AN ENERGY EFFICIENT SCHEME FOR CLOUD RESOURCE PROVISIONING USING TASK SCHEDULING STRATEGY BASED ON VACATION QUEUING THEORY

Mr. K Sathish
Assistant Professor of M.E
Department of CSE
K.Ramakrishnan College of Technology
Samayapuram
Tiruchirappalli,India
Email:- successathish89@gmail.com

M.Sarojinidevi
Student of M.E
Department of CSE
K.Ramakrishnan College of Technology
Samayapuram
Tiruchirappalli , India
Email:- sarodevi18@gmail.com

Abstract: **In Cloud Computing, Energy conservation is the major problem and it provides benefits such as reducing costs, increased reliability of the system and it also provides protection to the environment. Energy–aware scheduling is used to achieve these benefits. Existing energy-aware scheduling algorithms are not real time task oriented and also it lacks in system schedulability. Vacation queuing model is used for real-time, a periodic, independent task to solve this problem. The system which is proposed here can achieve energy optimization by combining the cloud resources with current exploitation. The eminence of computing hardware and nodes are well-organized with virtual network topologies. Vacation system is implemented with sojourn time to guarantee the schedulability of real-time tasks, efficiently. Energy conservation is achieved by switching the active host to sleep mode when the system does not perform any action. The task should be completed within the deadline and each user must provide the deadline to avoid rejection. The deadline is analyzed and acknowledgement is provided to the scheduler for each task completion. To implement automatic energy-saving management and performance optimizations in modern cloud data centers.**

Keywords: **Task scheduling, heterogeneous computing environment, schedule length, energy consumption.**

I. INTRODUCTION

Cloud computing is basically a service oriented architecture rendering easy access to all who make use of it[2]. An organization moving from traditional model to cloud model which has been switched from dedicated hardware and depreciate the service for certain period of time to use shared resources in cloud infrastructure and pay based on the usage. The cloud computing reduces the ground work of infrastructure costs, run the application faster with upgraded manageability and less maintenance, to rapidly adjust resources to meet inconsistent and capricious business demand and allow the user to differentiate their business instead of on infrastructure. The cloud providers use “pay-as-you-go” model. The varying characteristics of cloud making it different from other computing technologies are on-demand self-service, agility, autonomic computing, virtualization, parallel and distributed architecture. It is consisting of a family of interconnected and virtualized computers.

First the system is defined as Physical Machine (PM) as cloud servers, and instance or Virtual Machine (VM) as the virtual server provided to the users. And the users system the resources can be accessed from anywhere in any location.

So it is easy for the punter to preserve the records in the cloud. Cloud consists of a number of resources in any form i.e. cups, firewall, network are always dynamically allocated according to the sequence and requirements of the task, subtasks.

The scheduling of tasks in cloud means choose the best suitable resource available for execution of tasks or to allocate computer machines to tasks in such a manner that the completion time is minimized as possible[5].

In scheduling algorithms list of tasks is created by giving priority to each and every tasks where setting of priority to different tasks can be based on various parameters[4].

Tasks are then chooses according to their priorities and assigned to available processors and computer machines which satisfy a predefined objective function.

II. AN ENERGY EFFICIENT RESOURCE PROVISIONING USING TASK SCHEDULING

Task scheduling is the process in which arrived task [6] is uncertain at the run time and resource allocation is difficult when simultaneously many tasks arrive. The task components/task arrived is not known until it reaches the dispatching center node. Thus the execution time of the arrived tasks may not be identified. Therefore, the allocation[6] of tasks is carried-out on the fly when the application executes the project and focuses on the periodic scheduling and independency of real-time tasks. This approach dynamically creates virtual clusters dealing with the conflicts between the parallel and serial tasks in the system. The tasks are available dynamically for scheduling over time within the scheduler. The dynamic scheduling [6] is more flexible to be able to determine the run time in advance. Consider a case in which four states are viewed, running state, idle state, sleep state and recover state.

