Efficient Resource Allocation Strategies using DRIVE for Cloud

Sowmiya. S & Kuppuswami. S
Department of Computer science and Engineering, Kongu Engineering College, Perundurai, India

Abstract – A new Economic resource allocation model has been designed, that is efficient enough to allocate resource. Current existing economic resource allocation model has been criticized for their performance limitations and high overheads. This work addresses the poor performance of current economic allocation models, by reducing the failure rate, reallocation rate, thus increasing their occupancy and system utilization. This work concentrates on high performance resource utilization strategies that can be deployed in cloud systems. Market participants can use these high performance resource utilization strategies namely overbooking and second chance substitute provider for delivering services. These strategies are implemented in DRIVE meta-scheduler and evaluated using the performance benefits and economic implications. This work introduces a new co-allocation model where a DRIVE Metascheduler receive a job from cloud user, gets different resources from resources providers to execute job. Overbooking strategy along with co-allocation system minimizes theirs execution time of the job and maximizes theirs utilization rate.

Keywords – DRIVE meta-scheduler, Economic, resource allocation, co-allocation model, re-allocation, utilization

I. INTRODUCTION

Cloud computing is an emerging technology which uses internet to work. Data and Applications are maintained using Central remote servers in cloud. Applications can be used in Cloud computing without any installation of software. The users can access the Internet and send messages anywhere in the world. Centralized storage, memory, processing and bandwidth are more efficient computing that are allowed in Cloud computing. The best example is Google mail.

Economics is the study of the allocation of scarce resources between competing alternate users. In an economic market, allocation decision making is distributed among many interested parties, whether individuals and organizations, with no ultimate central control. It is driven by the availability of a resource and the need for that resource (demand).

Resource allocation is one of the most important and difficult tasks in Cloud systems. It is the task of spreading a finite group of resources across a user population and it forms the basis of modern economics. Economies are therefore equally well suited to distributed resource allocation. Some limitations of performance in current existing economic allocation models are addressed by reducing the rate of failure and by reducing reallocation rate, thus increasing the occupancy and the system utilization rate. This work concentrates on high performance resource utilization strategies and that can be deployed in cloud systems. High performance resource utilization strategies namely overbooking and second chance substitute provider can be used by market participants to deliver services. These strategies are implemented in DRIVE meta-scheduler and evaluated using the performance benefits and economic implications. A new co-allocation model has been introduced where a DRIVE meta-scheduler receives job from cloud user and get a different offers from resources providers to execute the job. Overbooking strategy along with co-allocation system improves the benefits of user by minimizing theirs job’s execution time and improve resources benefits by maximizing theirs utilization rate. Multiple resource providers offer co-allocation policy for grouping the resources to co-allocate a cloud user’s job. The following are the objectives of proposing overbooking strategy along with co-allocation system:

- improve users benefit by minimizing theirs job’s execution time
- improve resource benefit by maximizing theirs utilization rate

The rest of the document is organized as follows: Related works are reviewed in section II. Section III addresses existing economic models and its strategies. Section IV addresses the poor performance of current economic allocation models, by reducing the failure rate and reallocation rate, thus increasing the system utilization and occupancy. Section V describes the proposed resource co-allocation system to overcome the problem that occur in the existing methods. Section VI and VII dealt with parameter evaluation and strategy evaluation respectively. Finally, section VIII conclude its work.
II. RELATED WORK

Auctions are an efficient means of economic resource allocation [1] due to their ability to establish market prices in an open market. The process of economic allocation (auctioning) is transparent to the user. Economics [2] is the study of the allocation of scarce resources between competing alternate users. Auction system is an efficient means of economic resource allocation for their ability to fix market prices in an open market. Grid system is strongly well suited for resource allocation [3]. This work dealt with the characteristics of allocation patterns for different resource allocation mechanisms and their subsequent effect on resource utilization.

The Grid system is the technology for resource-sharing among a set of participants. The participants can be some of the resource providers and consumers. The properties of a bargaining mechanism and sealed-bid mechanism are tested and their experimental results are presented in [4]. The results produced indicates that this mechanism performs good.

Second price auction [5] is employed by eBay and Amazon, Where the bidder submits a price called reservation price and has the maximum price used to bid for him by proxy. Reservation price called as proxy bid is submitted by bidder, early in the auction phase. Later he will pay the second-price auction. For ending an auction eBay and Amazon uses different rules. Auction on eBay operates on hard close end time. Auction on Amazon operates as automatically and can be extended if needed for ten minutes.

