International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

System Balancing and Cost Management in Cloud Environment

Neha A. Bhandari

M. E. Second Year Student, Computer Department, VACOE, Ahmednagar, Maharashtra, India

Abstract: ***Cloud computing is combination of multiple machines which allows resource sharing, work with large data and hardware and software sharing. In this it provide reservation plan and on-demand plan for sharing of resources. At first step consumer go with reservation phase so it is prior plan and consumer pay not as much of amount but in on-demand plan consumer get hold of resource when ever consumer want it. So consumer pay lot amount or more amount than reservation plan. In practical scenario consumer demand can get change as task go for completion. Even producer can change cost of their resources. These two possibilities can occur in cloud computing resource sharing. During work completion if any provider is performing so many task at a time then its temperature will increase and because of temperature energy wastage can happen and cost will increase so green computing concept is applied which select another provider whose temperature is under normal condition. This paper will give idea about OCRP method which reduces cost of resource allocation and green computing to save cost and monitoring is performed which make certain about system work.***

Keywords: Cloud Computing, OCRP, Green Computing, Virtual Machine, Network Monitoring

1. Introduction

Everyday people are trying and working for easy and simple life. In networking many technologies are coming and cloud computing one of them which combine some computers together and work together. During this cloud support many features like resource sharing, providing hardware and software according to requirement [6]. Cloud work on large quantity of data. Transfer of data take place from one place to another place and same in case of resource sharing, resources get transfer from one user to another user. When user want resources then it get access of them from another user and for that user pay some amount. User who use and request resources are consumer and other way is users who bring resources are provider. Use of resources is based on cost factor. This cost is calculated based on how much resources are utilized or how much time it is been used or how many resources are utilized etc. From all perspective cost gets calculated and consumer get attach with related provider for particular resources. Once consumer is attach then provider deliver resources in different ways like provider grant resources at the start step or in middle of task completion. Depend on these, delivery of resources are categories in two phases.

- Reservation Phase
- On-demand Phase

In reservation phase consumer raise resources at start only then reserved resources get used and during task completion if consumer needs extra resources then he can acquire then again from provider but for this consumer need to compensate with high amount. During such scenario sometimes under and over provisioning conditions can occur. Resources which are booked before actual task are reserved but if consumer during planning of task book more quantity of resources which are not required during task completion then it is consider as wastage of resources, this situation is called as over provisioning condition or problem. But sometimes contradictory to this can happen like reserved resources are lacking during task completion then it is considered as under provisioning condition or problem. In cloud computing we need to overcome such situation to have balanced use of resources and to pay proper amount for resources. When multiple providers are providing resources or satisfying request then from all we need to find proper provider which can deliver resources at less amount than others. To select best provider we are doing network monitoring [16]. Network monitoring check all systems regularly and provide information to user to take suitable decision. Decision can be anything like which system is having more load or which system is less usable, detect problems associated with system or between system connection, what is temperature of system and many more. Here in this system temperature of system is considerably important because it help to detect hot spot and cold spot in system. Predefined values are taken which is considering as threshold for hot spot and cold spot detection. If temperature of system is below cold threshold then it is cold machine or if any machine's temperature is more than hot threshold then it is hot machine. Selecting other machine than hot spot and cold spot is green computing [15]. When proper provider is getting selected then we are considering these factors. Along with these factors we need to consider two more factors like cost of provider for resources can get change during task completion and even consumer can need more resources in future. Select a provider who will reduce cost for resources and which is not even cold spot or hot spot in our cloud. So in our system OCRP i.e Optimal Cost Resource Provisioning is used which will reduce cost of resource allocation[1]. For reduction it will use stochastic programming to perform calculations.

2. Literature Survey

Currently, much other research is going on in Cloud Computing which grows as delivering software and hardware services and any other resources over the internet.

Some similar type of paper present related work in the area of cloud and its resource provisioning methods. Rajkamal Kaur Grewal and Pushpendra Kumar Pateriya work on A Rule-based Approach for Effective Resource Provisioning in Hybrid Cloud Environment which improve scalability of private cloud on-demand and reduce the cost, this work on time also which set the time for public cloud and private cloud to provide the services within time and fulfill the request [2]. Another Yee Ming Chen and Shin-Ying Tsai present Optimal Provisioning of Resource in a Cloud Service which work on particle swarm optimization and develop Discrete Particle Swarm Optimization (DPSO) algorithm for tasks allocation. This method effectively generates an optimal task allocation so which will complete the tasks in a minimum period of time as well as using the resources in an efficient way, this is how DPSO algorithm has ability to obtain faster and feasible allocation [3]. Rajkumar Buyya, Saurabh Kumar Garg, and Rodrigo N. Calheiros present SLA-Oriented Resource Provisioning for Cloud Computing: challenges, Architecture, and Solutions which relates information about service Level Agreement (SLA). Proposed architecture from this paper supports integration of marketbased provisioning policies and virtualization technologies allocation of resources. This concentrate on customer driven service management, QoS requirements, define computational risk management to cover risks involved in the execution of applications, incorporation of autonomic resource management, utilization of Virtual Machine (VM) [4].