The scheduling[8] is planned with the measure in order to identify the degree of efficient task computation relative to the completion time of the application. The degree of task computation is used as a utilization value to identify a level of virtue for executing task on processor. This in turn implies the energy consumption of that processor efficiently. In this proposed approach, the task load is attuned dynamically without the prediction of running time of all the tasks. TSAST scheduling strategy is implemented to provide best tradeoff in scheduling the tasks. Job scheduling is global centralized - As cloud computing is a computing model which supply the centralized resource by the mirror service to multiple distributed applications, and this mirroring deployment can make heterogeneous procedures' executing of interoperate become easier, which used to be difficult to deal with. Therefore, virtualized technology and mirroring services make the task scheduling of cloud computing achieve a global centralized scheduling. Each node in the cloud is independent - In cloud computing, the internal scheduling of every cloud node is autonomous, and the schedulers in the cloud will not interfere with the scheduling policy of these nodes.

The scalability of job scheduling - The scale of resources supply from cloud provider may be limited in early stages. With the addition of a variety of computing resources, the size of the abstract virtual resources[6] may become large, and the application demand continues increasing. In the cloud, task scheduling must meet the scalability features, so that the throughput of the task scheduling in the cloud may not be too low. Job scheduling can be dynamically self-adaptive - Expanding and shrinking applications in the cloud may be necessary depend on the requirement. The virtual computing resources in cloud system may also expand or shrink at the same time. The resources are constantly changing, some resources may fails, and new resources may join in the clouds or restart.

The set of job scheduling[3] - Task scheduling is divided into two parts: one is used as a unified resource pool scheduling, and primarily responsible for the scheduling of applications and cloud API; the other is for the unified port resource scheduling in the cloud, for example, Map Reduce task scheduling. However, each scheduling consists of two two-way processes that are scheduler leases resource from cloud and scheduler call backs the requested resources after use. The former process is scheduling strategy and the latter one is call back strategy. The combination of the scheduling and call back resource strategy is the set of task scheduling.

III. AN ENERGY EFFICIENT RESOURCE PROVISIONING USING TASK SCHEDULING (WORKING)

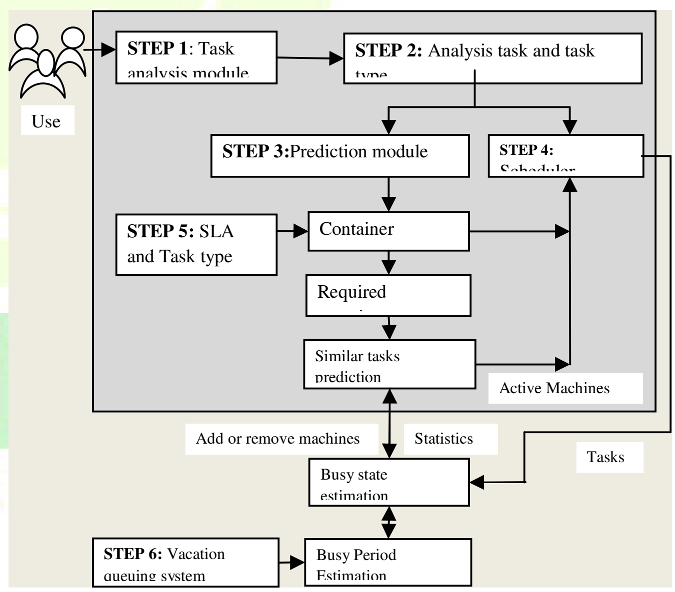

Figure 1.1 System Architecture

Users, server and scheduler as actors in the system. User module send the request the server module, it is in the form of service level agreement. This agreement contains type of request, time, and cost and so on. Fig 1.1 refers task analysis module in server finds each SLA of users. Then segregate each request and send to prediction module and scheduler. Prediction module communicates with container module to find the total number of available resources in cloud system. These details also sent to scheduler. The schedule analyzes active machine status. Then requests are considered as jobs. These jobs allocated in Vacation queue. In vacation queue contains four states running state, idle state, sleep state, recovering state.

Then predict busy time periods of each jobs. These slots are allocated to the machine based on TSAST algorithm. This algorithm clusters the similar task. Based on this framework can reduce the energy to system. The scheduler analyses the deadline and also provides acknowledgement for the completed tasks.

Following this, will describe stages of scheduling architecture .

An energy savings can be achieved in data centres through optimization mechanisms whose main objective is to minimize the energy consumption.

