International Journal of Advanced Research in IT and Engineering

ISSN: 2278-6244
Impact Factor: 6.111

AN OVERVIEW OF COST PROVISIONING STRATEGIES FOR CLOUD COMPUTING

Dr. A. V. Karthick*

Dr. K. Muthu Pandi**

Abstract: Cloud Computing is a collection of software, hardware, networking and storage resources that provides services based on contract between the user and Cloud Service Provider (CSP). Cloud Computing offers two types plans namely reservation and on-demand plan. Cost provisioning strategy is one of the most challenging task in cloud computing. The identifiable key advantage of cloud computing is billing system based on usage. The client wants to access the service for more than 3 years, the currency value and resources like manpower, hardware, software and storage cost will increase or decrease. If increases the cost of resources it will affect the CSP, on the other hand if decreases the cost of resources will affect the clients. Client needs are dynamic in nature; it cannot easily and accurately predict them. Collect the reasonable cost for accessing the cloud computing is one of the largest task. Client gathers information regard select which CSP gives service at low cost, maximum bandwidth, more number of attractive features and so on. Based on this features the client selects suitable service provider from the perfect competitive industry.

Keywords: cost of access, cloud service provider, client, time period, cost provisioning, challenging task.

*Teaching Assistant, Alagappa Institute of Management, Alagappa University, Karaikudi, Tamilnadu, India

**Assistant Professor, Department of Commerce, Sourashtra College, Madurai, Tamilnadu, India

INTRODUCTION

Cloud Computing is one of the predominant area in Information Technology highway. It is a utility based computing that convert the thoughts of large IT companies in to make software as most eye-catching one It is the next generation data center that provides different types [8] of applications based on the user needs. It gives diverse mixture of deployment and configuration. The biggest problem in cloud computing was jobs with variety of applications, dissimilar workload, size of the system and performance measurement of the system. Short term duration the time period is between 1month to 1 year and in long term plan for more than 1 year. CSP has many numerals of unique ideas to deploy their service in to the industry. It is widespread area coverage network that has multiple server rather than run in a single remote server or any local server. It provides error free, reliable and scalable service to their client.

In Optimal Cloud Resource Provisioning (OCRP) algorithm [10] Benders Decomposition technique has more number of steps. It splits the main module into more sub modules. It affects the quality of service the client waited more time for access the services. Needs and wants of the clients are increasing and varying one. It is uncertain to accurately predict the client demands. In the existing system like K-Means neighbor [9] algorithm was used for forecast the demand. Heuristics based algorithm does not provide effective utilization of system resources. Stochastic based program are used to forecast the uncertainty of client demands. It was difficult to the implement this method in real world problem.

Load balancing was one of the important part in cloud scheduling, it helps to minimize makespan, reduce waste of resources and cost reduction for accessing the service. Under provisioning and over provisioning of resource leads to waste of all the resources, maximize its execution time and increase the cost of access. The Software industry is dynamic one. The changes made by the user wants also reflects in the cost of access the service without disturbing the current work is one of the toughest job for the CSP.

LITERATURE SURVEY

There are so many cost provisioning strategy are used in cloud computing. Job grouping based scheduling algorithms like Tri Queue job Scheduling and Multi Queue job Scheduling algorithm helps to reduce the cost of access. The CSP provides better quality service to its clients. Cloud Computing needs was complicated when calculate the resources that request

made by Cloud Customers. It shows the vibrant behavior in utilize a different variety of services as a result of unstable utilization behavior of customers suffers it. The customer did not got any incentive for his loyalty. This approach doesn't produce a balanced result for all kinds of customers.

The cloud user mostly prefer reservation plan due [3] to its cheaper price. The drawback was highly unutilized capability of upcoming demands. The biggest problem was how to improve the usage of idle reserved resources. A rental based non-cooperative game was used to reallocate the redundant resources. Based on past demand [4] the CSP is used to analysis deterministic equivalent formulation to find optimum solution. It shows that sensitivity analysis of prices was used for both users and service provider measured in terms of increasing cost efficiency. It ignores factors such as random network delay and VM migration. This approach wants more improvement needed in computational performance especially in distributed system.

The Total Cost of Ownership [5] in cloud deployment was cheaper than on-premises or collocation. The overall costs used for cloud was not overlooked for small and middle-sized enterprises rather than large scale enterprises. In small and medium enterprises concentrates particularly in overhead owed due to the improper management and enlarged nodes. In cost efficient process they treated cloud computing as a discrete optimization problem for economic constraints in on-demand phase. The scale up of the service was decisive to keep the cost reduction. Betterment process is not a easy one. Fine tuning is a never ending process. The drawback identified in this paper was CSP does not automatically generate provision elasticity in cost cut intelligence. The machine learning or analytics it was more complex for progress and obvious in vibrant cost saving problem.

