

TASK SCHEDULING IN CLOUD COMPUTING

Snehal S. Deore¹, Gayatri V. Shimpi², Sadhana S. Patil³, Bhagyashri K. Samrut⁴

¹ Student, Information Technology, Sanjivani COE, Maharashtra, India

² Student, Information Technology, Sanjivani COE, Maharashtra, India

³ Student, Information Technology, Sanjivani COE, Maharashtra, India

⁴ Student, Information Technology, Sanjivani COE, Maharashtra, India

ABSTRACT

Now a days, cloud computing is a center of attraction of many researchers, scientist. To provide services online and with pay as you go makes it a strong. Still cloud computing has not reached to its best because of its some issues. Out of this one of the major is resource allocation. At a time so many customers are using the services which are provided by cloud computing provider. In this case the proper scheduling need to be maintain in between the available resources and coming requests. In this paper, we proposed an experimental performance analysis of the round robin and my broker, by this we try to see which one is achieve the optimum resource allocation.

Keyword : - cloud computing, round robin, IaaS, SaaS, Scheduling.

1. INTRODUCTION

Cloud Computing is the model of new trend computing where users have an access to utilize large computational resources as per their requirements. Cloud computing has three basic models i.e. IaaS, PaaS and SaaS through it provides services to customers. Every model has its different functionality as compare with other models. As like Infrastructure-as-a-Service clouds differ from other model, in that the user have an access to utilize compute, storage and network. In Platform as a Service model one more functionality get added and that is database. While in the Software-as-a-Service has all the layers in his model i.e. applications, database, compute, storage and network. The popularity of Cloud computing has a reason it offers small companies and organizations to have access to computing infrastructure without the need for prior investment. So, the risk of investment has been minimized by the Cloud computing. By this small institutions and individual teams can now have access to large computational resources with minor conditions. The Cloud computing has achieve this milestone because of the virtualization of the resources, at a time number of requests can be handled by the Cloud computing provider. Among the various issues in the Cloud computing one major issue is the resource allocation. Resource allocation is an integral and continuously evolving feature of many cloud computing and data center management problem. Consider the following related scenario. A cloud service provider currently allocates servers to tenant VMs based on CPU, memory and disk requirements of the VMs. At a later date, the service provider enhances the model and allocates network bandwidth resources as well to tenant VMs. Even later, the provider introduces a new fault-tolerant replication strategy, placing VM and data replicas intelligently across fault domains. In the allocation of resources and VM the resource allocation algorithm has perform a very important role. At this point, the VM allocation strategy depends on constraints that involve individual server capacity, network bandwidth capacity in the data center, as well as faultdomain definitions. Such varied and evolving resource allocation requirements are inherent not just to multi-tenant data centers. Such varied and evolving resource allocation requirements are inherent not just to multi-tenant data centers. Capacity planning for cloud services [1], VM placement in private data centers [2,3], network virtualization and virtual network embedding [4, 5, 6], multi-path routing [7,8], and data replica management [9], all utilize significant resource

allocation components. Broadly, they involve dividing and allocating resources subject to certain constraints such as guaranteed server performance, network performance, and fault tolerance requirements. The remainder of the paper is organized as follows. In Section 2, we present a literature review of the relevant studies according to the allocation strategy of cloud computing. Section 3 elaborates round robin and mybroker policy for cloud resource allocation. Section 4 presents the comparative experiments of these two algorithms involving the various parameters. Finally Section 5 draws conclusions and the future research directions.

2 LITERATURE SURVAY

The beginning of this section starts with an allocation example, where actually the resource allocation strategy makes the importance.

2.1 Execution Time

Different kinds of resource allocation mechanisms are proposed in cloud. In the work by Jiani et al. [10], actual task execution time and preemptible scheduling is considered for resource allocation. It overcomes the problem of resource contention and increases resource utilization by using different modes of renting computing capacities. But estimating the execution time for a job is a hard task for a user and errors are made very often [11]. But the VM model considered in [10] is heterogeneous and proposed for IaaS. Using the above-mentioned strategy, a resource allocation strategy for distributed environment is proposed by Jose et al. [12]. Proposed matchmaking (assign a resource to a job) strategy in [13] is based on AnySchedulability criteria for assigning jobs to opaque resources in heterogeneous environment in the cloud computing.

2.2 Virtual Machine (VM)

A system which can automatically scale its infrastructure resources is designed in [14]. The system composed of a virtual network of virtual machines capable of live migration across multi-domain physical infrastructure. By using dynamic availability of infrastructure resources and dynamic application demand, a virtual computation environment is able to automatically relocate itself across the infrastructure and scale its resources. But the above work considers only the non-preemptible scheduling policy. Several researchers have developed efficient resource allocations for real time tasks on multiprocessor system. But the studies, scheduled tasks on fixed number of processors. Hence it lacks in scalability feature of cloud computing [15]. Recent studies on allocating cloud VMs for real time tasks [16], [17], [18] focus on different aspects like infrastructures to enable real-time tasks on VMs and selection of VMs for power management in the data center. But the work by Karthik et al. [15], have allocated the resources based on the speed and cost of different VMs in IaaS. It differs from other related works, by allowing the user to select VMs and reduces cost for the user.

