

## Resource Management of IaaS Providers in Cloud Federation

Behnam Bagheri Ghavam Abadi, Mostafa Ghobaei Arani

*Department of Computer Engineering, Islamic Azad University of Mahallat,  
Mahallat, Iran*

*Department of Computer Engineering, Islamic Azad University of Parand, Tehran,  
Iran*

*Email: Behnamfix@gmail.com, mostafaghobaei@piaau.ac.ir*

### **Abstract**

*With the increasing use of cloud services, for both providers and users, saving energy and using needed resources regardless of time and place have got outmost importance. Resource provisioning in the cloud providers has come along availability of workloads in a given time. Most cloud providers conduct inquiries with a limited amount of resources that may be cause to reject the request of customers at the peak of workloads. Cloud federation is an approach to share resources, enhance the scalability and availability. High energy consumption is one of the current challenges for cloud providers; also it should be discussed and researched about providers' profit as an important issue. We have presented an approach in this paper to reduce power consumption for IaaS Providers, by choosing the most appropriate host and allocating the best virtual machine which leads to satisfy users requests, save energy and reduce the cost of resources. Simulation results show that the proposed approach, compared to the other similar approaches, causes to increase utilization and turns off idle servers to decrease consumed power which followed by an increase in providers' profit.*

**Keywords:** *Cloud computing, Cloud federation, IaaS provider, Power consumption*

### **1. Introduction**

Cloud computing technology is an important step in distributed processing. Cloud computing has many different aspects of sharing the software and hardware solutions which includes computing and storage resources, time of functional programs or complex application programs properties. Data centers are of the components in cloud computing [1-6]. A cloud is defined as a place over network infrastructure where information technology (IT) and computing resources such as computer hardware, operating systems, networks, storage, data bases, and even entire software applications are available instantly. Cloud computing is the use of cloud resources (hardware and software) that are delivered as a service over a network (typically the Internet) [3]. Providing Infrastructure as a Service (IaaS) through clouds has drawn much attention from key information technology (IT) players (*e.g.*, Amazon, Microsoft, IBM, HP) in the last few years. In general, public cloud providers adopt a pay-per-use model, where customers can “rent” virtualized computing resources, use these resources for a specific amount of time, and release them when they are not necessary any longer [7-10]. Increased demand in data center has led to an increase in services so that the large-scale data centers have been created. In the past, high-performance was the most important criterion in designing data centers. Today with development of cloud-based data centers and increasing demand for cloud services, the main problems are high energy consumption and the profit of the services. On the other hand the costs for establishment of the services must be such as providers have gain substantial profits to be able to earn a good income from investments.

However it should consider other challenges such as level of carbon production from the services which has an important role in weather by greenhouse effect [11, 12]. Idle servers are one of the most important factors for wasting energy on cloud data centers who have low utilization of consuming hardware resources, as servers doesn't have power fit. It means light workload or even idle servers consume large amount of energy compared to the time of maximum utilization of their resources [13].

In this paper, we propose an approach to reduce power consumption in data centers and increase profits for the satisfaction of service providers offering cloud services. The proposed approach is offered based on a real model for cloud federation; the approach derives from policy management techniques also maximizing utilization by improving the server centers. The proposed approach increases utilization as much as possible and reduces power consumption in data centers so increases profit for them, so that providers reject fewer requests that enhance their reputation.

The rest of this paper is organized as follows: The second section is devoted to the related works and the proposed approach is presented in third section. Evaluation and simulation results come in the fourth section and finally the fifth section includes conclusion and future discussions.

## 2. Related Works

Regarding to resource management in the cloud federation, most of the researches focus on maximizing IaaS providers' profit in cloud computing environment. In this section, we are going to consider an overview of related research on the approaches of resource management:

Goiri *et al.* [14] present a characterization of providers operating in a federated Cloud which helps to choose the most convenient decision depending on the environment conditions. These include outsourcing to other providers, renting free resources to other providers (i.e., insourcing), or turning off idle nodes to save power. This technique does not reduce financial costs related to the purchase and hosting of IT equipment, which will be amortized. This forces the providers to increase their prices to be profitable, which makes them less competitive in the Cloud market.

