IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

Bandwidth Aware Vm Migration Policy in Cloud Data Centers

Hetal Bhadja[1], Jigneshkumar Lakhani[2]

Department of Computer Engineering, Student of PG studies-MEF Group of Institutions, Rajkot, India[1]

Department of Computer Engineering, Faculty of PG studies-MEF Group of Institutions, Rajkot, India[2]

Abstract: cloud computing is a new paradigm for enterprise that can effective facilitate the execution of tasks.task scheduling is an important issue which is greatly degrade performance of cloud computing environment.cloud service provide and clients have different requirements and objectives in cloud environment resource availability and load on resources is day by day changed time to time .so scheduling resources in cloud is complicated problem.BATS algorithm is proposed for multiple data centers to overcome the time overhead and cost when replacement between two datacenters. after that Compare result with existing algorithms using cloud simulator. Observe the parameter performance.

Keywords: task scheduling; data center;virtual machine;

I. INTRODUCTION

Now a day, IT technologies are growing day by day and that's why the need of the computing are needed and the storage firstly. High efficient performance of cloud computing has become widely accepted approach. All the types of IT facilities are provided by cloud computing to the user as a "service". according to standard institute. Cloud computing provides Infrastructure as a Service (IaaS) which provides infrastructure like, storage and resources to user for various purpose, Platform as a service(PasS) provides platform to their clients for different application development and Software as a service(SaaS) which provides software directly from the cloud systems. So that users don't need to initially install the software. Amazon, EC2 are example of IaaS and PaaS where Salesforce is example of Saas. Cloud computing provide benefits by reducing the difficulty of setting up hardware, software and platform requirements. It also provides other benefits like efficiency, availability, cost-saving and scalable. Attractive feature of cloud computing is its elastic nature; service providers of cloud computing based company can be elastic in their service delivery to fulfill the end user requirements. Provide on demand services in pay-as-you-use manner so, user have to pay only for whatever they use.

Task scheduling algorithm is a one efficient method for matvhing the tasks or allocated to data center resources. Due to conflicting scheduling parameters are not absolutely exists. So scheduler can be used different application and try to make new scheduling algorithm according to parameter. A problem can be solved or taken time as a days hours or years also. The efficiency of an algorithm is evaluated by the amount of time necessary to execute it. The execution time of an algorithm is stated as a time complexity function relating the input. Scheduling problem can be further classified into two types as optimization problem and decision problem based on objective An optimization problem requires finding the best solution among all the feasible solutions in set. The main objective of decision problem is relatively easy. For a specified feasible solution problem needs optimal answer to whatever object is gain Clearly, optimization problem is harder than decision problem.

In recently cloud services is very demandful and increase day by day to fulfill the requirements of users. In cloud computing, data centers taken more time when data replacement between two virtual machines in multiple data centers Nowadays, data center can be having some limitation like limited bandwidth resources on the links and the computing capacity of the server, which can be change resource management problems. One classic issue is the task allocation problem, which involves variousconstraints, including performance, bandwidth, network, time and cost. Generally speaking, a task usually consists of two parts the computation workloads and the communication traffic. Different tasks running in the data center will compete for computing capability on the servers and bandwidth resources on the links. Obviously, to minimize the duration of communication traffic, locality is one important factor that we need to take into consideration. That is, we need to place the tasks as close to each other as possible, in order to lower the hops of communication between tasks. Furthermore, the tasks running in the same server will not need communication traffic between each other due to the internal communication within the server.

- However, the bandwidth capacity of each link is limited.
- More tasks running under the same link will lead to the lower
- Average bandwidth allocated to each task, which will decrease.
- The communication speed and increase the communication

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

- Duration. What is worse, the computing capability of a server
- Is limited. If tasks are packed together, it will also reduce the Computing capability allocated to each task. With the degraded Computation speed, the computation workloads will take more Time to complete. Geographically distributed data centers between countries that bandwidth can be manage difficult. Time duration peak power and cost factors are also considered when transmission between multiple data centers.
- On the other hand, if we apply load-balancing and allocate

The tasks evenly to the servers, the computing speed of each task and cost can be maximized [2]. However, the geographically separated tasks need more links to communicate with each other. With the limited bandwidth resources on the links, load balancing will considerably lengthen the communication time. Therefore, there is a tradeoff between the locality and task scheduling load balancing. Complexity class of the task scheduling problem categorizes as NP-complete relating large space with corresponding to large no of potential solutions and it takes longer time to get the optimal solution.

CHALLENGES AND ISSUES

It's generates number of the following challenge and issues, which are summarized as follow.

A. Increasing makespan:
An increasing makespan is required when a data is transmitting from one data center to another data center in local region length and time schedule is difficult to manage with native instructions.

B. Quality of service by cloud provider:
Qos is the collective effort of service performance, which determines the degree of satisfaction of user for the service. Qos is expressed by the following measures such as completion time, latency, execution price, throughput, reliability.