Nandini Mukherjee, Sunirmal Khatua work on Application-centric Resource Provisioning for Amazon EC2 Spot Instances. This paper introduces amazon's spot instances which will offer their unused resources at lower cost as effective change but with reduced reliability. Here customers bid on unused Amazon EC2 capacity and run those instances for as long as their bid exceeds the current spot price but spot price changes continues as per supply and demand, during this whose bids exceed then he gain access to the available spot instances. But reliability is compromised since service may become unavailable at any time without any notice to the customer. For this different check pointing schemes are used which increase reliability and decrease cost [7]. Amazon EC2 give less cost if resources are fully utilized which are from reservation plan [5].

Yanzhi Wang, Yue Gao, Sandeep K. Gupta and Massoud Pedram present An Energy and Deadline Aware Resource Provisioning, Scheduling and Optimization Framework for Cloud Systems, which consider large-scale, heterogeneous, multi-user environment cloud system, and profit maximization for the cloud service provider (CSP). During this they consider workloads are modeled as a collection of multiple task graphs with output dependencies, cloud platform is modeled as a weighted graph, request VMs and its placement, and CSP addresses deadline-aware resource provisioning, task scheduling and energy cost optimization in a holistic fashion [8].

Inderveer Chana, Rajni Aron work on Cost based Resource Provisioning Policy for Grids. This provides resource provisioning policy in XML schema which satisfies quality of service and minimize cost of resources which are required for execution [9]. Ian Foster, Ioan Raicu, Yong Zhao, Catalin Dumitrescu, Mike Wilde present Dynamic Resource Provisioning in Grid Environments which give idea about resource provisioning with various polices for allocation and deallocation. It makes use of batch scheduler to manage different clusters which receive individual tasks, dispatch them to idle and after finishing its task inform back to client [10].

Yang-suk Kee, Carl Kesselman work on Grid Resource Abstraction, Virtualization, and Provisioning for Time-targeted Applications which provide introduction of resource slot which give information about start time, duration, size and vector of attribute about resources. It is high level representation of resource in which start time and duration gives resource availability and its size provides spatial resource availability and finally attributes include different factors like processor type, processor clock speed, and memory capacity, software installations with some extra like cost, availability and reliability. This also specifies constraints on logical machines with their connectivity. Due to this resource slot, easily resource allocation is done at some time and utilizes it for certain duration [11].

Nagarajan Kandasamy and Dara Kusic present Risk-Aware Limited Lookahead Control for Dynamic Resource Provisioning in Enterprise Computing Systems. An optimization framework is develop for self managing behavior which support multiple QoS and Limited Lookahed control approach (LLC) is used to solve resource management as a sequential optimization with uncertainty. It makes use of predictions, handle explicit constraint [12].

3. System Architecture

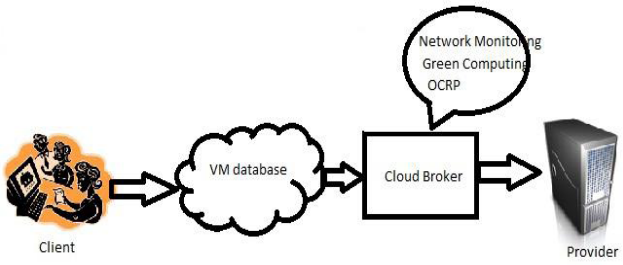


Figure 1: System Architecture

Our system is combination of multiple modules which form cloud environment. Cloud is form means some computers will provide resources on the requirement of other computers. Along with consumer and provider, we have broker and VM repository in our Cloud. VM repository is space or database to store create VM. VM is virtual Machine which will get create on requirement of consumer. As per requirement of consumer related VM allocation happen with request and for that we have VM repository. Here broker verify created VM from repository and start working. Broker performs network monitoring to get status of machines, from status hot spot and cold spot get identified. After all this checking OCRP method is applied to reduce cost of resource allocation. And VM get attach with proper provider which

give less cost for resource allocation and whose temperature is under normal machine temperature range.