Adel *et al.* [15] attempt to incorporate the outsourcing problem with option of terminating spot VMs within a data center. The main objective for increase profit is proposing policies that help making decisions when providers have different choices regarding incoming requests: rejecting, outsourcing, or terminating spot leases to free resources for more profitable requests. So that it agrees all requests based on probable demand by accepting, and this is a priority. They also have focused on specific policies applied by Cloud IaaS resource providers to decide when and how computational resources should be sold to make them available in the market for other IaaS providers.

Patel *et al.* [16] have presented two algorithms for VM allocation among multiple cloud providers. The first one allocates the resources to VM to improve the cloud service provider's profit. Second one allocates VM in order to balance the load among the multiple data centers in federated cloud environment. The first algorithm includes equations that help cloud providers in making decisions (i.e. when to outsource and insource resources to and from other providers). Simulation tests evaluate these equations to find impact of parameters on the providers' profit. The parameters include workload, free sold resources, cost of outsourcing, the amount of outsourced resources and cost of servers' maintaining.

Kumar *et al.* [17] have designed a data communication model for the members of a cloud confederation, and develop an algorithm and a model for cost calculation, which enhances the decision making process over all the VM types (on-demand, reserved, spot) to increase resources utilization and profit. They particularly focus on federation of IaaS providers that serve requests for VMs. Since every provider has limited amount of

capability, increase in load possibly will overload a provider's data center as well as may result in QoS infringement or users' request rejection. Providers that support dissimilar types of QoS and pricing plan for VMs, have the option of canceling their less portable VMs (spot VMs) in favor of additional profitable requests (on-demand VMs). Providers can also get advantage from federation through outsourcing requests to additional members of the federation by means of least load.

Michele *et al.* [18] propose and evaluate energy-aware allocation policies that focus at maximizing the revenue earned by Cloud providers. The maximization is achieved by improving utilization in the server centers, i.e., by powering excess servers off. The policies they have proposed are based on 1-dynamic estimates of user demand and 2-models of system behavior. The emphasis of the latter is on generality rather than analytical tractability; thus, we employ approximations for handling the resulting models. However, those approximations lead to algorithms that perform well under different traffic conditions and can be used in real systems.

### 3. The Proposed Algorithm

In this section, we propose an algorithm based on series of techniques reducing power consumption and in the first stage we arrange hosts according to CPU usage in a descending order. This will result that inquiries go to hosts with high workload so that cause to increase utilization. This will also lead to implement less host and higher utilization. At the last stage also it causes to increase the number of hosts in off mode. And higher number of off hosts means to avoid consumption of static energy in hosts. It means we have decreased consumption of energy in hosts practically however our requests have been operating. So cloud service providers should pay less for energy and this equals to more profit for them.

In the second stage, when we arranged and selected appropriate host, we should select the virtual machine (VM). Among the VM options inside the selected host, we should select the smaller VM according to power consumption and size of CPU and RAM, to apply request inside of it. So that larger virtual machines will save for later requests which should be executed next, larger requests that we cannot satisfy with small VMs. so this will also save larger virtual machines and will increase utilization of virtual machines.

The policies are done for each type of virtual machine but if they cannot run for the reserved and spot type, their requests will reject. Adjacent providers assist in the on-demand type providers. in this case, that request will send for left and right providers then they consider and decide based on the policy to do it or not. As they are working based upon their own benefit, if it will not be successful they will send it to the next neighbor and the process will be continued till it gets to the first provider. Now provider can select the best option according to the price list which has sent by neighbor providers.

In fact the proposed algorithm is an algorithm which we could reach to a beneficial green cloud method. It means we have reduction in energy consumption and increase in profit. The general process proposed algorithm has been shown in Figure 1.

