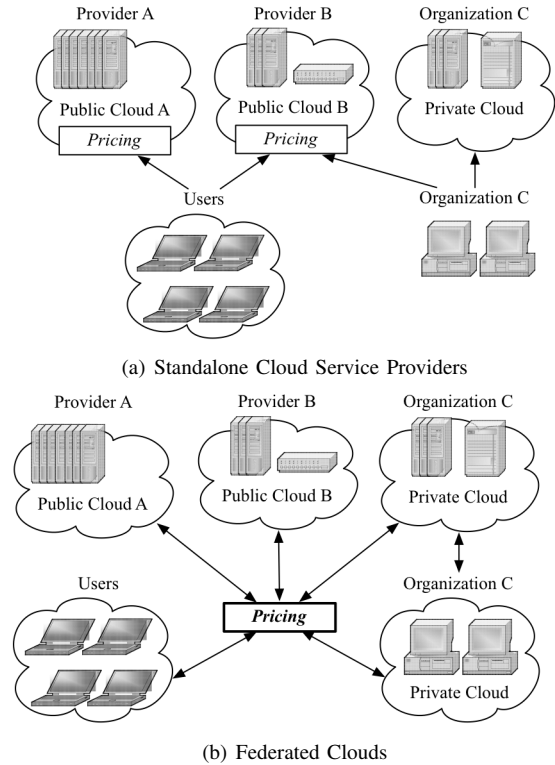(a) Standalone Cloud Service Providers

(b) Federated Clouds

Figure 1: Pricing in Standalone and Federated Clouds

impossibility theorem [8], no mechanism can achieve all four properties together. Accordingly, related works have traded incentive compatibility [6], [13], economic efficiency [7] or budget-balance [10].

In contrast to resource sharing systems used in research and academic communities or for personal benefit, cloud computing has been put into commercial use and its economic model is based on pricing. Federated clouds can be formed by combining private clouds to provide users with resizeable and elastic capacities [4]. Currently, companies such as Amazon operate as standalone clouds service providers (Figure 1.a). However, in a federated cloud (Figure 1.b), any globally distributed user can both offer and use cloud services. A user is either an individual, a group, or an organization, depending on the application context.

3. Proposed Dynamic Pricing Scheme

In a resource market with a large number of providers (sellers) and users (buyers), fixed pricing does not reflect the current market price resource price due to the changing demand and supply. This leads to lower user welfare and to imbalanced markets, e.g. under-demand. Figure 2 shows the welfare lost by a seller that uses fixed pricing. In the case of under-demand, the fixed price tends to be higher than the market price and buyers may look for alternative resources. In the case of over-demand, the fixed price
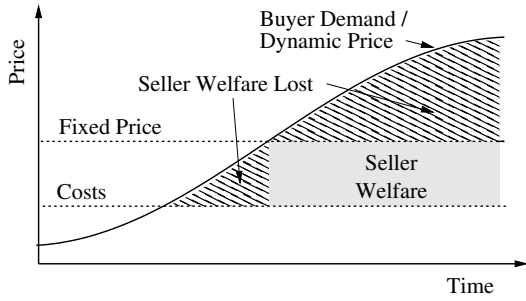
514

Figure 2: Fixed Pricing Limits Seller Welfare

limits the seller welfare, which could be increased by using a higher resource price.

In a federated clouds market, *dynamic pricing* sets resource payments according to the forces of demand and supply. Moreover, the use of dynamic pricing facilitates sellers to provide multiple resource types. Early cloud services such as Sun Grid Compute Utility were restricted to one resource type, e.g. CPU time [3]. More recent services, such as Amazon S3 and EC2, introduced more resource types, i.e. storage and bandwidth. Currently, Amazon has expanded its offer to 10 different virtual machine instance configurations, with different prices for each configuration, and practice tiered pricing for storage and bandwidth [1]. We see this as the first step towards dynamic pricing, where users can request for custom configurations with multiple resource types.

In the context of federated clouds, we propose a strategy-proof dynamic pricing mechanism for allocating shared resources with multiple resource types. We assume a federated cloud resource market where rational users can both provide (sellers) and utilize resources (buyers). Rational users represent either an individual or an organization. Interoperability provides the buyers with uniformity and elasticity. Thus, a buyer request for a large number of resources can be met by more than one seller.