OCRP method works by calculating cost of phases. Here we have three phases namely reservation phase, expanding phase and on-demand phase. Reservation phase reserve resources at start of work and then consumer will consume them in expanding phase. Expanding phase is operation phase of resources. This phase gets continue till consumer has resources. Once need of extra resources arrives then on-demand support extra resources. Like this individual cost all three phases get calculated and final cost of resource allocation get decided. This final cost equation get form by stochastic integer model. While forming cost equation we consider demand and cost from provider as changeable parameters. This total cost is combination of cost of three phases namely reservation cost, expanding cost, on-demand cost and cost required to reserve resources at first provision.

Reduce cost

F. Cost = first.provision.cost + reservation.cost + expanding.cost + on-demand.cost

This cost is depend on number of VM created, number of resources allocated in particular phase. Then this get solve by benders decomposition method. Benders decomposition will divide equation into master equation and sub equation of master equation.

4. Algorithm

Algorithm starts with network monitoring of system, apply green computing concept and finally fond cost of resource allocation.

1. Perform network monitoring, check parameters of system and find out if system is overloaded or underloaded.
2. If system is overloaded or underloaded then take any action to have balancing in network.
3. Then perform green computing method. In this check temperature of system.
4. If temperature > hot threshold then hot spot.
5. If temperature < cold threshold then cold spot
6. Find all hot spot and cold spot and select normal machine to allocate resource.
7. Once provider is selected then calculate cost.
8. Minimize cost
   F.Cost = first.provision.cost + reservation.cost + expanding.cost + on-demand.cost
9. Solve this problem by dividing equation into master and sub-problem and solve these parallel using stochastic integer programming [13][14].
10. Allocate resources to respective consumer.

5. Model Formation

Cloud is prepared with consumer, provider, and broker and VM database. As per requirements VM gets select. After this we can have green computing which will check temperature of system if it is desirable then we can have OCRP for that provider.

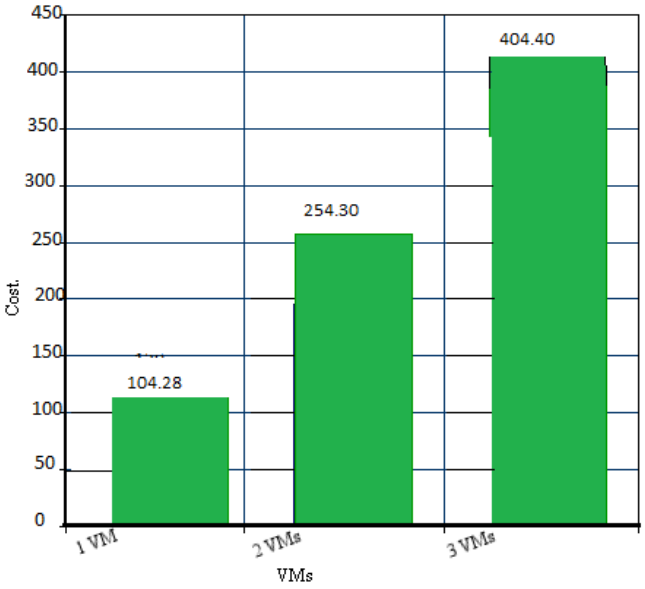
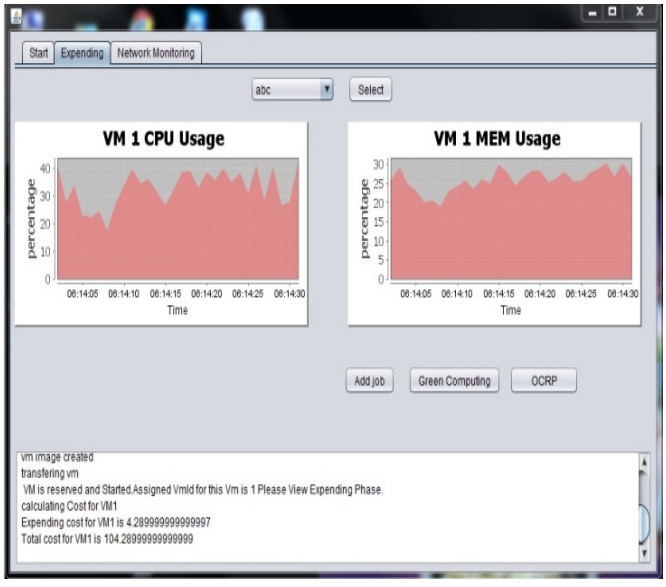


Figure 2: Cost Calculation

This graph is number of VM versus Cost. If requirement of consumer increases the number of VM get increase. VM has its cost associated with it and it is based on its type. Graph shows if VM number increases then cost will get increase.