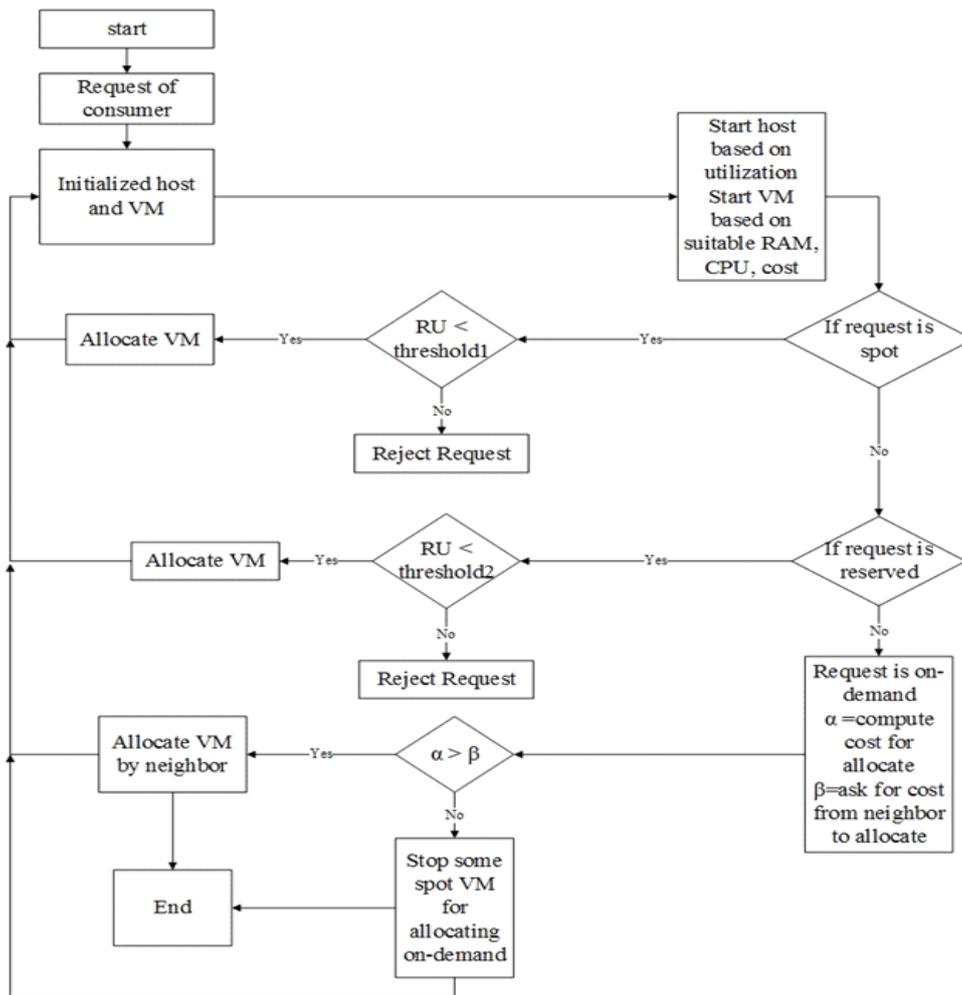


Figure 1. The Proposed Algorithm

#### 4. Performance Evaluation

In this section, we have used a Cloudsim toolkit [19] to simulate proposed approach. We used it to model the virtualized resources needed for cloud computing which have real data workload. We consider 8 types of VMs for simulation whose specifications have been shown in Table 1.

Table 1. Types of Virtual Machines Used in the Simulation

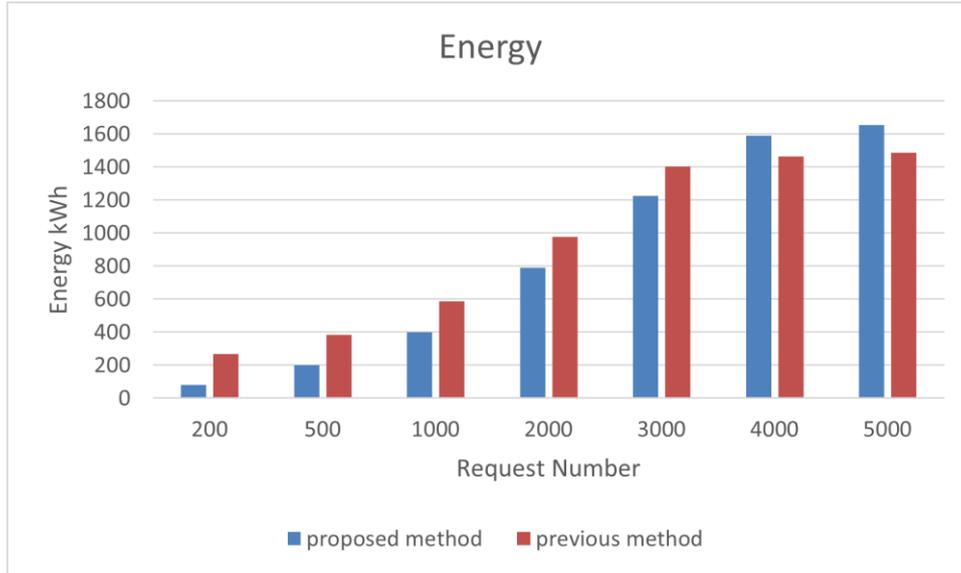
Type	MIPS	PE	RAM	BW	SIZE	COST
1	2500	1	3500	100Mits/s	1GB	1300
2	2000	1	2500	100Mits/s	1GB	1250
3	2200	1	1600	100Mits/s	1GB	1200
4	500	1	800	100Mits/s	1GB	600