In a previous paper, we propose a mechanism design problem that describes a resource sharing system where rational users can be both buyers and sellers of resources, and a reverse auction-based pricing and allocation scheme, which we prove formally to be individual rational, incentive compatible and budget-balanced [12]. A mechanism that is both individual rational and incentive compatible is known as *strategy-proof*. Auctions are usually carried out by a third party, called the *market-maker*, which collects the bids, selects the winners and computes the payments. Since this paper focuses on the economical and computational advantages of dynamic pricing, we consider for simplicity a centralized market-maker, to which sellers publish resources, and buyers send requests. In order to improve scalability, Section 5 shows our insights

into distributed auctions, where more than one market-makers are able to auction at the same time. Given that buyers and sellers are globally distributed, it is practical to adopt a peer-to-peer approach, where, after pricing and allocation, buyers connect to sellers to use the resources paid for.

4. Impact of Dynamic Pricing

We evaluate the proposed dynamic pricing scheme both for economic and computational efficiency. Using simulation, we compare our pricing scheme with fixed pricing, currently used by many cloud providers. We implement our framework as an application built on top of FreePastry [11], an open-source DHT overlay network environment. FreePastry offers a discrete-event simulator which is able to execute applications without modification of the source code. This allows us both to simulate large systems, and to validate the results in a deployment over PlanetLab [2].

For simplicity, we use a centralized market-maker to compare the efficiencies of the two pricing schemes. A centralized implementation has the advantage of allowing the measurement of economic and computational efficiency with a simple setup for a large simulated network. Moreover, the use of a peer-to-peer substrate such as FreePastry allows us to address the scalability issue in our future work. Thus, our simulated environment contains one market-maker and 10,000 nodes, where each node can be seller and buyer. Publish and request messages are sent to the market-maker node using the FreePastry routing process, which then performs the reverse auctions using the first-come-first-serve policy and computes the payments using the algorithm described in [12].

Economic systems measure efficiency with respect to user's valuations for resources (utility). Consequently, in a Pareto efficient system, where economic efficiency is maximized, a user's utility cannot improve without decreasing the utility of another user. Economic efficiency is a global measure and represents the *total* buyer and seller welfare. More specifically, there are two factors that affect the economic efficiency: *i)* average user welfare; and *ii)* number of successful requests, for buyers, and number of allocated resources, for sellers.

4.1. User Welfare

The *user welfare* is determined by the difference between the user utility and payment. In our proposed scheme, the user utility is the same as the published price, since both buyers and sellers are truthful, according to the incentive compatibility property of our pricing scheme. In the case of fixed pricing, we also consider a truthful buyer, i.e. the published request

price represents the buyer's utility. However, we do not make the same assumption about sellers, which have a fixed resource price that may differ from the seller's utility. Thus, in our experiment, we compare only the average buyer welfare when using fixed and dynamic pricing, respectively.

For the first experiment, we consider a balanced market, where supply and demand are equal. Thus, we assume that the market-maker receives events with an interarrival time of $1s$, where an event has equal probability of being a buyer request or a seller resource publish. Events are uniformly distributed between 10,000 FreePastry nodes, and contain of a number of resource types uniformly distributed between 1 and 3, chosen randomly from a total of 5 resource types. The number of items for each resource type is generated according to an exponential distribution with mean 10. For sellers, we assume 100 as the fixed price, while in the case of dynamic pricing we vary the price by 10%, 20% and 50%, i.e. the price is generated according to a uniform distribution between 90 and 110, 80 and 120, and 50 and 150, respectively. Buyer price is varied according to the same percentage, shown in Figure 3(a) as *Price Variation*. The simulation runs for 600,000 events, which, for an arrival rate of $1s$, give a total simulation time of approximately seven days. To reduce sampling error, we run our experiments three times and compute the average.

