A Survey on Dynamic Resource Allocation For Virtual Machine In Cloud

Kachhadiya Khushbu V.[1], Gayatri S Pandi(Jain)[2]

[1] *Master Of Engineering, Computer Engineering, L.J.I.E.T, Gujarat, India*
[2] *Head Of Department, Computer Engineering, L.J.I.E.T, Gujarat, India*

ABSTRACT

Cloud computing becomes well liked among cloud users by contribution of various resources. Cloud computing provides on-demand access to computational resources which together with pay-per use business models, enable application providers seamlessly scaling their services. Virtualization technology provides server resources allocation dynamically based on the application request. In this paper, we use hybrid cloud environment for managing workload like MapReduce. VM migration provide high availability and avoid potential performance degradation.

Keyword: - *Cloud Computing, Dynamic Resource Allocation , VM migration*

1. Introduction

Cloud computing is rapidly being adopted by business and IT organizations since it offers a high degree of scalability, convenient pay-as-you-go services, device independency, location independency and low cost computing. Cloud Computing can be defined as "On demand delivery of infrastructure, applications, and business processes in a security-rich, shared, scalable, and based computer environment over the Internet for a fee"

In cloud computing, Resource allocation is the process of appointing available resources dynamically to the required cloud applications. In cloud the resource allocation is based on the infrastructure as a service (IaaS) [1]. In cloud environment, VMs are usually configured during the initialization with the amount of resource (CPU, memory) specified. Resource over-provision is one of the solutions to lower the SLA violation from users point of view but usually leads to poor infrastructure utilization. In contrast, underprovision could lead to potential performance degradation despite the increased utilization [2]. Static allocation cannot deliver good results if most VMs exhibit a similar workload behavior. Dynamic resource allocation is able to cover such scenarios. It periodically reallocates VMs to servers. Live-migration technology is used to move a VM during operation from one server to another without noticeable service disruptions [3].

2. Related Work

In first paper, Mohammadreza Mesbahi proposed resource management system we design resource allocation algorithm which dynamically allocated resource for cloud user in server. In this proposed work for effective resource allocation used adapt load balancer which presented below in algorithm which makes decision on the basis

of skew value of all VMs and future predicted load on server. Scheduling algorithm increases the efficiency in real time cloud computing services. Skewness is used to measure the uneven utilization of multiple resources of each VMs. Skewness measured by Hot and Cold spot. load prediction algorithm predict the CPU load and measure the load every minute using past history.[1]

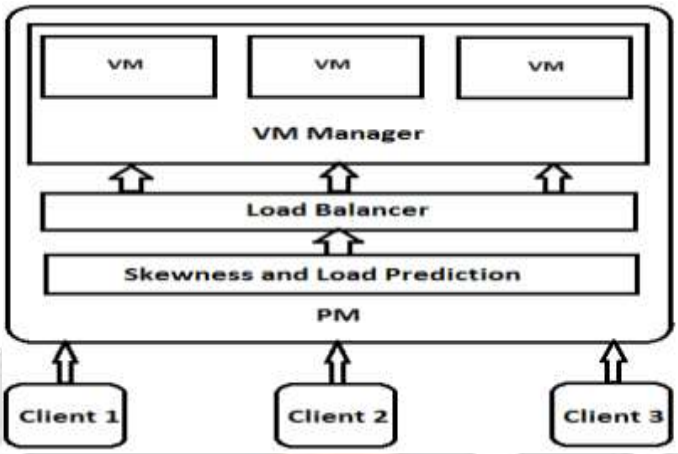

Fig -1: Dynamic resource allocation system on single PM

In second paper, Yuda Wang proposed dynamic resource allocation system is proposed which combine traditional batch processing task with long running VM workload in hybrid cloud computing. VM migration is used to ensure high availability and avoid potential performance degradation. A new protocol is proposed to deal with the VM placement problem and we conduct optimizations to reduce the serial requests waiting time. they build a dynamic resource allocation mechanism for VM in order to improve the holistic cluster utilization.[2]