5	600	1	1200	100Mits/s	1GB	700
6	600	1	1200	100Mits/s	1GB	700
7	500	1	1500	100Mits/s	1GB	650
8	760	1	2000	100Mits/s	1GB	800

Test results have been calculated by given average 800 hosts and 1000 virtual machines on a daily basis. The results have been estimated in 10 working days. We have considered the scenario to assess efficiency of offered algorithm that shows the impact of changes in request number.

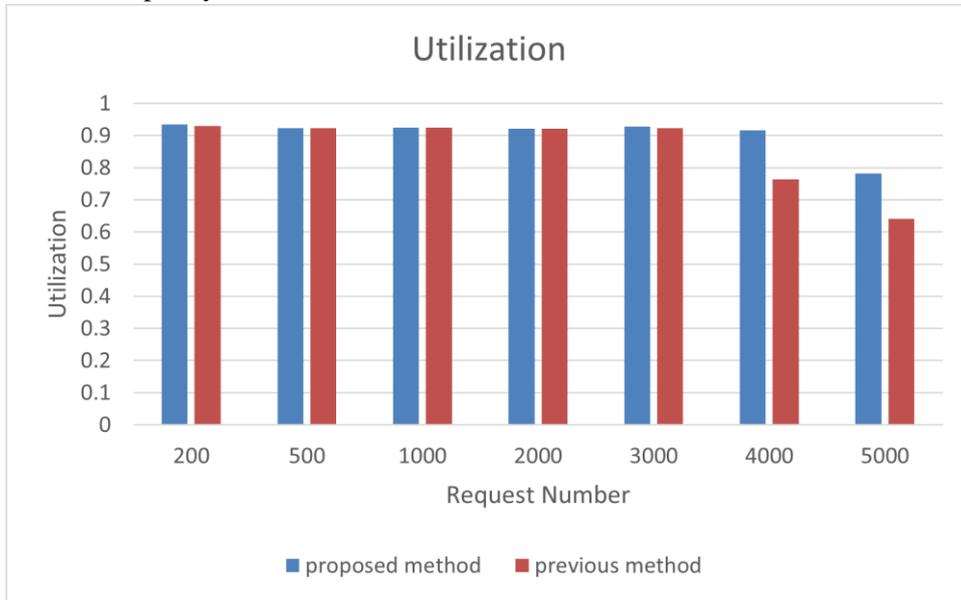
#### 4.1 Experimental Results

In this scenario, we examined energy consumption and profits due to change in the number of requests; several parameters have been considered as: energy, utilization, number of rejected requests, profit and profit impact on energy in the charts and the results are compared to the basic approach. Figure 2 shows the energy consumption for both proposed approach (proposed method) and basic approach. The X axis shows the number of presumptive requests that start with 200 requests and at last it reaches to 5000. The Y axis displays energy in terms of kilowatt per hour. As seen in this chart due to the increase in the number of requests, energy consumption is increasing and this is a normal process because whatever the provider runs more requests, more energy is consumed. In the mode of 4,000 and 5,000 requests, our energy consumption is more than the basic approach because the workload has increased in host and fewer requests are rejected compared to basic approach.