Database is a collection of information that is organized so that it can easily be accessed, managed, and updated. In one view, databases can be classified according to types of content: bibliographic, full-text, numeric, and images.

Datacenter class models the core infrastructure level services (hardware, software) offered by resource providers in a Cloud computing environment. It encapsulates a set of compute hosts that can be either homogeneous or heterogeneous as regards to their resource configurations (memory, cores, capacity, and storage). Furthermore, every Datacenter component instantiates a generalized resource provisioning component that implements a set of policies for allocating bandwidth, memory, and storage devices.

Datacenter Broker is a broker, which is responsible for mediating between users and service 6 providers depending on users' QoS requirements and deploys service tasks across Clouds. The broker acting on behalf of users identifies suitable Cloud service providers through the Cloud Information Service (CIS) and negotiates with them for an allocation of resources that meets QoS needs of users. The researchers and system developers must extend this class for conducting experiments with their custom developed application placement policies.

A CIS is an entity that provides resource registration, indexing, and discovering capabilities.

Virtual Machine is an instance of a VM, whose management during its life cycle is the responsibility of the Host component. As discussed earlier, a host can simultaneously instantiate multiple VMs and allocate cores based on predefined processor sharing policies (space-shared, time-shared).

Every VM component has access to a component that stores the characteristics related to a VM, such as memory, processor, storage, and the VM's internal scheduling policy, which is extended from the abstract component called VM Scheduling.

- DB Datacenter Broker
- CIS Cloud Information services
- V1, V2, V3 are the virtual machines

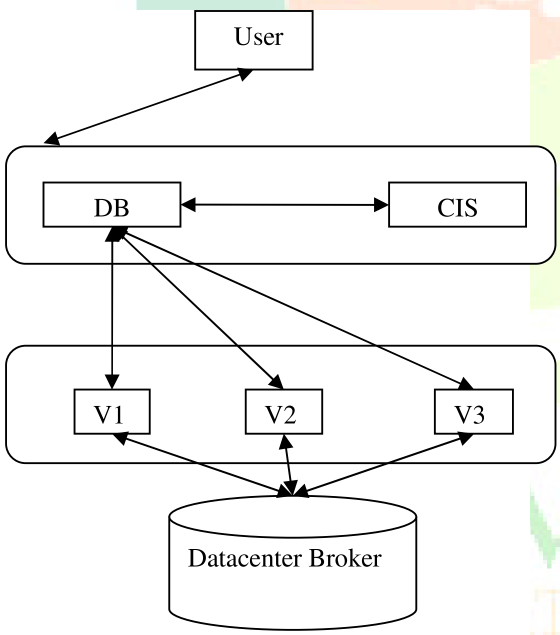

Figure 1.1 Stages of Scheduling

IV COMPONENTS

The various components used here are:

(A) Cloud Framework Construction

(B) workload Heterogeneity

(C)vacation M/G/1 Queue System

(D) Energy Consumption Expectation

(E)Evaluation Criteria For Task Scheduling

(A) Cloud Framework Construction

The cloud service provider is responsible for maintaining an agreed-on level of service and provisions resources accordingly. A CSP, who has significant resources and expertise in building and managing distributed cloud storage servers, owns and operates live Cloud Computing systems, it is the central entity of cloud. Cloud provider activities for utilizing and allocating scarce resources within the limit of cloud environment so as to meet the needs of the cloud application. It requires the type and amount of resources needed by each application in order to complete a user job.

The order and time of allocation of resources are also an input for an optimal resource allocation. Cloud user represents a person or organization that maintains a business relationship with, and uses the service from, a cloud provider. Users, who stores data in the cloud and rely on the cloud for data computation, Cloud consists of both individual consumers and organizations. Cloud consumers use Service-Level Agreements (SLAs) for specifying the technical performance requirements to be fulfilled by a cloud provider. In this module, scheduler gets the tasks from users and analyzes task types.