2.3 Gossip

Cloud environment differs in terms of clusters, servers, nodes, their locality reference and capacity. The problem of resource management for a large-scale cloud environment (ranging to above 100,000 servers) is addressed in [19] and general Gossip protocol is proposed for fair allocation of CPU resources to clients. A gossip based protocol for resource allocation in large scale cloud environments is proposed in [9]. It performs a key function within distributed middleware architecture for large clouds. In the thesis, the system is modeled as a dynamic set of nodes that represents the machines of cloud environment. Each node has a specific CPU capacity and memory capacity. The protocol implements a distributed scheme that allocates cloud resources to a set of applications that have time dependent memory demands and it dynamically maximizes a global cloud utility function. The simulation results show that the protocol produces optimal allocation when memory demand is smaller than the available memory in the cloud and the quality of the allocation does not change with the number of applications and the number of machines. But this work requires additional functionalities to make resource allocation scheme is robust to machine failure which spans several clusters and datacenters. But in the work by Paul et al. [19] cloud resources are being allocated by obtaining resources from remote nodes when there is a change in user demand and has addressed three different policies to avoid over-provisioning and under provisioning of resources. Recent research on sky computing focuses on bridging multiple cloud providers using the resources as a single entity which would allow elastic site for leveraging resources from multiple cloud providers [19]. Related work is proposed in [18] but it is considered only for preemptible tasks. Yang et al. [18] have proposed a profile based approach for scaling the applications automatically by capturing the experts' knowledge of scaling application servers as a profile. This approach greatly

improves the system performance and resource utilization. Utility based RAS is also proposed for PaaS in [12]. In paper [8], Gossip based co-operative VM management with VM allocation and cost management is introduced. By this method, the organizations can cooperate to share the available resources to reduce the cost. Here the cloud environments of public and private clouds are considered. They have formulated an optimization model to obtain the optimal virtual machine allocation. Network game approach is adopted for the cooperative formation of organizations so that none of the organizations wants to deviate. This system does not consider the dynamic co-operative formation of organizations. Related work is discussed in [19] that use desktop cloud for better usage of computing resources due to the increase in average system utilization. The implication for a desktop cloud is that individual resource reallocation decisions using desktop consolidation and decision based on aggregate behavior of the system

2.4 Round Robin

Round robin is the scheduling algorithm used by the CPU during execution of the process. Round robin is designed specifically for time sharing systems . It is similar to first come first serve scheduling algorithm but the pre-emption is the added functionality to switch between the processes. A small unit of time also known as time slice or quantum is set/defined. The ready queue works like circular queue. All processes in this algorithm are kept in the circular queue also known as ready queue. Each New process is added to the tail of the ready/circular queue. By using this algorithm ,CPU makes sure, time slices (any natural number) are assigned to each process in equal portions and in circular order, dealing with all process without all process without any priority. It is also known as cyclic executive. The main advantage of round robin algorithm over first come first serve algorithm's that it is starvation free. Every process will be executed by CPU for fixed interval of time (which is set as time slice) . So in this way no process left waiting for its turn to be executed by the CPU.

3. ROUND ROBIN AND MY BROKER POLICY

3.1 Round Robin Policy

Let, we take a pseudo code review for round robin and my broker policy.

```

Round_Robin_Load_Balancing ()
{
Initialize all the VM allocation status to AVAILABLE in the VM state list;
Initialize hash map with no entries;
While(new request are received by the Data Centre Controller)
Do
{
Data Center Controller queue the requests;
    Data Centre Controller removes a request from the beginning of the queue;
If(hash map contain any entry of a VM corresponding to the current requesting user base && VM allocation status
== AVAILABLE)
{
The VM is reallocated to the user base request;
}
}

```

Else

```
{
Allocate a VM to the user base request using Round Robin Algorithm;
Update the entry of the user base and the VM in the hash map and the VM state list;
}
}
}
```

3.2 MyBroker Policy

In this policy the resources has been allocate as per the following pseudo code.

```
//package cloudreports.extensions.brokers;
import cloudreports.extensions.brokers.Broker;
import java.util.List;
public class MyBroker extends Broker {
public MyBroker(String name) throws Exception {
super(name);
}
@Override
public List<Integer>getDatacenterIdList() {
return super.getDatacenterIdsList();
}
@Override
public int getDatacenterId()
{
return getId();
}
}
```

4. PERFORMANCE ANALYSIS

In this section we can to see the comparative analysis for both the policies parameter wise.



