

MANAGING COST AND PERFORMING BALANCING AT CLOUD PLATFORM

Neha A. Bhandari

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

Abstract

Group of computers for performing some function which are beneficial for users is called as Cloud computing. Cloud computing provide more facility like reduce cost of components, low cost for resource allocation and perform multiple operation on huge quantity of data. Out of these many facilities, cloud gives reservation and on-demand facility for resource allocation. In which reservation provide resource at initial implementation and on-demand provide according to user wish. In this sometimes cloud come across over-provisioning and under-provisioning situations and reducing cost of resource allocation is also important so with these many issues Optimal Cloud Resource Provisioning OCRP algorithm help to reduce cost and handle previous situation. Stochastic integer model and benders decomposition algorithm is used for cost reduction. Along with this, balancing is performed in cloud by checking bandwidth use and temperature of machines. These values get checked and as per situation some actions take place to do balancing in network.

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

1. INTRODUCTION

Cloud Computing is collection of multiple computers for providing facilities to different users of cloud. Mainly there are three types of cloud namely public cloud, private cloud and hybrid cloud. Public cloud is accessible to everyone but private cloud is restricted to certain users and hybrid cloud is combination of both public cloud and private cloud. Cloud computing act like distributed environment means any user can use cloud facility from any location and work on huge data [13]. Cloud give three services to user namely Platform as a Service (PAAS), Software as a Service (SAAS) and Infrastructure as a Service (IAAS) [2]. PAAS provide platform service in which platform is supplied to user to run application, to check output etc. Then SAAS provide software service means user can rent needed software instead of buying it forever. And finally IAAS give infrastructure service in which Virtual platform concept got emerge, user can get virtual things to work like storage, servers etc.

Reservation and on-demand plans are used in cloud computing to share resources in between them. In reservation, its name only suggest that user need to perform reservation at initial stage then use that reserved resources in future work so for this lees amount user need to pay but sometimes user need more resources in future so for that on-demand plan works. In on-demand plan user order as per task requirement and pay more amount. During reservation if user reserve extra resources which are not needed in future so it is wastage of resources and this situation is called as over-provisioning of resources and opposite to this is under-provisioning situation. In case of under-provisioning, user of cloud gets need of extra

resources other than resources which are book in reservation plan. Means resources which are book in reservation are not sufficient to task so user need extra. During all this, requirement of user is always in changing fashion and even cost of resource provider can vary so demand of user and cost of provider is consider as a fluctuating factor or uncertainty parameters of cloud environment. Remembering this entire uncertainty factors and above mention situation of cloud we need to reduce cost of resource allocation in cloud environment. OCRP Optimal Cloud Resource provisioning algorithm considers these entire scenario and reduce cost of resource allocation optimally [1].

With reduction of cost for resource allocation, balancing in network is also important. Continuous network monitoring helps us to find out load on each system and then according to load we are performing balancing in network. During monitoring we are also checking temperature of all machines, from this temperature we can find out hot spot and cold spot from network. If machines temperature is more than some predefined value then it is consider as hot spot and if its temperature is below than some other predefined then it is consider as sold spot. Remove these hot and cold spot from network will reduce energy consumption in network. So goal of our proposed system is performing balancing in network, reducing energy consumption and reducing cost of resource allocation in cloud environment.

2. RELATED WORK

Much research is going on in this field which gives related knowledge about our system. Yang-suk Kee, Carl Kesselman propose resource provisioning with respect to time [3]. New concept is rising in this which presents logical machine and resource availability time according to user perspective. This give starting time of resource allocation, how long it is going to continue, related attribute of this and size or width of resources. Yee Ming Chen and Shin-Ying Tsai propose discrete particle swarm optimization which performs meta-heuristic optimization [4]. This algorithm minimizes execution cost and communication cost in between provider and consumer. Rajni Aron, Inderveer Chana introduce [5] Cost based Resource Provisioning Algorithm which increase profit of consumer. In this authorization and authentication is provide by GSI Grid Security Infrastructure. Resource Manager (RM) take care of resource allocation, check accessibility etc. XML schema is used to present resource allocation policy. Ioan Raicu, Yong Zhao, Catalin Dumitrescu, Ian Foster, Mike Wilde introduce some policy related to resource acquisition and resource release [6]. It makes use of adaptive provisioning, reduces waiting time for queue in allocation and introduces dynamic resource provisioning. Sivadon Chaisiri, Bu-Sung Lee, and Dusit Niyato [7] introduce OVMP Optimal Virtual Machine Placement which reduces cost of allocation with satisfaction of consumer and provider. It presents EVF expected value formulation which considers average value of fluctuating parameters to solve cost formula. Laura Grit, David Irwin, Aydan Yumerefendi, Jeff Chase [8] introduces virtualization concept with its variation in implementation technique. Zhen Xiao, Weijia Song, and Qi Chen [9] propose skewness algorithm which predict load, implement hot spot migration and perform cold spot removal. For hot spot removal BG algorithm is introduce which collect load information into a metrics. As per this matrix virtual machine get allocated to another machine. Sandeep Bhargava, Swati Goyal [10] give idea about various types of virtualization methods like full virtualization and paravirtualization technique. Limited Lookahead Control (LLC) is used to solve provisioning and management of resources in paper [11]. LLC make use of prediction, perform optimization in discrete domain, handle various constraints and also satisfy quality of control. Zenon Chaczko, Venkatesh Mahadevan, Shahrzad Aslanzadeh and Christopher Mcdermid give brief concept on load balancing [12]. Load balancing is performed to improve performance, reduce delay and long response time etc. [12] give information about some static and dynamic load balancing algorithm.

