A Detailed Study on Virtual Machine Consolidation Techniques in Cloud Computing

[1] Anusuya G, [2] John Peter S
[1] Research Scholar,ManonmaniamSundaranar University,Tirunelveli, Tamilnadu
[2] Associate Professor and Head,Department of Computer Science ,
St.Xavier's College,Palayankottai, Tirunelveli,Tamilnadu
[1]anushyabalamurugan@gmail.com,[2] jaypeeyes@rediffmail.com

Abstract:

Virtual machine consolidation techniques have been studied for a long time in the domain of cloud computing. The goal of VM consolidation is to minimize energy consumption by switching servers to the sleep mode, when they are not utilized as expected. This paper aims to provide a detailed study of the existing VM consolidation approaches and outline the notable advantages and disadvantages. Furthermore, it highlights that VM placement sub problem for an under loaded host and an overloaded host should be handled in a different way.

Index Terms: virtual machine consolidation, virtual machine placement

1. Introduction

Recent research efforts have explored the depth of virtualization in cloud computing. One of the major advantage and achievement of virtualization is green computing. Potential virtualization techniques are believed to increase the server utilization rates. Virtual machine consolidation techniques prove to help in building and managing ecofriendly virtualized data centers. Virtual machine consolidation objectives range from simple issues like minimizing resource wastage, minimizing SLA violations, maximizing resource utilization to complex issues that include minimizing data center power cost and maximizing network size.

Virtual machine (VM) consolidation is mainly achieved by live VM migration and switching off the underutilized physical machines. Live VM migration involves selecting a VM to be migrated and a target host on which the selected VM is to be placed. Switching off the underutilized physical machines requires detection of such hosts.

2. Related Work

Virtual machine consolidation is aimed at building and managing anenergy efficient dynamic cloud datacenters. It is based on the concept of packing as manyvirtual machines into fewer number of physical machines. It aims at achieving the dual objective; maximizing resource utilization and minimizing power consumption. From the existing literature, VM consolidation is found to be a NP hard problem and a wide range of heuristic, probabilistic and genetic algorithms have been proposed.

Farahnakian et al. [6] have applied reinforcement learning based method that minimizes the number of physical machines due to the current resources requirement. The proposed method learns from past knowledge to decide the time when a physical machine should change its state. The advantage is that no prior knowledge or prediction of workload is required.Farahnakian et al. [7] have proposed k-nearest neighbor regression algorithm that predicts the amount of future CPU usage thereby minimizing the number of physical servers on a data center. The objective of this method is twofold, reducing SLA violation and reducing energy cost.

Ashraf et al. [8] have presented a novel ant colony system algorithm that devises VM migration plans. VM migration plans are used for minimizing overprovisioning of physical machines by migrating VMs on underutilized ones. They have focused on maximizing the number of released physical machines. Feretto et al. [9] have proposed an LP formulation and heuristics to control VM migration that gives priority to virtual machines with stable capacity. But the proposed method will enable the migration only if the VM demands a change in its capacity. Hwang et al. [12] have formulated the VM consolidation problem as a multi capacity stochastic bin packing problem and have tried to solve it using heuristics.

Feller et al. [15] have proposed fully decentralized dynamic VM consolidation schema based on peer to peer (P2P) network of virtual machines. The proposed method periodically consolidates VMs on a least number of physical machines. The advantage of this method is that it supports increasing number of physical machines as well as virtual machines.Sedaghat et al. [17] have investigated a decentralized approach using P2P consolidation protocol. They have considered various dimensions of resources and unpredictable environment. The advantage of this method is twofold;concurrency and distributed control.

2.1 Live Migration

Live Migration of a virtual machine may also lead to a change in its resource requirements [1]. Live migration is performed to avoid SLA violation that is caused by an overloaded host. VM migrations also consume timecreating overheadswhich may lead to SLAviolation [2]. VM migration being the integral part of VM consolidation, are unavoidable. The number of VM migrations should be kept minimum.