Amazon Elastic Compute Cloud (EC2) offers three kinds of offer to clients like reservation, on-demand and spot options. In this paper [7] they give three options, each option has different cost and it provides various benefits to its client. Spot price means the pricing based on spot option. It is more flexible and costlier one. The cost of access the reservation and on-demand phase was decided due to demand and supply of the present resources in EC2. Both reservation and on-demand plans are more stable one than the cost of spot option. The biggest crucial decision was taken by client was made on effective purchase from the available different plan. In Amazon EC2 Environment, they are using two types of

algorithm namely long-term provisioning algorithm and short-term provisioning was designed for forecast the feasible solution. The time duration for long provisioning in on-demand category was between 1 to 3 years, where as in reservation category has more number of years. The time duration for short term provisioning was within few minutes to several limited hours. Cost of access the short period aims to give service at reduced total cost.

In this paper [9] the author forecasted the client future demand was unpredictable. K-nearest algorithm was used for predict the demand with the help of Probability Distribution Function. Optimal Virtual Machine Placement (OVMP) algorithm was used to identify the feasible solution for VM placement and resource allocation. The objective function was derived with the help of Stochastic Integer Programming Model. The main task of objective function was to identify and reduce the provisioning cost. The Benders Decomposition was used to break down the large problem into smaller one to find the objective function easily and minimize the time to compute the problem.

In OCRP algorithm focused to minimize the cost of resource provisioning along with security mechanisms. They taken three like predict the resources, calculate the cost of resources and demanding extra resources for reserved. This paper [11] aims to reduce the cost of access their service in reservation plan. The major disadvantage in reservation phase was not easy achieves the aim, due to uncertainty of client forecasted demand and cost of service. An Optimal Cloud Resource Provisioning (OCRP) algorithm was implemented with the help of stochastic programming approach.

The Optimal Virtual Machine Placement (OVMP) algorithm was replaced by OCRP algorithm. OVMP algorithm was applied in both the resource provisioning and VM placement. To improve the number of provisioning stages they introduced OCRP algorithm.

The OCRP algorithm provides improved result in terms of the number of provision stages and also it reduces the cost. It has two step process, first they generalized is divided into the multiple stage formulation. Second stage they consider different approaches to obtain the solution for computing resources provisioning and finally, the performance evaluation was extended to consider various realistic scenarios. When the number of scenarios was numerous, it's difficult to obtain the solution. They formulated stochastic integer programming [12] was transformed into the Deterministic Equivalent Formulation (DEF).

Pricing terms [13] are dynamically estimated by the system. Supply/demand factors are considered in the pricing policy.

The system divided into five major modules. They are cloud providers, cloud consumers, resource provisioning, price optimization and term management. Cloud providers module designed to manage resources and clients. Cloud consumers module designed to utilize resources under the providers. Resource provisioning module was designed to allocate resources based on user requirements. Price optimization module was designed to estimate resource price values. Term management module was designed to manage payment and resource usage terms.

The OCRP and HCOC algorithm [14] for computing resources in short term as well as a long-term plan. Numerical studies showed that OCRP and HCOC algorithm implement, was to find out the minimum total cost of resource provisioning in cloud computing environments. The optimal solution obtained from OCRP was obtained by formulate and solve stochastic integer programming with multistage resource. Benders decomposition approach was applied to divide OCRP problem into sub problems solved parallelly. The processing cost was almost similar in Round Robin (RR) and Equal Spread Current Execution (ESCE) scheduling algorithms [15] these parameters was bit improved in Throttled algorithm.

Cloud computing provides an dynamic infrastructure [2] media content providers utilize these resources. Media content provider collect charges for resources like bandwidth rented from the cloud. The problem identified in this paper was quantity of resources allocated within the cloud are not correctly measured and how to reduce the in reservation phase. This paper [6] explains how CSP face the market competition in terms of instance pricing in Infrastructure as a Service. The reserved instance pricing consist of on-demand price in addition to reservation fee like two part tariffs, different cloud service provider gives variety of price combinations the customer choice depends on its usage.

It ignores capacity availability and market consists of various factors. They consider only pricing and omit other factors. It ignores stable market prices in three part tariffs. Both two part and three part tariffs does not produce an optimum service pricing. It does not significantly affect from Bertrand market competition. They assume all market production was identical one.

Cloud Analyst [16] was equipped with comprehensive GUI built in Java Swing. The various metrics was used to measure the performances of the cloud system like overall response time, time of service, loading done by data center and cost of services. The ESCE algorithm [17] was improved by response time and processing time of the job. The job was allocated in a prompt balanced manner. There is no under usage of virtual machines. It helps to reduce the cost of virtual machine and the data transfer.