3. SYSTEM DESIGN

Our system is made of many several components namely cloud consumer, cloud provider, cloud broker and virtual machine (VM) repository. Cloud consumer is one who wants to access resources or use resources to accomplish its task. Consumer places its requirement to access resources. As per

requirement Virtual machines will get form and VM repository will store created VM. Cloud provider will supply required resources to consumer. For this resource consumer pay some amount.

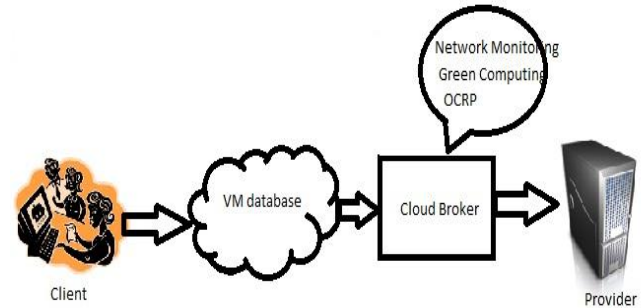


Fig. 1 System View

Cloud broker works like broker from real life. This broker will check requirement of consumer and find proper provider who can satisfy consumer. Will finding particular provider, broker perform network monitoring, check bandwidth access of all machines in network, find hot and cold spots from network and then allocate VM to that particular provider. While doing all this it also consider uncertainty parameters of system and reduce cost of resource allocation. Cost is depending on plans and phases of resource allocation. If timing factor is considered with plans then it is called as phase of allocation. Total three phase are there. Reservation phase is from reservation plan and when consumer use resources from reservation plan then it named as expanding phase. And finally on-demand phase is when consumer order more resources in on-demand plan. Provider offer all these phases to consumer and along with this provide multiple stages of resource allocation, multiple offers on resources, contracts etc. Cost of each phase is calculated from number of resources consumer is accessing in that phase, number of virtual machines is created, and consumer's selected contract and stage. But cost of expanding phase should be less than cost of reservation phase. Total cost of allocation is form from all these three phases and cost required to reserve resources at first reservation. At first reservation stage consumer is totally unaware of uncertainty parameters of system. Stochastic integer programming is applied to form equation of total cost in this algorithm [14][15], it observe uncertainty parameters and then give minimum cost.

$$TCost = Cost_first_provisioning + Cost_Reservation_Phase + Cost_Expanding_Phase + Cost_On-demand_Phase$$

So our proposed system is used to reduce cost of this resource allocation. But this equation of total cost is bound with some constraints. Constraints like consumer demand should get satisfy, first reservation is totally free from uncertainty parameters, amount of resources in reservation phase is more

than in expanding phase, and quantity of total resources is less than all allocated resources. Once constraints are checked then benders decomposition algorithm is applied. This will divide original problem into many small sub problems so that we can easily solve these entire problem.

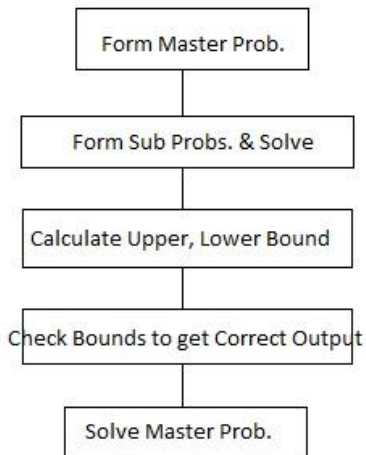


Fig. 2 Benders Decomposition

Benders decomposition [15] starts with creating master problem. Generally one technique is applied to form master problem. Select particular step for master problem which is having more constraints so in our case expanding phase is having more constraints so expanding phase cost is consider as a master problem and remaining parameters from total cost formula will form other two sub problems. Later in next step solve sub-problems which are formed now. Next is to find upper bound and lower bound. This upper bound is calculated from sub-problem which is solved in previous step and lower bound is calculated from master problem. To check this bounds one tolerance value is decided and if difference of upper bound and lower bound is less than tolerance means if they are close then it is consider as a correct output otherwise we need to work for next iterations. And last step is to solve primarily detected master problem. Like this bender decomposition will solve total cost problem easily and we can reduce total cost of resource allocation.