2.2 Load Prediction and Detection

Belaglazov et al. [4] have presented a novel approach that detects an overloaded host by maximizing the mean inter- migration time under the QoS goal. The approach is based on a Markov chain model and it is found to be optimal for any known, stationary workloads. The proposed approach is heuristically adapted to handle unknown non stationary workloads based on the multi size sliding window estimation technique. Khanna et al. [10] have proposed a workload management algorithm that is triggered when a physical server becomes overloaded or under loaded. The algorithm aims at minimizing the number of physical servers. The algorithm also guarantees that SLAs are not violated and migration costs are minimized.

Belaglazov et al. [1, 3] have shown that to improve the quality of VM consolidation, maximizing the time intervals between VM migrations from overloaded hosts is important. The authors have also proposed adaptive heuristics due to the variability of VMs' load for detecting overloaded physical machines. For this, both threshold based and prediction based methods are considered; Median Absolute Deviation(MAD), Interquartile Range(IQR), Local Regression(LR), Robust Local Regression(RLR). The performance of the algorithms that detect overloaded physical machines using the above said methods are evaluated using CloudSimframework [18].The authors have also performed statistical analysis and concluded that host overload detection algorithms based on local regression outperform the threshold based algorithms due to a decreased number of VM migrations and SLA violations.

2.3 VM Selection

When an overloaded host is detected, the candidate VMs are to be selected for migration. Khanna et al. [10] defined a VM selection criterion based on CPU utilization of VMs. They proved that selecting VMs based on this criterion can result in minimized migration costs. Masoumzadeh et al. [16] have proposed a fuzzy Q-Learning technique so as to make optimal decisions to select virtual machines for migration. The approach integrates multiple VM selection criteria to benefit from all advantages and possible contributions in long term learning.

Belaglazov et al. [1] have considered severeal heuristics: The Minimum Migration Time Policy(MMT) migrates a VM that requires the minimum time to complete a migration relatively to the other VMs residing in the physical machine. The Random Selection Policy(RS) randomly selects a VM to be migrated from the overloaded physical machine. The Maximum Correlation Policy(MCP) is based on the idea of selecting candidate VMs that have the highest correlation of the CPU utilization with the other VMs. All three selection algorithms have been simulated and followed by a detailed statistical analysis. From the presented algorithms, MMT algorithm has been proven the best.

2.4 VM placement

Once the candidate VMs are selected for migration, then comes the sub problem of VM placement. Several VM allocation mechanisms have been found in the existing literature. Few authors have considered the VM placement problem as a special version of the bin packing problem.Belaglazov et al. [3] have suggested Modified Best Fit Decreasing heuristic, which considers the VMs in decreasing order of load and places each of them to the physical machine with best energy efficiency as well as sufficient capacity to host it.

Verma et al. [5] develoedpMapper architecture that manages virtualized resources that are heterogenous in nature. The authors aimed at optimizing the mapping of VMs to PMs with respect to energy consumption and number of migrations. They presented three algorithms namely Min Power Parity, Min Power Placement with History and pMap. The first algorithm offered best energy efficiency whereas the second and third algorithms aimed at avoiding unnecessary migrations thereby decreasing the number of migrations.

3. Conclusion

We have presented a detailed study on the VM consolidation techniques. The heterogeneity nature and extreme dynamicity of the domain under study led to the lack of fruitful comparative analysis. We also emphasize that VM placement algorithms for under loaded and overloaded physical machines must be differentiated. The former case, there is no need of VM selection and in case if a suitable target physical machine is not available, the underloaded physical machine can be prioritized to accommodate VMs from other overloaded hosts as FFD has been proven to have good theoretical bounds. Moreover the under loaded physical machine is not prone to SLA violations. The latter case, candidate VMs should be selected and it demands a tradeoff between SLA violation cost and migration cost.Such algorithms that handle both the cases differently are yet to be developed and analyzed. With the ever growing density and complexity of the cloud datacenters, VM consolidation techniques that gives a better tradeoff between energy and performance continue to be emerging.

References:

[1] A.Belaglazov,R.Buyya. "Optimal online deterministic algorithms and adaptive heuristics for energy and performance sufficient dynamic consolidation of virtual machines in cloud data centers", Concurrency and Computation: Practice and Experience, Vol. 24, no. 13, pages 23-50, 2011.