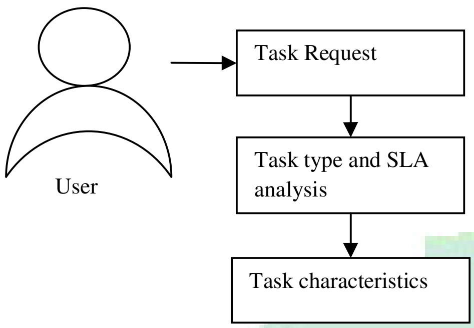

Figure 1.2 Cloud Frame Work Constructions

(B) workload Heterogeneity

The scheduler is responsible for assigning incoming tasks to active machines in the cluster. It also reports the average number of tasks in the queue during each control period to help the controller make informed decisions. The prediction module receives statistics about the usage of all resources (CPU and memory) in the cluster and predicts the future usage for all of them. Implement technique include the similar kind of workloads which is having fixed number of parameters such as length, number of CPU's required and buffer size of input and output files, etc.

Workload consolidation in data centers is achieved with the help of virtualization, which helps in saving the energy as unused nodes can be put to sleep mode or by shut down of the machines. The proposed approach uses the varying workloads and simulation results show the increased number of completed jobs. This approach does not give any information about the energy consumption of the data center and SLA.

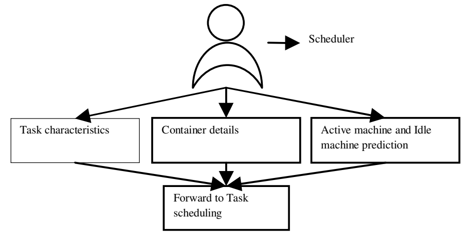

Figure 1.3 Workload Heterogeneity

(C) vacation M/G/1 Queue System

In this module can allocate task based on vacation M/G/1 queue system and contains various states can implement finite capacity M/G/1/K queue with the variation that the server goes for vacations when it is idle. This service model is referred to as one providing exhaustive service, as the server cannot go for a vacation until all the jobs presently in the system have been served.

This is the service model being considered here as this leads to a simple analytical model. Note that it is also possible to have a gated service model where the server only serves those customers that it finds in the system when it first starts service following its vacation. It then leaves for vacation again.

The four states are described as follows:

Running state: The state when a compute node is working.

Idle state: If there are no tasks arriving at a compute node, the node goes through an idle period to avoid frequent switches from the deep sleep state. The threshold of idle period isT_i^{idl}.

Sleep state: After the idle period ofT_i^{idl}, if there are no incoming tasks, the compute node goes into sleep state.

Recovering state: When a task arrives at the compute node under sleep state, the compute node needs to recover and then start to execute the task.

The recovering state is a transition state—in this state, a compute node is woken up from the sleep state and transitioned to the running state.

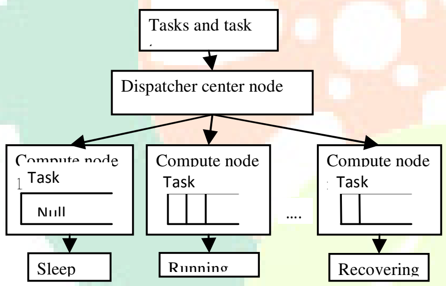

Figure 1.4 Vacation M/G/1 Queue Systems

(D) Energy Consumption Expectation

In this module compute sojourn time that consists of waiting time in the local queue of a compute node. And service time of the node performing tasks. The sojourn time probability distribution function is first derived for a Markovian queue; with both batch arrivals and batch departures that admits a product-form in networks of such queues.

The reversed process is then considered and it is confirmed that the unconditional reversed sojourn time has the same distribution as the forwards sojourn time. Using conditional forward and reversed sojourn times, a result is obtained for the sojourn time distribution on overtake-free paths in product-form networks of such batched queues. Then calculate the busy cycle and busy period, calculate mean power. Finally set the reasonable threshold idle time.

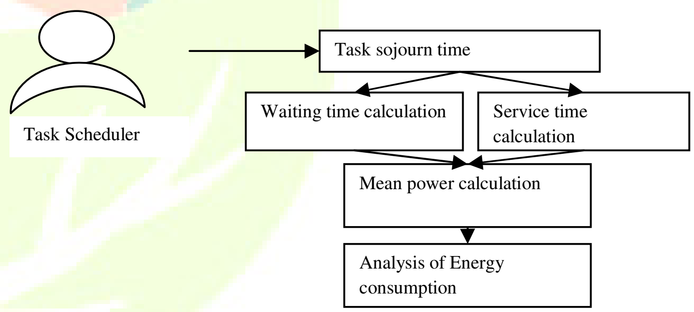

Figure 1.5 Energy Consumption Expectations

(E) Evaluation Criteria For Task Scheduling

In this module can implement Task Scheduling Algorithm based on Similar Tasks (TSAST). TSAST follows the partitioned or non-hierarchical clustering approach. It involves partitioning the given data set into specific number groups called Clusters.