Network Monitoring is continuous process which is going on in our system. In this we are going to check different parameters of system like bandwidth, access of data, temperature of system etc. In any system transfer of data from one machine to another is common so it need bandwidth for data transfer or data access. Bandwidth will identify if any machine or overloaded or not. If its data access is more than some fixed value then it is called as overloaded but if its bandwidth is less than fixed value then it is normal work mode of system. Like this machines get classified and then overloaded machine get normalized by removing its load or by

moving its load to different machine. While doing this care need to take that other machine should not get overloaded.

Green Computing is applied to restrict consumption of energy in network. Here network monitoring will check temperature of all machines and then classified machine either into hot spot, cold spot or normal machine. To classify some fixed values are consider as tolerance values. If hot and cold spot is detected then remove them to reduce energy consumption. Hot spot is like whose temperature is more than threshold or it is performing more task so required more VM so its temperature is more. And cold spot is like which is not doing any work, just like idle machine so either keeping it on we can remove cold spot to save energy consumption. Some techniques are applied to remove them. Find another destination machine to transfer load or VM of hot and cold spot but we need to check that destination machine should not get hot again. After finding destination machine we can move VM of hot/cold spot. Or we can terminate VM from machines or if nothing is working correctly then we can shutdown physical machine itself. Like this green computing is applied to save energy in network and along with monitoring we can add efficiency in cost reduction.

4. CONCLUSIONS

Like this in our proposed system we are implementing OCRP Optimal Cloud Resource Provisioning, network monitoring and green computing. Network monitoring will monitor system and find out normal heavy or light loaded system. By moving load of heavy machine to light loaded machine we are going to perform balancing in network. Green computing is applied to reduce consumption of energy in network. It detect hot spot and cold spot and its removal will reduce consumption of energy usages. After balancing and energy saving, we will reduce cost of resource allocation by OCRP. It initially detects total cost formula from finding cost of all individual phases and plans. This total cost formula get solve by benders decomposition algorithm which will reduce difficulty in solving cost formula and find proper solution for cost. It divide problem into sun problems and give us minimized cost for resource allocation. While reducing cost we are considering uncertainty parameters of system which are demand of consumer which is always in changing fashion and cost of provider which is also in changing fashion.

ACKNOWLEDGEMENTS

I sincerely thanks to all those who supported in my work directly or indirectly I thanks to Mr. Mandar C. Kshirsagar for his valuable guidance.

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]. 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
- [3]. Yang-suk Kee, Carl Kesselman,Grid Resource Abstraction, Virtualization, and Provisioning for Time-targeted Applications
- [4]. 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
- [5]. Rajni Aron, Inderveer Chana, Cost based Resource Provisioning Policy for Grids, Proceedings of the World Congress on Engineering 2011 Vol II
- [6]. Ioan Raicu, Yong Zhao, Catalin Dumitrescu, Ian Foster, Mike Wilde , Dynamic Resource Provisioning in Grid Environments
- [7]. Sivadon Chaisiri, Bu-Sung Lee, and Dusit Niyato, Optimal Virtual Machine Placement across Multiple Cloud Providers
- [8]. Laura Grit, David Irwin, Aydan Yumerefendi, Jeff Chase Virtual Machine Hosting for Networked Clusters: Building the Foundations for .Autonomic. Orchestration
- [9]. Zhen Xiao, Senior Member, IEEE, Weijia Song, and Qi Chen, Supplementary File: Dynamic Resource Allocation using Virtual Machines for Cloud Computing Environment
- [10]. Sandeep Bhargava, Swati Goyal, Dynamic Load Balancing in Cloud Using Live Migration of Virtual Machine
- [11]. Dara Kusic and Nagarajan Kandasamy ,Risk-Aware Limited Lookahead Control for Dynamic Resource Provisioning in Enterprise Computing Systems
- [12]. Zenon Chaczko 1, Venkatesh Mahadevan 2, Shahrzad Aslanzadeh 1 and Christopher Mcdermid, Availability and Load Balancing in Cloud Computing
- [13]. Ian Foster, Yong Zhao, Ioan Raicu, Shiyong Lu, Cloud Computing and Grid Computing 360-Degree Compared
- [14]. Shabbir Ahmed, Two-stage stochastic integer programming:A brief introduction
- [15]. George B. Dantzig and Gerd Infanger, Large-Scale Stochastic Linear Programs: Importance Sampling and Benders Decomposition, TECHNICAL REPORT SOL 91-4, March 1991