[2] I.Rodero, H.Viswanathan,E.K.Lee, M.Gamell,D.Pompili, M.Parashar. "Energy- efficient thermal aware autonomic management of virtualized HPC cloud infrastructure", Journal of Grid Computing, Vol. 10, no. 3, pages 447-473, 2012.

[3] A.Belaglazov, J.Abawajy, R.Buyya. "Energy aware resource allocation heuristics for efficient management of data centers for cloud computing", Future Generation Computer Systems (FGCS), Vol.28, no.5, pages 755-768, 2012.

[4] A.Beloglazov, R.Buyya, "Managing Overloaded Hosts forDynamic Consolidation of Virtual Machines in Cloud Data Centers UnderQuality of Service Constraints," IEEE Transactions on Parallel and DistributedSystems , Vol. 24, no. 7, pages 1366-1379, 2013.

[5] A.Verma,P.Ahuja, A.Neogi, "pMapper: Power and migration cost aware application placement in virtualized systems", Middleware 2008, pages 243-264, 2008.

[6]F.Farahnakian, P.Liljeberg, J.Plosila, "Energy-efficient virtual machines consolidation in cloud data centers using reinforcement learning", Parallel, Distributed and Network-based Processing – 22ndEuromicro International Conference, pages 500-507, 2014.

[7] F.Farahnakian, T.Pahikkala, P.Liljeberg, J.Plosila, " Energy aware consolidation algorithm based on k-nearest neighbor regression for cloud data centers", Utility and Cloud Computing, 6th IEEE/ACM International Conference,2013.

[8]A.Ashraf, I.Porres, "Multi objective dynamic virtual machine consolidation in the cloud using ant colony system", International Journal of Parallel, Emergent and Distributed Systems, Vol.33, no.1, pages 103-120.

[9] T.C.Ferreto, M.A.Netto, R.N.Calheiros, C.A.D. Rose, " Server consolidation with migration control for virtualized data centers", Future Generation Computer Systems, Vol.27,no.8, pages 1027-1034, 2011.

[10] G.Khanna, K.Beaty, G.Kar, A.Kochut, "Application Performance Management in Virtualized Server Environments", Proceedings of the 10th IEEE/IFIP Network Operations and Management Symposium, pages 373-381,2006.

[11] Y.Gao, H.Guan, Z.Qi, Y.Hou, L.Liu, " A Multi objective ant colony system algorithm for virtual machine placement in cloud computing" , Journal of Computer and System Sciences", Vol.79,no.8, pages 1230-1242, 2013.

[12] I.Hwang, M.Pedram, "Hierarchical virtual machine consolidation in a cloud computing system", Cloud Computing 2013 IEEE International Conference, pages 196-203,2013.

[13]J.V.Wang, N.Ganganath, C.K.MichaelTse, "A heuristics based VM allocation mechanism for cloud data centers", IEEE International Symposium on Circuits and Systems, pages 1-4, 2017.

[14] F.Farahnakian, A.Ashraf, T.Pahikkala, P.Liljeberg, J.Plosila, et al. "Using ant colony system to consolidate VMs for green cloud computing", Services Computing, IEEE Transactions on Services Computing", Vol.8, no.2, pages 187-198, 2015.

[15] E.Feller, C.Morin, A.Esnault, "A case for fully decentralized dynamic VM consolidation in Clouds", IEEE International Conference on Cloud Computing Technology and Science (Cloudcom)",pages 26-33, 2012.

[16] S.S.Masoumzadeh, H.Hlavacs, "Integrating VM selection criteria in distributed dynamic VM consolidation using fuzzy Q-learning", Network and Service Management, 9th International Conference, pages 332-338, 2013.

[17] M.Sedaghat, F.Hernandez, Rodriguez,E.Elmroth, S.Girdzijauskas, "Divide the task, multiply the outcome: Cooperative VM consolidation", Computing Technology and Science (Cloudcom), IEEE 6th International Conference, pages 300-305, 2014.

[18] N.Calheiros, R.Ranjan, A.Belaglazov, C.A.F.D. Rose, 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, pages 23-50, 2011.