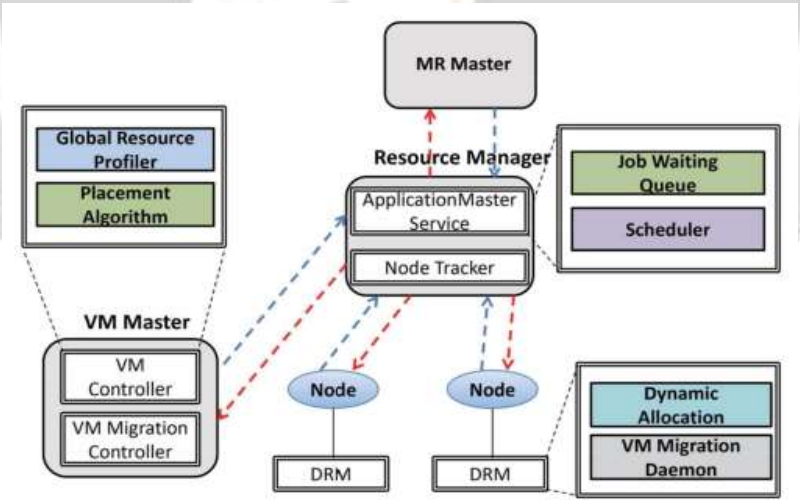

Fig -2: System Architecture

In third paper, Andreas Wolk evaluate the applicability of DSAP in a deterministic environment. DSAP is a linear program, calculating VM allocations and live migrations on workload patterns known a priori. DSAP can increase consolidation efficiency at the price of system stability and service quality degradation. Frequent infrastructure reallocations are required to reduce the average server demand, leading to a high number of VM live-migrations. [3]

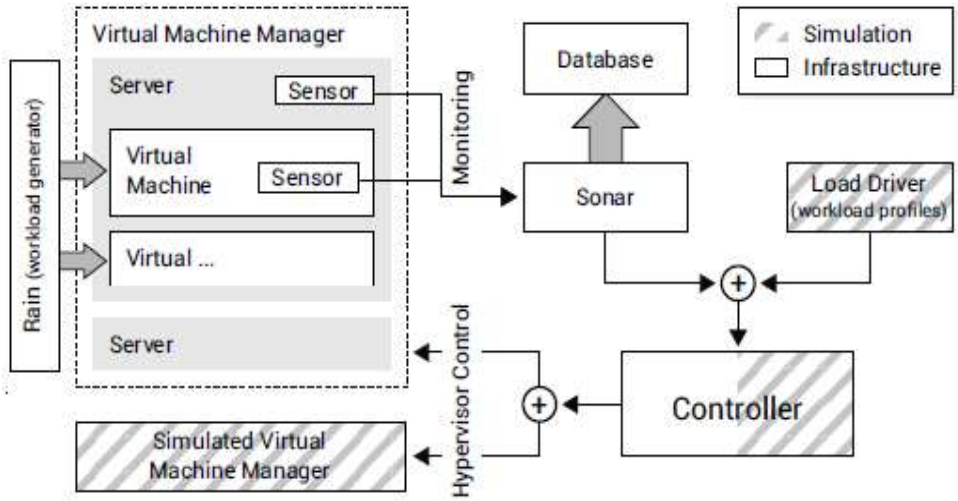

Fig -3: Simulation and experimental infrastructure setup

In fourth paper, Javier Espadas attempts to establish formal measurements for under and over provisioning of virtualized resources in cloud infrastructures, specifically for SaaS platform deployments and proposes a resource allocation model to deploy SaaS applications over cloud computing platforms by taking into account their multitenancy, thus creating a cost-effective scalable environment. a tenant-based model is presented to tackle over and underutilization when SaaS platforms are deployed over cloud computing infrastructures. This model contains three complementary approaches: (1) tenant-based isolation which encapsulates the execution of each tenant, (2) tenant-based load balancing which distributes requests according to the tenant information, and (3) a tenant-based VM instance allocation which determines the number of VM instances needed for certain workload, based on VM capacity and tenant context weight. [4]