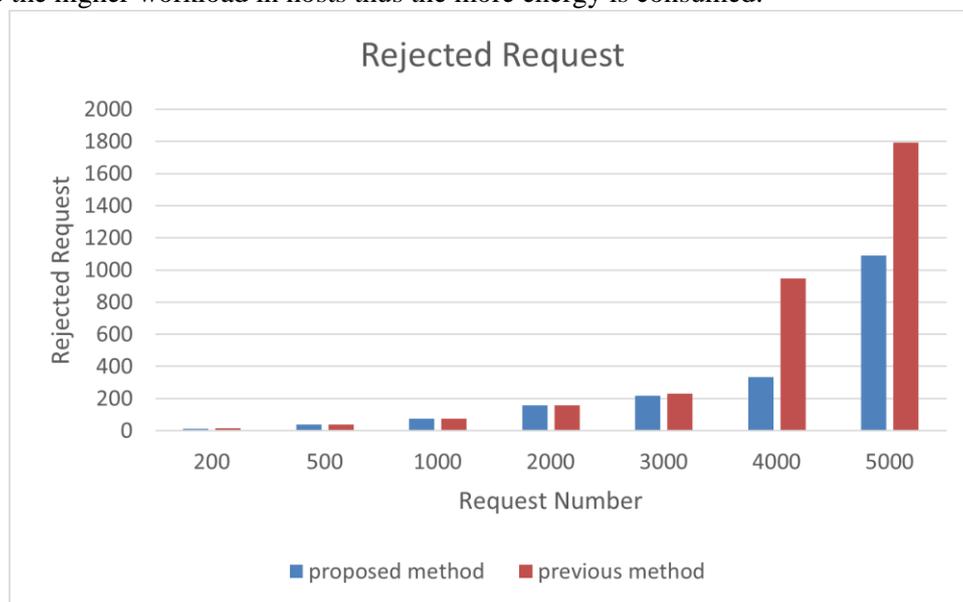
**Figure 2. The Comparison of Energy Consumption with for Varying Request Number**

In Figure 3, whatever the number of requests are increasing, utilization remains constantly and has not changed because there is considered a specific policy. The utilization is related to the active hosts. However in the last mode with 5,000 requests, it causes a series of larger requests to be implemented and a large number of requests to be rejected. Practically some small VMs in the hosts have not been executed, so utilization decreases. Utilization has increased compared to the basic approach, because we use the maximum host policy.



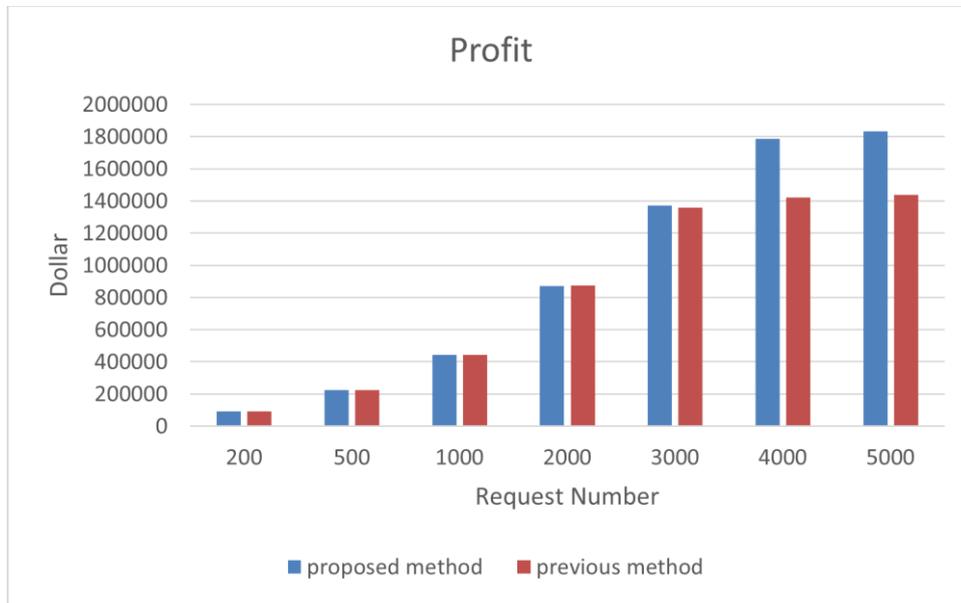
**Figure 3. The Comparison of Utilization in Provider with Varying Request Number**

In Figure 4, whatever the number of requests has increased, as well as the number of rejected requests has been increased which is the best state in the 4000 requests. As you can see in the mode of 5000 requests, the number of rejected requests has increased that can cause customer dissatisfaction also there might be occurred some violations of the Service Level Agreement. As you can observe, the number of rejected requests compared to the basic approach has been reduced and there are more requests running that will cause the higher workload in hosts thus the more energy is consumed.