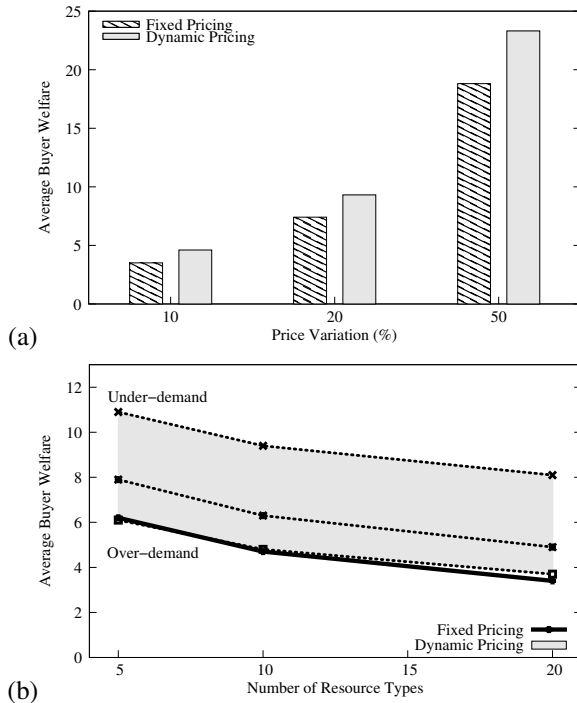The results plotted in Figure 3(a) show that dynamic

(a)

(b)

Figure 3: Dynamic Pricing Increases Buyer Welfare

pricing increases the average buyer welfare. Given that the mean buyer utility is 100, and a *theoretical* maximum welfare for an item is achieved when having the minimum payment, i.e. $100 - Price\ Variation$, we derive that the maximum welfare equals the price variation. Thus, using the proposed dynamic pricing mechanism increases buyer welfare by approximately 10%, when compared to fixed pricing.

Figure 3(b) shows the average buyer welfare when varying the number of resource types to 5, 10 and 20, while the price variation is set to 20%. We consider 3 market conditions: *Under-demand*, when supply is greater than demand, *Balanced market*, when supply equals demand, and *Over-demand*, when supply is less than demand. To simulate different market conditions, we vary the probability of a request event. Thus, in the case of a balanced market, the probability is set to 50%, while for under-demand is 33%, and for over-demand is 66%. For fixed pricing, the average buyer welfare is constant for all market conditions. However, we see that when using a dynamic scheme, the buyer welfare is increased to up to 200%.

4.2. Successful Requests

Next, we study the influence of multiple resource types and different market conditions for the proposed dynamic scheme, and compare its performance with fixed pricing. We vary the number of resource types in a request to 5, 10, and 20, while the price variation is set to 20% and the market condition varies from under-demand, to over-demand. We measure the performance of the pricing scheme given by the percentage of successful buyer requests and the percentage of the allocated seller items.

As shown in Figure 4(a), in the case of fixed pricing, the percentage of successful buyer request is close to 50% for all market conditions, since the buyer item price is uniformly distributed with the mean equal to the seller item price. However, the percentage of successful buyer requests decreases when the number of resource types increases since the number of sellers that are allocated to satisfy a request also increases.

In contrast, when using dynamic pricing, the percentage of successful buyer requests varies under different market conditions, according to the forces of supply and demand. Thus, when supply is greater than demand, the percentage of successful buyer requests is higher than in the case of a balanced market, while for over-demand the percentage decreases further. The proposed pricing scheme achieves a higher percentage of successful buyer requests, shown in Figure 4(a) by the area between under-demand and over-demand.

In Figure 4(b) we show the percentage of allocated seller items as different areas based on the market
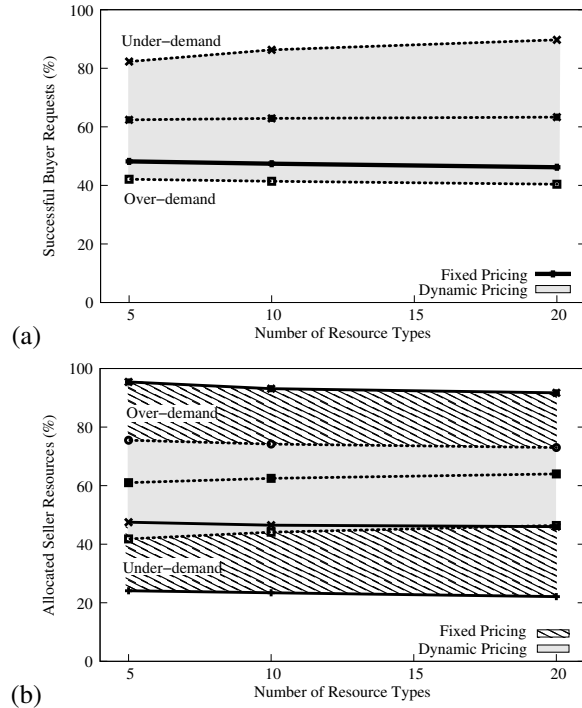
516

(a)

(b)

Figure 4: Dynamic Pricing Increases Efficiency

condition. Although fixed pricing obtain better results in the case of over-demand, our results show that the average performance of dynamic pricing is higher than fixed pricing.