For sharing the resources such as computers, software etc., commercialization of large scale cloud system and grid system [6] requires some mechanisms. Resource requests that are quickly and efficiently servicing, is critical to the efficiency of such Grid based utility computing and communication providers. Secure auction [7] that have potential uses like eVoting, computational resource allocation and FCC spectrum auctions. SGVA privacy preserving auction scheme can conduct Combinatorial auctions.

Cloud and Grid systems can use different high performance resource utilization strategies [8]. The different strategies are overbooking [9], advanced reservation, just in-time bidding and second chance substitute providers. These strategies can be used for delivering of services. Large scale distributed computing had changes with the adoption of utility computing models by the commercial cloud providers. For improving the performance, resource occupancy and to optimize overall utilization, various proposed strategies can be employed in DRIVE [10], [11] meta-scheduler.

III. HIGH PERFORMANCE RESOURCE UTILIZATION STRATEGIES

Auction systems are an efficient means of economic resource allocation due to their ability to establish market prices in an open market. The Vickrey auction is also a sealed-bid protocol, except that the winning bidder pays the amount of the second best bid (second price). The Vickrey protocol is most appropriate for computational economies as the dominant strategy is for providers to bid their true value, with results in no counter speculation, communication overhead is also minimal as bidders bid their true value in a single sealed bid.

A. Economic Resource Allocation

Economics is the study of the allocation of scarce resources between competing alternate users. Auction system is an efficient means of economic resource allocation for their ability to fix market prices in an open market. The most appropriate protocol used for computation economies is Vickrey protocol.

B. Overbooking

The Overbooking is to exploit overestimation of required job execution time. The main notion of overbooking is to schedule more number of additional jobs. Overbooking strategy used in economic model can improve system utilization rate and occupancy. In overbooking strategy every job is associated with release time and finishing deadline, as shown in Fig 1. Here successful execution will be given with fee and penalty for violating the deadline.

C. Second Chance Substitute Providers

It will be a waste of resources to re-execute the auction process again if the winner of the conducted auction cannot meet his/her obligations, when there is enough capacity available from non winning providers. If this case occurs, the losing bidders of that auction can be given a second chance to win the auction. Here the auction can be re-computed without defaulting the bidder. Thus this process can increase the system utilization and can highly reduce the allocation failures which was caused by overbooking.

D. DRIVE

DRIVE stands for Distributed Resource Infrastructure for a Virtual Economy. DRIVE is a distributed economic meta-scheduler and it is used to efficiently allocated resources. DRIVE is designed for an efficient and richer economic model for allocation. Now a days Economic markets are most commonly used
as a means of resource allocation system. DRIVE uses Vickrey auction protocol to allocate resources. Vickrey auction is the second price sealed bid protocol. For establishing the market prices quickly and efficiently for goods/services, DRIVE is designed to use auction protocols. This Distributed DRIVE meta-scheduler is different from other existing meta-scheduler. Combinatorial auction is also supported by DRIVE. DRIVE supports for privacy and verifiability. The Overbooking and second chance substitute provider strategies are implemented in DRIVE. During negotiation in auction between the auctioneer and bidders, the latency occurs. In order to minimize the effect of latency, a progressive two phase contract mechanism was proposed. That two phase contract structure is shown in Fig.2. Once the DRIVE determines the winner at the end of auction, it undergoes a two phase contract model.

Fig. 2: Reverse auction lifecycle in DRIVE.

In phase 1, a tentative agreement is created between the user and the winning bidder. Before the redemption this agreement must be hardened into a binding agreement or the contract. Contract defines the particular levels of service to be delivered along with a set of rewards and penalties for honoring or breaking the agreement (phase 2).

IV. PROBLEM DESCRIPTION

Economics is the study of the allocation of resources that are scarce between competing alternate uses. Auction system is an efficient means of economic resource allocation. It has ability to fix market prices in an open market. Most often proposed efficient means of resource allocation is the Computational economies. But their adoption is criticized for their poor performance and high overheads. This work, addresses the poor performance of current economic allocation models by introducing some of the strategies. The strategies considered are overbooking, second chance substitute providers and along with co-allocation for service delivery. These strategies can reduce the failure rate and reallocation rate, thus increases the occupancy and system utilization. The proposed strategies can be implemented in DRIVE. Thus the co-allocation of resources with these strategies increases the system utilization rate, increases the occupancy and thereby improve the provider’s revenue.