One snapshot is given here which is cloud broker. In start tab, broker is performing reservation of VM and in expanding phase consumer will utilize resources. Two graphs in figure give idea about how much memory and CPU is used. When we add new job into this then it will also take some more resources and finally OCRP will calculate cost of this entire step.

6. Conclusion

Thus we have studied cost reduction and having balance in network. Consumer ask resources and provider provide them. In cloud reservation plan and expanding plan are available to allocate resources. In our system we have reservation phase, expanding phase and on-demand phase to

allocate resources. During every phase of allocation cost is associated and final cost is calculated from combination of all three phases. Then OCRP is applied to reduce cost of allocation. During this entire network monitoring always keep its task in running mode. It regularly monitors system and report identified parameters. From this parameter green computing is check. It identifies hot machine and cold machine. So proper provider is selected which is not hot spot and cold spot. And from graph we got clear idea about cost reduction because of OCRP and Green computing.

7. Acknowledgement

I like to thanks all people for their importance guidance about project work. I am thankful to those who help me during my paper publishing and project dissertation. I am thankful to journal for giving me opportunity to publish my paper. I am sincerely thankful to my head of department, my guide and other staff members from department for supporting me.

References

[1] Sivadon Chaisiri, Bu-Sung Lee, and Dusit Niyato, Optimization of Resource Provisioning Cost in Cloud Computing, VOL. 5, NO. 2, APRIL-JUNE 2012.
[2] SRajkamal Kaur Grewal and Pushpendra Kumar Pateriya, A Rule-based Approach for Effective Resource Provisioning in Hybrid Cloud Environment, International Journal of Computer Science and Informatics ISSN (PRINT): 2231 -5292, Vol-1, Iss-4, 2012.
[3] Yee Ming Chen, Shin-Ying Tsai, Optimal Provisioning of Resource in a Cloud Service, IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 6, November 2010.
[4] Rajkumar Buyya, Saurabh Kumar Garg, and Rodrigo N. Calheiros, SLAOriented Resource Provisioning for Cloud Computing: Challenges, Architecture, and Solutions, 2011 International Conference on Cloud and Service Computing.
[5] Amazon EC2, http: //aws.amazon.com/ec2.
[6] Ian Foster, Yong Zhao, Ioan Raicu, Shiyong Lu, Cloud Computing and Grid Computing 360-Degree Compared.
[7] Sunirmal Khatua, Nandini Mukherjee, Application-centric Resource Provisioning for Amazon EC2 Spot Instances.
[8] Yue Gao, YanzhiWang, Sandeep K. Gupta and Massoud Pedram , An Energy and Deadline Aware Resource Provisioning, Scheduling and Optimization Framework for Cloud Systems.
[9] Rajni Aron, Inderveer Chana, Cost based Resource Provisioning Policy for Grids, Proceedings of the World Congress on Engineering 2011,Vol II.
[10] Ioan Raicu, Yong Zhao, Catalin Dumitrescu, Ian Foster, Mike Wilde, Dynamic Resource Provisioning in Grid Environments.
[11] Yang-suk Kee, Carl Kesselman,Grid Resource Abstraction, Virtualization, and Provisioning for Time-targeted Applications.
[12] Dara Kusic and Nagarajan Kandasamy ,Risk-Aware Limited Lookahead Control for Dynamic Resource Provisioning in Enterprise Computing Systems.
[13] Shabbir Ahmed, Two-stage stochastic integer programming:A brief introduction.
[14] George B. Dantzig and Gerd Infanger,Large-Scale Stochastic Linear Programs: Importance Sampling and Benders Decomposition, TECHNICAL REPORT SOL 91-4,March 1991.
[15] Zhen Xiao, Senior Member, IEEE, Weijia Song, and Qi Chen, Supplementary File: Dynamic Resource Allocation using Virtual Machines for Cloud Computing Environment.
[16] Zenon Chaczko, Venkatesh Mahadevan, Shahrzad Aslanzadeh and Christopher Mcdermid, Availability and Load Balancing in Cloud Computing.