RESEARCH ARTICLE   OPEN ACCESS

Efficiency Enhancing Resource Scheduling Strategies in Cloud Computing

Neha Solanki[1], Dr. N. C. Barwar[2]

1(Computer Science Department, Jai Narain Vyas University, and Jodhpur

Abstract:

The clouds have changed the patterns of traditional way of using the software and infrastructures. In cloud computing job scheduling is used to schedule certain jobs to particular resources at particular time. In this paper, some scheduling strategies are discussed which schedule computing resources in a way that providers achieve high resource utilization, low power consumption and users meet their applications performance requirements with minimum expenditure.

Keywords **— Scheduler, Virtual Machine, Round Robin.**

I. INTRODUCTION

The enterprises can rapidly complete some business and reduce a lot of cost by using clouds. The cloud computing system is divided into consumers, service providers and resource providers, which is currently the major way to layer the cloud computing. The service providers want to minimize the cost of using the resources offered by the resource providers, and to reduce the response time for consumers. An application operated in the cloud consisting of one or more services which is sent to the service provider stating two main constraints, time and cost [1, 2]. Service/job scheduling is one of the most important methods to achieve these. The actual processing time is longer than the original estimated time because of delays occurring on the provider's side. As the cloud computing are primarily operated by the principle of paying by time, so the service provider want to reduce the delay and improve the quality of their service [3]. The ultimate goal of scheduling in cloud computing is to have efficient resource utilization.

II. SCHEDULING PROCESS IN CLOUD COMPUTING

The process dealing with normal service requests from consumers is as follow:

a) The cloud service provider receives the service request of user's [4]
b) Execution of the receiving.
c) The process of service/job scheduling.
d) The process of resource allocation.

Service/job scheduling work is carried out in the step(c). The service providers have a huge number of users; they have to deal with massive data [5], which are more difficult to schedule. The requests from users must be scheduled efficiently, so scheduler needs to calculate a proper sequence to response those requests.

A. Main Components in Service/Job Scheduling

The main components performing Service/job scheduling are shown in Fig. 1, which is composed of classifier, scheduler, and summary component.
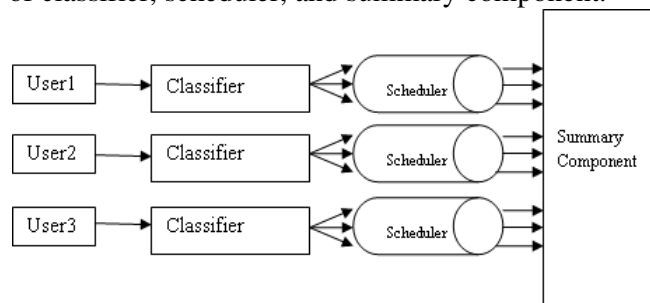

Fig. 1 Component in Service/Job Scheduling

a) Classifier. This component receives all requests from consumers, analyzes them and

classifies these requests into job units which are the units can be scheduled directly. The classifier analyzes the requests to decide how many job units every request can be divided to, and then classifier sends all job units to the next component scheduler.

b) Scheduler. Every scheduler has several schedule units, it push every job unit into the appropriate schedule unit according to the saturation and idle of every schedule unit. The schedule unit executes some algorithms to the job units, and then adjusts the job units' position in the queue to optimize the result.

c) Summary component. This component is the last one in the service provider. It summaries the job units from several schedule units in a cycle time, and sends them to the source provider.