Fig-1: Resource utilization in My Broker Policy



Fig-2: Resource utilization in Round Robin Policy



Fig-3: Output of MyBroker Policy



Fig-4: Output of Round Robin Policy

5. CONCLUSIONS

As Cloud computing has many benefits over the time, but still it facing the various issues. In this worked we address the resource allocation by round robin and mybroker policy. An experimental results shows that round robin is a comparatively optimum method to allocate a resources in cloud computing.

6. REFERENCES

- [1] Microsoft Assessment and Planning Toolkit. <http://www.microsoft.com/map/>.
- [2] J. Mudigonda, P. Yalagandula, M. Al-Fares, and J. Mogul. SPAIN: COTS Data-Center Ethernet for Multipathing over Arbitrary Topologies. In Proceedings of the 7th Symposium on Networked Systems Design and Implementation (NSDI), San Jose, CA, Apr. 1810.
- [3] G. Lee, N. Tolia, P. Ranganathan, and R. H. Katz, "Topology-Aware Resource Allocation for Data Intensive Workloads", In Proceedings of the AsiaPacific Workshop on Systems (APSys), 1810.
- [4] H. Ballani, P. Costa, T.Karagiannis, and A. Rowstron, Towards Predictable Data Center Networks", In Proceedings of the Annual Conference of the ACM Special Interest Group on Data Communication (SIGCOMM), Toronto, Canada, Aug. 1811.
- [5] H. Ballani, P. Costa, T.Karagiannis, and A. Rowstron, Towards Predictable Data Center Networks", In Proceedings of the Annual Conference of the ACM Special Interest Group on Data Communication (SIGCOMM), Toronto, Canada, Aug. 1811.
- [6] H. Ballani, P. Costa, T.Karagiannis, and A. Rowstron, Towards Predictable Data Center Networks", In Proceedings of the Annual Conference of the ACM Special Interest Group on Data Communication (SIGCOMM), Toronto, Canada, Aug. 1811.
- [7] C. Guo, G. Lu, H. J. Wang, S. Yang, C. Kong, P. Sun, W. Wu, and Y. Zhang, "SecondNet: A Data Center Network Virtualization Architecture with Bandwidth Guarantee". In Proceedings of the International Conference on emerging Networking EXperiments and Technologies (CoNEXT), 1810.
- [8] J. Mudigonda, P. Yalagandula, M. Al-Fares, and J. Mogul. SPAIN: COTS Data-Center Ethernet for Multipathing over Arbitrary Topologies. In Proceedings of the 7th Symposium on Networked Systems Design and Implementation (NSDI), San Jose, CA, Apr. 1810.
- [9] M. Al-Fares, A. Loukissas, and A. Vahdat. A Scalable, Commodity Data Center Network Architecture. In Proceedings of the Annual Conference of the ACM Special Interest Group on Data Communication (SIGCOMM), Seattle, WA, Aug. 1808. [9]Hadoop Rack-Aware Replica Placement. <http://bit.ly/Nc0V8F>.
- [10] Jiyani et al.: Adaptive resource allocation for preemptable jobs in cloud systems (IEEE, 1810), pp.31-36.
- [11] HadiGoudarzi and MassoudPedram: Maximizing Profit in Cloud Computing System Via Resource Allocation: IEEE 31 st International Conference on Distributed Computing Systems Workshops 1811: pp,1-6.
- [12] K. Jang, S. Han, S. Han, S. Moon, and K. Park, SSLShader: Cheap SSL Acceleration with Commodity Processors. In Proceedings of the 8th Symposium on Networked Systems Design and Implementation (NSDI), Boston, MA, Mar. 1811.
- [13] Karthik Kumar et al., "Resource Allocation for real time tasks using cloud computing", ICCCN, 1811 IEEE, pp. 1-7, July 1811.

- [14] Wei-Tek Tsai Qihong Shao Xin Sun Elston, J. Service-oriented cloud computing. In world congress on services, pages 473-478, 1810.
- [15] K.H Kim et al. Power-aware provisioning of cloud resources for real time services. In international workshop on Middlelware for grids and clouds and e-science, pages 1-6, 1809.
- [16] RerngvitYanggratoke, FetahiWuhib and Rolf Stadler: Gossip-based resource allocation for green computing in Large Clouds: 7th International conference on network and service management, Paris, France, 24-28 October, 1811
- [17] P.Ruth,J.Rhee, D.Xu, R.Kennell and S.Goasguen, "Autonomic Adaptation of virtual computational environments in a multi-domain infrastructure", IEEE International conference on Autonomic Computing, 1806,pp.5-14.
- [18] Yang wt.al A profile based approach to Just in time scalability for cloud applications, IEEE international conference on cloud computing ,1809,pp 9-16.
- [19] AndrzejKochut et al. : Desktop Workload Study with Implications for Desktop Cloud Resource Optimization,978-1-4244-6534-7/10 1810 IEEE.