5. Conclusions and Future Work

This paper discusses current resource allocation models for cloud computing and federated clouds, and shows that dynamic pricing is more suitable for federated sharing of computing resources, where rational users may both provide and use resources.

Using fixed pricing, the average user welfare is constant, since the user utility is also constant. In contrast, when using dynamic pricing, the average user welfare fluctuates with the computed payments, according to the resource demand. Moreover, a dynamic pricing scheme is able to balance the number of successful requests and the number of allocated resources depending on the market condition. For example, resource contention in the case of over-demand is balanced by increasing the resource price. Similarly, buyers are incentivized by a lower price when the market condition is under-demand. Overall, dynamic pricing achieves better economic efficiency both with higher average user welfare, and a higher number of successful buyer requests and allocated seller resources. From our experiments we find that buyer welfare is increased up to 200%, while the percentage of succesful requests is also increased up to 200%.

Even though the pricing algorithm is polynomial, scalability becomes an issue as the number of resource types in a request increases. We are currently implementing a scheme that uses distributed auctions, where multiple auctioneers can allocate different resource types at the same time. Specifically, by taking advantage of distributed hash tables, we aim to create an overlay peer-to-peer network which supports resource discovery and allocation using the proposed dynamic pricing mechanism.

References

[1] Amazon Web Services. http://aws.amazon.com, 2009.
[2] A platform for developing, deploying, and accessing planetary-scale services. http://planetlab.org, 2009.
[3] Sun Cloud Computing Initiative. http://www.sun.com/solutions/cloudcomputing, 2009.
[4] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia. Above the Clouds: A Berkeley View of Cloud Computing. Technical Report UCB/EECS-2009-28, EECS Department, University of California, Berkeley, USA, 2009.
[5] R. Buyya, C. S. Yeo, and S. Venugopal. Market-oriented cloud computing: Vision, hype, and reality for delivering it services as computing utilities. In *Proc. of the 10th Intl. Conf. on High Performance Computing and Communications*, pp. 5–13, Dalian, China, 2008.
[6] B. N. Chun and D. E. Culler. Market-based Proportional Resource Sharing for Clusters. Technical Report UCB/CSD-00-1092, EECS Department, University of California, Berkeley, USA, 2000.
[7] K. Lai, B. A. Huberman, and L. R. Fine. Tycoon: A Distributed Market-based Resource Allocation System. Technical Report cs.DC/0404013, HP Labs, Palo Alto, USA, 2004.
[8] R. B. Myerson and M. A. Satterthwaite. Efficient Mechanisms for Bilateral Trading. *Journal of Economic Theory*, 29(2):265–281, 1983.
[9] S. J. Nielson and S. A. Crosby. A Taxonomy of Rational Attacks. In *Proc. of the 4th Intl. Workshop on Peer-to-Peer Systems*, pp. 36–46, Ithaca, USA, 2005.
[10] N. Nisan. Bidding and Allocation in Combinatorial Auctions. In *Proc. of the 2nd ACM Conf. on Electronic Commerce*, pp. 1–12, Minneapolis, USA, 2000.
[11] A. Rowstron and P. Druschel. Pastry: Scalable, Decentralized Object address, and Routing for Large-Scale Peer-to-Peer Systems. In *Proc. of the IFIP/ACM Intl. Conf. on Distributed Systems Platforms*, pp. 329–350, Heidelberg, Germany, 2001.
[12] Y. M. Teo, and M. Mihailescu. A Strategy-proof Pricing Scheme for Multiple Resource Type Allocations. In Proc. of the 38th Intl. Conf. on Parallel Processing, pp. 172-179, Vienna, Austria, 2009.
[13] R. Wolski, J. S. Plank, J. Brevik, and T. Bryan. Analyzing Market-Based Resource Allocation Strategies for the Computational Grid. *Intl. Journal of High Performance Computing Applications*, 15(3):258–281, 2001.
[14] C. S. Yeo and R. Buyya. A Taxonomy of Market-based Resource Management Systems for Utility-driven Cluster Computing. *Software: Practice and Experience*, 36:1381–1419, 2006.