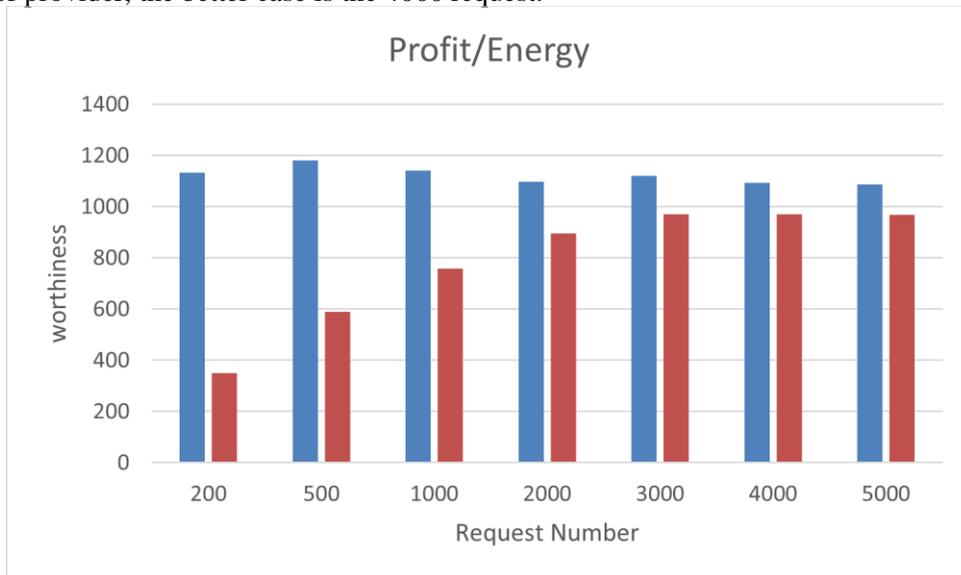
**Figure 4. The Comparison of Reject Request with Varying Request Number**

In Figure 5, whatever the number of incoming requests increases and more requests run, the profit rate increases up to 4000 requests which is the best state in the graph. Profit rate has not changed in 5000 requests because for in the mode of 5000 requests, the number of rejected requests has increased. Therefore the most reasonable mode is 4000 requests where the state's most profitable. As you can see in 4000 and 5000 requests in comparison with the basic approach, more requests have been run with rising workloads and maximum hosts using we gain more profit.



**Figure 5. The Comparison of Provider Profit with Varying Request**

In Figure 6, we consider the amount of profit per energy. As seen in the figure. Proposed method is not sensitive to the number of requests but in the basic approach, rate of profit/energy will be different depending on the number of requests, the profit gives a different energy. In other words, in the proposed approach an increase in the number of requests leads to increase energy consumption. Also the level of profit increases by reducing the number of rejected requests therefore the profits /energy will be the same and equal due to the number of different requests. However, if we want to look at their both dimensions, the best interest and energy of the 500 request, but if we only look at the profit for provider, the better case is the 4000 request.



**Figure 6. The Comparison of Profit/Energy between Basic Method and Proposed Method for Varying Request**

## 5. Conclusion and Further work

With daily increasing use of cloud services, saving energy and using needed resources regardless of time and place, for both providers and users, have great importance. Maximum use of active available resource is a suitable approach to increase utilization and improve cloud systems functions. In this paper we tried to offer a new approach to improve quality of cloud services and reduce rejected requests. Our proposed approach is based on resource management and profitability techniques for various types of virtual machines which focuses on the customer requests and offers the best host and VMs to meet their requests. Also it is an appropriate method to minimize rejected requests, to have best operation and being the best option among the others as a green cloud method. The bandwidth and correlation index in selecting VM, selecting hosts with high efficiency also using various thresholds are all matters which has not mentioned in this article and it needs further discussion, debate and exploration in the future.