C. Data transfer cost between data centers :
The cost foran under-serviced request is set to be equal to the service price for one hour divided by the maximum number of requests that can violate the response time in each charg cycle. The migration cost is considered to be linearly related to the migration latency.

D. Consume more energy:
Energy on brown sources of electricity and supplement/diversify the sources of the data center that can be generate the power like that renewable sources such as wind and solar at the datacenters. They try to run the virtual machine on the basis of minimum number of the physical machine so power server less usage can be lead to reduce the power consumption

E. Time overhead:
Time overhead meand consume more time for resource utilization and virtual machine migration between multiple data centers. Solving a multiple objectives scheduling problem is difficult than the single objective scheduling problem.

F. Efficiency:
It's important to detect under loaded VM because they will be operating with a bad resource utilization efficiency. In this case all the VMs from this one data center should be migrated to another data center and than that data center should be switched off or set to a low-power mode so in the terms of resource utilization not utilized efficient.

PARAMETERS OBSERVED

It's generates number of the following challenge and issues, which are summarized as follow.

A. Compilation time:
The model has to be designed to allocate the resources for batch of tasks so that the completion time is minimal.

B. Power:
The model has to be designed to manage the cloud environment for maximal resource utilization which leads to minimal power consumption as well.

C. Size:
The model must work for any size of batch requests. i.e scalable. The data centers are over provisioned to meet the unexpected workload variations

D. Makespam:
Makes pan is the total length of the schedule that is, when all the jobs have finished processing. the problem is presented as an online problem as dynamic scheduling, online algorithm needs to make the job scheduling according to which job is first execute and their scheduling dependency also scheduled.

E. Resource utilization:
The model must enable to provide uninterrupted services to customers. Services are available round the clock. If any resource is shut due to any reasons the work load on that resource must be migrated to another nearby, lightly loaded resource to meet the service level agreements. During the migration of task migration cost and the down time should be minimized.

F. Task dependency:
The model must consider communication dependency among tasks for a realistic modeling of the execution environment.

G. Reliability:
When meta-heuristic algorithms are applied on the model the outcome may not be guaranteed. However, one has to design the system for maximal reliability.

H. Bandwidth:
Bandwidth is what determines how much stuff can pass through your network at once. It affects how long time taken for download and upload the information

E. network capacity:
Capacity management under approach tends to center on predicting peak usage needs like hardware software resources and ensuring that in-house data centers can meet those demands according to capacity.

II. EXISTING APPROACHES

Weiwei Lin1, Chen Liang1, James Z. Wang2 and Rajkumar Buyya [1] have proposed a novel approach for obtain better performance. to obtain better Performance and to minimize time, they propose a bandwidth aware divisible task scheduling algorithm. A bounded multiport task scheduling algorithm constraints can be proposed in the paper. By solving this model, the optimized allocation scheme that schedule how no of tasks are assign to each virtual machine. On the basis of the optimized allocation scheme is generated and bandwidth-aware task-scheduling (BATS) algorithm, is proposed. They have used bandwidth and capacity parameter in their approach.

Hadi Goudarzi and Massoud Pedram [2] have proposed a force directed scheduling algorithm to minimize the total force. Considering a distributed cloud system can have considered the no of datacenters which can be geographically located. An offline solution based on force-directed scheme is proposed which can be decide the application placement for long periods of time. The online solution is for virtual machine migration and placement for geographically distributed datacenters based on predictions about the application life duration, workload capability, dynamic energy cost, and renewable energy generation capacities at different located data center in the cloud. They have been used type of server and power supply of VM and cost of migration.

Xin Sun_, Sen Su_, Peng Xu_, Shuang Chi_ and Yan Luo[3] have proposed resource scheduling problem to a bounded multi-dimensional knapsack problem (B-MDKP), taking into account the requirement dependency among multidimensional resources. that assign the resources based on the energy optimize methods. They have used CPU, memory, storage, and network Bandwidth parameters

B.Parkavi I, G.Malathl and Ching-Hsien Hsu[4] have proposed a Abstraction based K tier Partitioning algorithm is proposed for k tier data center with multiple computing capacities by solved large scale application using abstraction refinement. Abstraction scheduler keeps track of free intervals on all computing nodes in data center and uses information to schedule blocks of tasks in job. Independent tasks in a parallel job are scheduled simultaneously which leads to idea of cost of maintaining set of free intervals on nodes in data center.

XiaonianWu, MengqingDeng, RunlianZhang, BingZeng, Shengyuan Zhou [5] have proposed a task scheduling algorithm based on Qos driven for cloud computing first of all each and every task has priority assign and sort the task by priority after that algorithm evaluate the completion time of each task on different services.and schedules each task on services and completed task according to sorted queue.

Mohsen Amini Salehi 1,*,†, Bahman Javadi 2 and Rajkumar Buyya1[6] have proposed a method resource provisioning VM based lease preempting used we model the Then, to reduce the impact of these side effects, we propose and compare several policies that decide the set of the leases in which preempting VM-based leases from external users and allocating them to the local ones.