V. PROPOSED RESOURCE CO-ALLOCATION SYSTEM

A. DRIVE Architecture with Co-Allocation System

This proposed work focuses on resources co-allocation using DRIVE in cloud computing. Co-allocation of resources is the process where more than one resources are simultaneously allocated to the user. Co-allocation model is integrated within the DRIVE meta-scheduler to increase system utilization and occupancy. As shown in Fig. 3., a new co-allocation model can be introduced, where the DRIVE receives job from cloud user, get a different resources from resources providers to execute the job and then later it dispatches the job to the selected resources based on the best provided offers to improve resource and user benefits. The objective of integrating overbooking with the co-allocation system are as follows:

- improves benefit of users by minimizing theirs jobs execution time
- improves benefit of resources by maximizing resource utilization rate

Co-allocation model have three types of participants namely cloud users, Meta-scheduler and resource providers. Number of processors, speed of processing, local scheduling etc., may differ from one resource providers to the rest of other resources providers. DRIVE the meta-scheduler gets the respective job from the cloud users, then selects the resources as required for the job from the resources providers and then the final step is to submit the job to the selected resources.

To meet a job’s requirement, resources from different resource providers must be composed with various types of resources. DRIVE meta-scheduler is responsible for meeting a job’s requirement. The resources that are composed determines how much work the meta-scheduler should send to each resource provider. The main aim of DRIVE meta-scheduler is to meet the user’s requirement, maximize the rate of resource utilization, minimize the job’s execution time and thus maximize system utilization.

B. Experimental Economy

The pricing function and penalty function are two types of experimental economy used in this work. These function depends on local information known by the provider.

1) Pricing Functions:

Any one of the following functions can be used to provide rewards for honoring the agreement. These functions are based on projected conditions and current
conditions. $P_{\text{unit}}$ is the price per job unit, $b$ denotes a bid and its range is $(b \in (0,B))$.

- Random: The unit price can be random, irrespective of any other factors.

$$P_{\text{unit}} = \text{Random} \ (0,B) \quad (1)$$

- Constant: The unit price is constant.

$$P_{\text{unit}} = c, \ c \in (0,B) \quad (2)$$

• Available Capacity: The unit price is formulated based on projected provider capacity at the time when the job would execute. $U_{\text{provider}}$ is the projected utilization of the provider, $U_{\text{job}}$ is the utilization of the requested job, and $C_{\text{provider}}$ is the total capacity of the provider.

$$P_{\text{unit}} = \frac{U_{\text{provider}} + U_{\text{job}} \times B}{C_{\text{provider}}} \quad (3)$$

2) Penalty Functions: Penalty functions can be constant or dynamic. Different penalty functions are:

- Constant: Penalty unit is constant irrespective of other factors.

$$P_{\text{defaulter}} = c, \ c \in R \geq 0 \quad (4)$$

- Job units: $J_{\text{units}}$ is the number of units in a job and $c$ is a constant penalty per unit.

$$P_{\text{defaulter}} = J_{\text{units}} \times c, \ c \in R \geq 0 \quad (5)$$

- Win price: $Price_{\text{win}}$ is the price that is to be paid by the bidding winner

$$P_{\text{defaulter}} = Price_{\text{win}} \quad (6)$$

- Substitute Price: $Price_{\text{substitute}}$ is the price that is to be paid by the substitute bidding winner

$$P_{\text{defaulter}} = Price_{\text{substitute}} \quad (7)$$

VI. PARAMETERS FOR EVALUATION

In this paper, the proposed strategies are evaluated with respect to

- Allocation occupancy
- System utilization
- Economic implications
Here the occupancy is stated as the number of contracts that are satisfied i.e., the task that is allocated and completed during the auction. System utilization is defined as the amount of host’s resource capacity used by the winner of the auction. Economic implication has its two types namely, pricing function and penalty function. Economic implication must increase the provider’s and consumer’s revenue.

VII. STRATEGY EVALUATION

A theoretically comparison is made between different strategies with system utilization rate and their allocation occupancy with respect to number of contracts created. Here the strategies are denoted as Overbooking (O), second chance Substitutes (S) and Co-allocation (C). A Guaranteed strategy (G) is implemented to make a comparison with other proposed strategies. Each of the strategy uses Vickrey protocol to allocate tasks. Average system utilization rate is produced by running each strategy three times.

- **G**: Bidders can bid for resources that never goes beyond the allotted capacity. Here bids are guaranteed. So, bidders are not allowed to reject the contract made. There exists no use of second chance substitute providers strategy. Using this strategy 34.41% of contracts can be completed and its maximum system utilization rate will be 29.65%.

- **O**: Bidders can bid for resources that goes beyond the allotted capacity. Bidders can accept/reject contracts made at the time of contract creation. Using this strategy 84.5% of contracts can be completed and its maximum system utilization rate will be 61.90%.

- **O + C**: Bidders can bid resources that goes beyond the allotted capacity and additionally they can bid more than one resources at the time of auction. So that it increases the system utilization and occupancy further. Using this strategy 84.5% of contracts can be completed and its maximum system utilization rate will be 75.63%.