## References

- [1] A. Kertesz, J. D. Dombi, and A. Benyi. "A Pliant-based Virtual Machine Scheduling Solution to Improve the Energy Efficiency of IaaS Clouds", *Journal of Grid Computing*, (2015), pp.1-13.
- [2] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, "A view of cloud computing", *Communications of the ACM*, vol. 53, no. 4, (2010), pp. 50-58.
- [3] S. S. Manvi and G. K. Shyam, "Resource management for Infrastructure as a Service (IaaS) in cloud computing: A survey", *Journal of Network and Computer Applications*, vol. 41, (2014), pp. 424-440.
- [4] A. Fereydooni, M. G. Arani and M. Shamsi, "EDLT: An Extended DLT to Enhance Load Balancing in Cloud Computing" *International Journal of Computer Applications*, vol. 108, no. 7, pp. 6-11.
- [5] A. Ghazi, M. G. Arani, H. Babaei "A New Framework for the Evaluation of QoS in Cloud Federation. ", *International Journal of Computer Applications*, vol. 107, no. 1, pp. 44-49.
- [6] B. S. Taberi, M. G. Arani and M. Maeen, "ACCFLA: Access Control in Cloud Federation using Learning Automata" *International Journal of Computer Applications*, vol. 107, no. 6, pp. 30-40.
- [7] J. A. Wickboldt, R. P. Esteves, M. B. de Carvalho and L. Z. Granville, "Resource management in IaaS cloud platforms made flexible through programmability", *Computer Networks*, vol. 68 (2014), pp. 54-70.
- [8] M. Uddin, Y. Darabidarabkhani, A. Shah and J. Memon, "Evaluating power efficient algorithms for efficiency and carbon emissions in cloud data centers: A review", *Renewable and Sustainable Energy Reviews*, vol. 51, (2015), pp. 1553-1563.
- [9] L. Liu, H. Wang, X. Liu, X. Jin, W. B. He, Q. B. Wang and Y. Chen, "GreenCloud: a new architecture for green data center", *Proceedings of the 6th international conference industry session on Autonomic computing and communications industry session*, ACM, (2009).
- [10] A. Gandhi, M. Harchol-Balter, R. Das and C. Lefurgy, "Optimal power allocation in server farms", *ACM SIGMETRICS Performance Evaluation Review*, vol. 37, no. 1, (2009), pp. 157-168.
- [11] M. Fallah, M. G. Arani "ASTAW: Auto-Scaling Threshold-based Approach for Web Application in Cloud Computing Environment", *International Journal of u- and e- Service, Science and Technology (IJUNESST)*, vol.8, no.3, (2015), pp.221-230.
- [12] Í. Goiri, J. Guitart and J.Torres, "Economic model of a Cloud provider operating in a federated Cloud", *Information Systems Frontiers*, vol. 14, no. 4, (2012), pp. 827-843.
- [13] A. N. Toosi, R. N. Calheiros, R. K. Thulasiram and R. Buyya, "Resource Provisioning Policies to Increase IaaS Provider's Profit in a Federated Cloud Environment", *High Performance Computing and Communications (HPCC)*, 2011 IEEE 13th International Conference on, (2011).
- [14] K. S. Patel and A. K. Sarje. "VM provisioning policies to improve the profit of cloud infrastructure service providers", *Computing Communication & Networking Technologies (ICCCNT)*, 2012 Third International Conference on, (2012), pp. 1-5.
- [15] A. K. Das, T. Adhikary, M. A. Razzaque, E. J. Cho, and C. S. Hong, "A QoS and profit aware cloud confederation model for IaaS service providers", *Proceedings of the 8th International Conference on Ubiquitous Information Management and Communication*, (2014), p. 42.
- [16] M. Mazzucco and D. Dyachuk, "Optimizing cloud providers revenues via energy efficient server allocation", *Sustainable Computing: Informatics and Systems*, vol. 2, no. 1, (2012), pp. 1-12.

- [17] D. C. Erdil, "Autonomic cloud resource sharing for intercloud federations", *Future Generation Computer Systems*, vol. 29, no. 7, (2013), pp. 1700-1708.
- [18] J. Tong, E. Hai-hong, M. Song, and J. Song, "Host load prediction in cloud based on classification methods", *The Journal of China Universities of Posts and Telecommunications*, vol. 21, no. 4, (2014), pp. 40-46.
- [19] R. N. Calheiros, R. Ranjan, A. Beloglazov, C. A. De Rose and R. Buyya, "CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms", *Software: Practice and Experience*, vol. 41, no. 1, (2011), pp. 23-50.

## Authors



**Behnam Bagheri Ghavam Abadi**, he received the B.S.C degree in Software Engineering from Azad University dolatabad isfahan, Iran in 2011, and M.S.C degree from Azad University of mahallat, Iran in 2015, respectively. His research interests include Cloud Computing.



**Mostafa Ghobaei Arani**, he received the B.S.C degree in Software Engineering from IAU Kashan, Iran in 2009, and M.S.C degree from Azad University of Tehran, Iran in 2011, respectively. He's Currently a PhD Candidate in Islamic Azad University, Science and Research Branch, Tehran, Iran. His research interests include Grid Computing, Cloud Computing, Pervasive Computing, Distributed Systems and Software Development.