Anton Beloglazov, Jemal Abawajyb, Rajkumar Buyya1[7] have proposed energy-aware allocation provide data center resources to client applications in a way that improves energy efficiency for the data center, while sending the quality of service decline. In particular, proposed architecture and principles of energy efficient management; energy-efficient resource allocation policies and scheduling algorithms considering QoS expectations and power usage characteristics of the devices; and a number of open research challenges, addressing which can bring substantial benefits to both resource providers and consumers.

LI Hongyou1, WANG Jiangyong1, 2, PENG Jian1*, WANG Junfeng1, LIU Tang1 have proposed propose two algorithms energy aware scheduling algorithm using workload consolidation technique and the Energy aware Live Migration algorithm using Workload- aware Consolidation Technique . the two algorithms are based on the fact that multiple multidimensional resources such that cpu memory and bandwidth shared by users concurrently in cloud data centers and heterogeneous workloads having different characteristic. Both algorithms investigate the problem of consolidating. They use virtual machine using minimum number of physical machine so, less use of physical server that can be minimize the power consumption.

III. PROPOSED FRAMEWORK

Main aim of this research is BATS algorithm is implemented for multiple data center.inn share physical resources like cpu, bandwidth, memory, time delay and cost on a single physical host. multiple virtual machines can share the bandwidth of data center by virtualization. many system parameters such as power, memory space, network bandwidth and efficiency of the task scheduling. frequent data exchange between no of hosts, nodes and data centers that make feasible. Task is performed multi objective. Each and every data center has a scheduling node and that node decides how many tasks should be specifically allocated and transferred to each VM.in that optimize allocation scheme is generated by the system. Algorithm will not violate the real situation when task execution time is less than task transmission time.in such case they will certainly exist until the state that the previous task is already finished.in that concept of broker

is used. Which works instead of broker; all the details are related to the VM creation, the VM destruction, and task submission to these VMs. In Cloud Sim, broker is java object and work as mediator in between internet and data center It accepts a list of VMs and a list of cloudlets to be completed to user , and then binding the tasks to virtual machines properly in a certain way and submitting all cloudlets. When scheduling done it give the result to the user and cloudlets works are finished. The algorithm is deployed on Broker and works out the optimized task allocation scheme on the basis of information of VM(computing power and bandwidth) and submits the optimized allocation scheme so that cloud broker allocate the tasks to virtual machine. After applying BATS algorithm for multiple data centers if there is a situation occurred overloaded load into the virtual machine then that virtual mach in tasks are allocating to another virtual machine of data center.

In that issues regarding above papers are resolved makespan, quality of services cost bandwidth for the multiple data centers can be handled. In this proposed approach we first make an BATS efficient algorithm for multiple data centers .no of tasks are assign to no of virtual machines to decided by scheduling node in multiple data center.in that multiple VMs can share the bandwidth of data center by using network virtualization. We compared BATS with three other algorithms We conducted experiments with four scheduling algorithms under the same condition and compare the performance of the algorithms

ACKNOWLEDGMENT

I hereby thank the individuals who directly or indirectly helped us in completion of our work from good to better and at last to the best. who has constantly inspired us for this work. Last but not the least, I would like to thank our parents and faculty members who have done a wonderful job in presenting this wonderful work which gives the information about the existing and upcoming devices and technologies that will soon become the part of our routine.

REFERENCES

[1]. Lin, Weiwei, et al. "Bandwidth-aware divisible task scheduling for cloud computing." Software: Practice and Experience 44.2 (2014): 163-174.
[2]. Goudarzi, Hossein, and Massoud Pedram. "Geographical load balancing for online service applications in distributed datacenters." Cloud Computing (CLOUD), 2013 IEEE Sixth International Conference on. IEEE, 2013.
[3]. Sun, Xin, et al. "Multi-dimensional resource integrated scheduling in a shared data center." Distributed Computing Systems Workshops (ICDCSW), 2011 31st International Conference on. IEEE, 2011.
[4]. Parkavi, B., G. Malathy, and Rm Somasundaram. "Abstraction based K-tier partitioning scheme for parallel job scheduling in multiple cloud centers." Electronics and Communication Systems (ICECS), 2014 International Conference on. IEEE, 2014.
[5]. X. Wu, M. Deng, R. Zhang, B. Zeng, and S. Zhou, "A Task Scheduling Algorithm based on QoS-Driven in Cloud Computing," Inf. Technol. Quant. Manag., vol. 17, pp. 1162–1169, 2013.
[6]. Salehi, Mohsen Amini, Bahman Javadi, and Rajkumar Buyya. "Resource provisioning based on preempting virtual machines in distributed systems." Concurrency and Computation: Practice and Experience 26.2 (2014): 412-433.
[7]. Beloglazov, Anton, Jemal Abawajy, and Rajkumar Buyya. "Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing." Future generation computer systems 28.5 (2012): 755-768.
[8]. Hongyou, Li, et al. "Energy-aware scheduling scheme using workload-aware consolidation technique in cloud data centres." Communications, China 10.12 (2013): 114-124.