- **S + O**: Bidders can bid resources that goes beyond the allotted capacity. Incase of rejecting the contract or if the bidding winner cannot meet their obligation, the losing bidders may be provided with a second chance substitute provider strategy. This over comes the failure rate that occur due to overbooking. Using this strategy 98.95% of contracts can be completed and its maximum system utilization rate will be 79.55%.

- **S + O + C**: Same as S+O but additionally the bidder can bid for more than one resources at the time of auction. The losing bidders can also have opportunity to bid more than one resources at the time of auction. Using this strategy 98.95% of contracts can be completed and its maximum system utilization rate will be 95.46%.

A. Simulation Tool To be Used

CloudSim is an extensible simulation toolkit and it is a framework for modeling and simulation of cloud computing infrastructure and application services. CloudSim allows the cloud users to concentrate on some system design issues and they need not get concerned about the low level details that are related to the cloud-based infrastructures and its application services. Cloudsim can be used for,

- Cloud resource provisioning
- Managing data center
- Research Activities
- Optimization of Cloud Computing

VIII. CONCLUSION

The success of Cloud computing can be measured by effective utilization of resources for the different computational jobs. The high performance computing economic resource allocation models has been re-motivated by the commercial cloud providers. There are numerous resources available on Cloud. These resources can be effectively utilized with the co-allocation of resources strategy. The co-allocation model has two main objectives namely execution time and resources utilization rate. In co-allocation model resources are composed from various resources providers for co-allocating a cloud user’s job. Thus these proposed co-allocation model with other strategies can improve the allocation performance and thereby increases the system utilization rate and occupancy. Provider revenue also increases with the use of these proposed strategies, for their increased allocation rate.

IX. REFERENCES

[1] Schubert L. (2010) “The Future of Cloud Computing”. European Commision, information society and Media.

[2] Simon Mark Davy (2008) “Decentralised Economic Resource Allocation For Computational Grids “. The University of Leeds School of Computing.

[3] Krawczyk S. and Bubendorfer K. (2008) “Grid Resource Allocation: Allocation Mechanisms and Utilization Patterns”, Victoria University of Wellington, Kelburn, Wellington, New Zealand .

[4] Schotter A. and Radner R. (1989) “The Sealed-Bid Mechanism: An Experimental Study” New York Universifv, New York and Murray Hill, New Jersey.

[5] Roth A.E and Ockenfels A. (2002) “Last-minute bidding and the rules for ending second-price auctions: Evidence from ebay and amazon auctions on the internet,” American Economic Review, vol. 92, no. 4, pp. 1093–1103.

[6] Welch I., Bubendorfer K. and Chard B. (2006) “Trustworthy Auctions for Grid-style Economies” Victoria University of Wellington, New Zealand.
[7] Palmer B., Bubendorfer K. and Welch I. (2007) “Anyone Can Hold An Auction”, Victoria University Wellington, New Zealand.

[8] Chard K., K. Bubendorfer K. and Komisarczuk P. (2010) “High occupancy resource allocation for grid and cloud systems, a study with drive,” in Proceedings of the 19th ACM International Symposium on High Performance Distributed Computing (HPDC '10). New York, NY, USA: ACM, pp. 73–84.

[9] Birkenheuer G., Brinkmann A. and H. Karl H. (2009) “The gain of overbooking,” in Proceedings of the 14th International Workshop on Job Scheduling Strategies for Parallel Processing (JSSPP), ser. LNCS, vol. 5798, pp. 80–100.

[10] Chard K. and Bubendorfer K. (2008) “A distributed economic metascheduler for the grid,” in Proceedings of the 8th IEEE International Symposium on Cluster Computing and the Grid (CCGRID'08). Washington, DC, USA: IEEE Computer Society, pp. 542–547.

[11] Kyle Chard and Kris Bubendorfer, “High performance resource allocation strategies for computational economies,” IEEE Transactions on Parallel and Distributed Systems, 2012.

[12] Bubendorfer K., Palmer B. and Welch I. (2008) “Trust and privacy in grid resource auctions,” in Encyclopedia of Grid Computing Technologies and Applications, E. Udoh and F. Wang, Eds. IGI Global.

[13] John W. and James F. (2010) “Cloud Computing Implementation, Management and Security” CRC Press.

[14] Sutherland I. E. (1968) “A futures market in computer time,” Communications of the ACM, vol. 11, no. 6, pp. 449–451, 1968.

[15] Stoica I. and Rabkin A. (2009) “Above the Clouds: A Berkeley View of Cloud Computing” University of California, Berkeley.