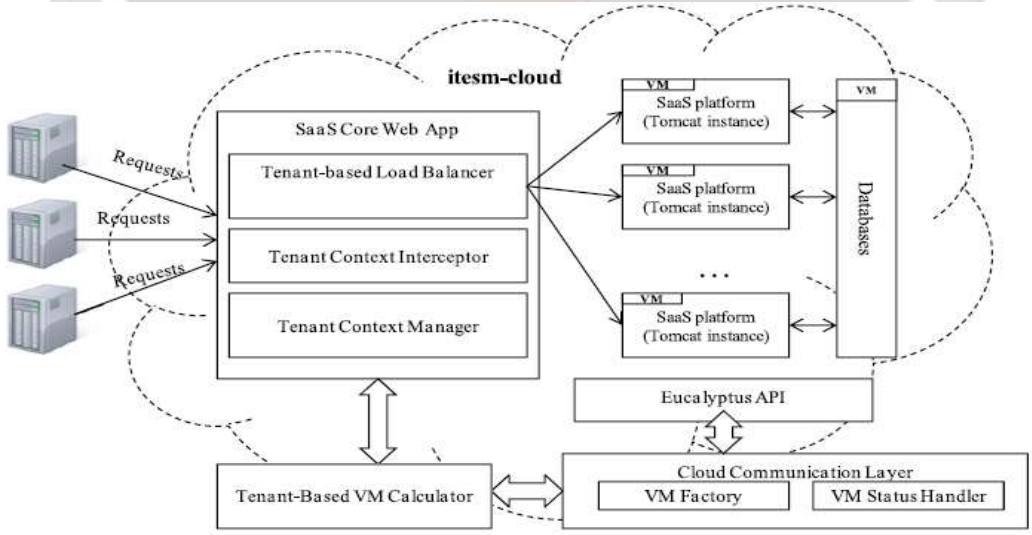

Fig -4: Tenant-based VM instance creation architecture

In fifth paper, Weiwei Lin propose a threshold-based dynamic resource allocation scheme for cloud computing that dynamically allocate the virtual resources (virtual machines) among the cloud computing applications based on their load changes (instead of allocating resources needed to meet peak demands) and can use the threshold method to optimize the decision of resource reallocation. The scheme for cloud computing that monitor and predict the resource needs of the cloud applications and adjust the virtual resources based on application's actual needs. Threshold-based dynamic resource allocation scheme consists of two procedures, Datacenter and Broker. The broker procedure runs on user's machine with the application. The datacenter procedure, which works as the manager of the cloud computing resources, runs on the datacenter's central computer. These two procedures interact with each other to dynamically manage the virtual resources for cloud applications. [5]

3. CONCLUSIONS

Cloud computing is widely used technology. It allows business customers to scale up and down their resource usage based on need. Dynamic resource allocation is widely used in VM migration and it is better then static allocation. Using this we can reduce server demand and use minimum energy consumption and it gives optimize performance. Using live migration server overload can be avoided. Dynamic resource allocation provide memory isolation , re-allocation and on demand cpu control for cpu intensive workload and VM failover.

4. REFERENCES

[1]. Mohammadreza Mesbahi , Amir Masoud Rahmani , Anthony Theodore Chronopoulos "An efficient dynamic resource allocation strategy for VM environment in cloud " International Conference on Pervasive Computing (ICPC) , 2015, DOI: 10.1109/PERVASIVE.2015.7087186.

[2]. Yuda Wang, Renyu Yang, Tianyu Wo, Wenbo Jiang, Chunming Hu "Improving Utilization through Dynamic VM Resource Allocation in Hybrid Cloud Environment" Parallel and Distributed Systems (ICPADS), 2014 20th IEEE International Conference on , 2014 , DOI: 10.1109/PADSW.2014.7097814.

[3]. Andreas Wolke and Lukas Ziegler "Evaluating dynamic resource allocation strategies in virtualized data centers" IEEE International Conference on Cloud Computing, 2014, DOI: 10.1109/PERVASIVE.2014.7088671.

[4]. Javier Espadas , Arturo Molina , Guillermo Jimenez , Martin Molina , Raul Ramirez , David Concha "A tenant-based resource allocation model for scaling Software-as-a-Service applications over cloud computing infrastructures" DOI: 10.1016/j.future.2011.10.013.

[5]. Weiwei Lin, James Z. Wang, Chen Liang, Deyu Qi "A Threshold-based Dynamic Resource Allocation Scheme for Cloud Computing" DOI: 10.1016/j.proeng.2011.11.2568.