Each cluster is associated with a centre point called centroid. Each point is assigned to a cluster with the closest centroid. Proposed dynamic VM allocation algorithm using clustering.

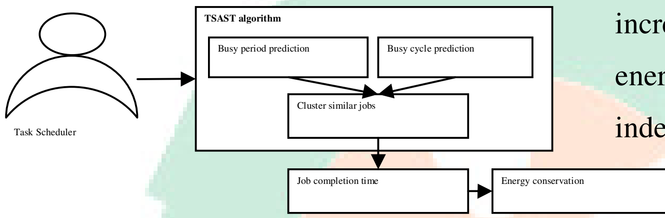

Figure 3.6 Evaluation Criteria for Task Scheduling

IV CONCLUSION

The efficient task scheduling algorithm is implemented in cloud computing for scheduling the task. There are two efficient approaches Vacation queue and TSAST algorithm. Using these algorithms, the analysis of any incoming task supplied by the user is done and then the scheduler allocates the task to the real-time controller. Before placing the task in a real time controller, the scheduler first checks whether the task can be completed by the real time controller or not. i.e., the real time machine is checked whether it has enough resources to complete the task or not.

If there is insufficient resources in real time controller, then the scheduler allocates that task to the VM controller and the VM controller use the VM machines to complete the task. Otherwise, the real time controller allocates the task to the Physical Machine (PM) for task completion. The efficiency of the system performance is increased by using this algorithm. The energy-aware scheduling is analyzed for independent, periodic real-time tasks in al clouds. The scheduling objective is to improve the system's schedulability for real-time tasks and to save energy. Here, the virtualization technique is deployed and TSAST is used to save energy efficiently. To extend work to implement various scheduling algorithms with deadline constraints and implement automatic energy-saving management and performance optimizations in modern cloud datacenters.

References

1. **Ankita Atrey, Nikita Jain and Iyengar (2013),'A Study on Green Cloud Computing' *International Journal of Grid and Distributed Computing* Volume 6, No.6, page no-93-102.**

2. **Devendra Singh Thakur(2014),' Energy Efficient Task Scheduling Algorithms In Cloud Data Center' , *A thesis submitted in partial fulfillment of the requirements for the degree of Master of Technology,* Volume 18,page no-18-25.**

3. **Ji Lia,b, Longhua Fenga,b, Shenglong Fang(2014), 'An Greedy-Based Job Scheduling Algorithm in Cloud Computing', *Journal Of Software*, Volume 9, No.4, page no-921-925.**

4. P Kowsik, K.Rajakumari (2014), 'A Comparative Study on Various Scheduling Algorithms in Cloud Environment', *International Journal of Innovative Research in Computer and Communication Engineering,* Volume 2, Issue 11.

5. Lipsa Tripathy, Rasmi Ranjan Patra (2014), 'Scheduling In Cloud Computing' *International Journal on Cloud Computing: Services and Architecture (IJCCSA),* Volume 4, No.5, page no-21-27.

6. Mehiar Dabbagh, Bechir Hamdaoui, Mohsen Guizani, Ammar Rayes(2015), 'Energy-Efficient Resource Allocation and Provisioning Framework for Cloud Data Centers' *IEEE Transactions on Network and Service Management*, Volume 12, No. 3, page no-377-391.

7. Marjan Kuchaki Rafsanjani and Amid Khatibi Bardsiri (2012),'A New Heuristic Approach for Scheduling Independent Tasks on Heterogeneous Computing Systems' *International Journal of Machine Learning and Computing*, Volume 2, No. 4, page no-371-376

8. Er. Shimpy, Mr. Jagandeep Sidhu(2014), 'Different Scheduling Algorithms In Different Cloud Environment', *International Journal of Advanced Research in Computer and Communication Engineering* ,Volume 3, Issue 9,page no-78-85.