In the moving average method [18] constraint k made a decision on demand for the next 5 minutes interval based on the log history of the previous observations. Exponential Moving Average Model was used to determine its feasibility. The drawback identified in this paper was forecast workload more than 5 minutes time interval exceeds, it does not predict.

CONCLUSION

The study indicates that most of the research paper concentrates reservation phase and ignores on-demand phase. The cost of access the on-demand is very high when compared to reservation phase. No one study reveals how to minimize the on-demand phase. Future plan is to develop a low cost on-demand provisioning strategy. They also aim at balancing under provisioning and over provisioning of resources problems.

REFERENCES

[1] Mohammad Aazam, Eui-Nam Huh, “Cloud Customers Historical Record Based On-Demand Resource Reservation”, Architectures for Networking and Communications Systems (ANCS), ACM/IEEE Symposium,(2015) pp.207 – 208.

[2] Amr Alasaad, Kaveh Shafiee, Sathish Gopalakrishnan, and Victor C.M. Leung, “Prediction-based Resource Allocation in Clouds for Media Streaming Applications”, GC'12 IEEE Workshop: First International workshop on Management and Security technologies for Cloud Computing, (2012) pp.753 – 757.

[3] Jiang Fu, Pingping Yao, Kaiyang Liu, Jun Peng, “A Differentiated Pricing Mechanism for Idle Resources Allocation in Reservation Clouds”, IEEE Control and Decision Conference (CCDC), (2015) pp.4236 – 4241.

[4] Jonathan Chase and Dusit Niyato, “Joint Optimization of Resource Provisioning in Cloud Computing”, Services Computing, IEEE Transactions, (2015) pp.1 – 1.

[5] Yuan-Chih Yu, "The Cost-Efficient Awareness for Cloud MapReduce", 3rd IEEE International Conference on Future Internet of Things and Cloud, (2015) pp.573 - 578.

[6] Jorn Kunsemoller, Holger KarlSonja Brangewitz, Claus-Jochen Haake, "Provider Competition in Infrastructure as a Service", IEEE International Conference on Services Computing, (2014) pp.203 - 210.

[7] Sivadon Chaisiri, Rakpong Kaewpuang, Bu-Sung Lee, and Dusit Niyato, "Cost Minimization for Provisioning Virtual Servers in Amazon Elastic Compute Cloud", 19th Annual IEEE International Symposium on Modelling, Analysis, and Simulation of Computer and Telecommunication Systems, (2011) pp. 85-95.

[8] Rajkumar Buyya, Rajiv Ranjan and Rodrigo N. Calheiros, "Modeling and Simulation of Scalable Cloud Computing Environments and the CloudSim Toolkit: Challenges and Opportunities", IEEE, (2009) pp.1-11.

[9] P.Aradhya, K.Shivaranjani, "Exploring Resource Provisioning Cost Models in Cloud Computing", IJCTT, (2013) Vol.4 Issue10, pp.3512-3516.

[10] G.Menaga, S.Subasree, "Development of Optimized Resource Provisioning On-Demand Security Architecture for Secured Storage Services in Cloud Computing", IJESIT, (2013) Vol.2, Issue 3, pp.172-180.

[11] Poobalan, V.Selvi, "Optimization of Cost in Cloud Computing Using OCRP Algorithm", IJETT, (2013) Vol.4 Issue 5, pp.2105-2107.

[12] Muthulakshmi.M, Raja.G, "Optimization of Resource Cost and Service Price Scheme", IJREAT, (2013) Vol.1, Issue 1, pp.1-3.

[13] D.Poornima, K.Gandhimathi "Enhancement of Resource Pricing and Provisioning Scheme in Commercial Clouds", International Journal of Scientific & Engineering Research, (2013) Vol.4, Issue 5, pp.1205-1210.

[14] Karthika M.T., Neethu Kurian, Mariya Seby, "Comparison of Load Balancing and Scheduling Algorithms in Cloud Environment", IJITEE, (2013) Vol.3, Issue 1, pp.164-166.

[15] Tejinder Sharma, Vijay Kumar Banga, "Efficient and Enhanced Algorithm in Cloud Computing", IJSCE, (2013) Vol.3, Issue-1, pp.385-390.

[16] Bhathiya Wickremasinghe, Rajkumar Buyya, "CloudAnalyst: A CloudSim-based Tool for Modelling and Analysis of Large Scale Cloud Computing Environments", MEDC Project Report, (2009).

[17] Jaspreet kaur, "Comparison of load balancing algorithms in a Cloud", IJERA, (2012) Vol. 2, Issue 3, pp.1169-1173.

[18] Brian Bouterse, Harry Perros, "Scheduling Cloud Capacity for Time-Varying Customer Demand", Cloud Networking (CLOUDNET), IEEE 1st International Conference, (2012) pp.137 – 142.