B. The Process in Scheduler

The scheduler does the core process in the scheduling process, and implements some algorithms or strategies. The actual process of scheduling for service requests is a process of state transition. At a particular time, there are four job units in the scheduler with the order job unit1, job unit3, job unit4, and job unit2. By analyzing the priority of the job units waiting in the schedule unit, the order of these four task units changes into job unitl1, job unit2, job unit3 and job unit4.
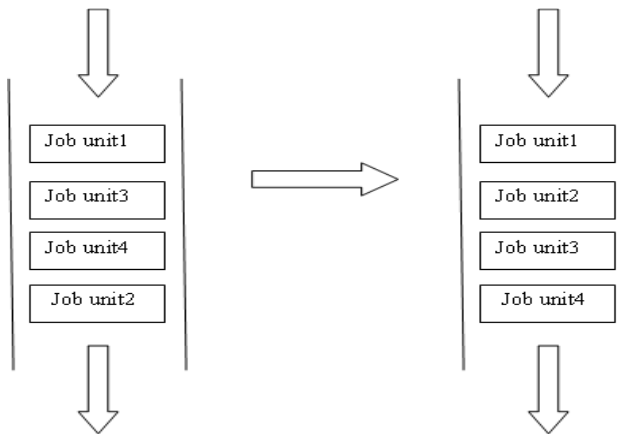

Fig. 1 Process of State Transition

III.   SCHEDULING ALGORITHMS

A. Round Robin

It uses the concept of time quantum or slices. Here the time is divided into multiple slices and each node is given a particular time quantum or time interval and in this quantum, the node will perform its operations. Using this algorithm, the scheduler allocates one virtual machine to a node in a cyclic manner. The round robin scheduling in the cloud is very similar to the round robin scheduling used in the process scheduling. The scheduler starts with a node and moves on to the next node, after a virtual machine is assigned to that node. This is repeated until all the nodes have been allocated at least one virtual machine and then the scheduler returns to the first node again.

B. Preemptive Priority

Preemption is a process that removes one or more previously scheduled activities according to certain criteria and re-allocates freed resource capacity to a new activity. A preemption policy is normally used for scheduling high priority activities when a capacity shortage appears. Priority is an important issue of job scheduling in cloud environments. As a scheduling policy, preemption has wide applications in many areas (e.g. Process scheduling, bandwidth allocation, manufacturing scheduling).

C. Shortest Job First

An SJF algorithm is simply a priority algorithm where the priority is the inverse of the next CPU burst. That is, the longer the CPU burst, the lower the priority and vice versa. SJF policy selects the job with the shortest processing time first. Shorter jobs are always executed before long jobs. One major issue with SJF is the need to estimate the processing time of each job.

IV.   CONCLUSIONS

Three types of scheduling strategies in cloud computing are discussed they are Round-Robin, Preemptive Priority, and Shortest Job First. Round-Robin scheduling utilizes all the resources in a balanced order. It has good response time but its power consumption is high. Preemptive priority scheduling is very complex but it has good makespan. Shortest Job First scheduling has

problem of starvation for long running jobs. Comparisons of these algorithms are show in table I.

TABLE I
COMPARISON OF SCHEDULING ALGORITHMS IN CLOUD

Features	Round-Robin	Preemptive Priority	Shortest Job First
Fairness	✓		
Complexity		✓	
Power Consumption	✓		
Starvation			✓
Response Time	✓		
Makespan		✓	

REFERENCES

1. *Gabor Kecskemeti, Peter Kacsuk "Automatic Service Deployment Using Virtualisation" 16th Euromicro Conference on Parallel, Distributed and Network-Based Processing.*

2. *Hai Zhong, Kun Tao, Xuejie Zhang "An Approach to Optimized Resource Scheduling Algorithm for Open-source Cloud Systems" 2009 Third International Conference on Multimedia and Ubiquitous Engineering.*

3. *SungHo Chin, JongHyuk Lee, SeongJin Song, Byungkwen Song, HeonChang Yu "Dynamic Service Scheduling for Workflow Applications" Third International Conference on Semantics, Knowledge and Grid.*

4. *Luqun Li "An Optimistic Differentiated Service Job Scheduling System for Cloud Computing Service Users and Providers" 2009 Third International Conference on Multimedia and Ubiquitous Engineering.*

5. *Wuu-Yee Chen, Tsang-Yean Lee, Nai-Wen Kuo,Yuan-Hwa Mei, Shan-Kai Chen, Yu-Chen Yan, "Job Schedule Model Based on Grid Environment" 2009 Fifth International Joint Conference on INC, IMS and IDC.*