ISSN:0975-9646

A Survey of Resource Scheduling Algorithms in Green Computing

Arshjot Kaur, Supriya Kinger
Department of Computer Science and Engineering, SGGSWU,
Fatehgarh Sahib, Punjab,
India (140406)

Abstract-Cloud computing has emerged as an optimal way of sharing and providing resources over the internet. Its service models have helped a lot in providing sources easily to the users. Green computing is now gaining a lot of importance and is an intense situation where all the major issues need to be resolved. Rapid growth of the internet, increasing cooling requirements and increased rate of power has led to adoption of green computing. Implementation of green computing has become important so as to secure the environment. Emphasis has been laid on virtualisation, power management, material recycling and telecommunicating. Still,a lot needs to be done. The work habits of computer users and business can be modified to minimise adverse impact on the global environment. Job scheduling is one of the challenging issues in green computing. A lot of work needs to be done on deciding the priority of jobs and time issues so as to provide efficient execution of user jobs. Assigning the right priority with reduced time and less energy consumption needs to be focussed on.

Keywords- Cloud Computing, Green Computing, Job Scheduling

1. INTRODUCTION

A. Cloud computing

Rather than having local servers or personal devices to handle applications, cloud computing relies on sharing the computing resources. Basically the word cloud is used as a metaphor for "internet". Therefore, cloud computing can be defined as the internet based computing where resources are retrieved from the web based tools and connections instead of direct connection to the server[1].

B. Types of cloud

Cloud Computing is classified into three types :

1. Public cloud : In public cloud, service and infrastructure are provides off-site over the internet. Services may be offered on pay by usage mode. Public cloud is used when we need to test and develop application mode or when we are doing collaboration projects. Example of public clouds is IBM's blue cloud.
2. Private cloud: in private cloud, services and infrastructure are maintained on a private network. These clouds are said to be the most secure and offer greatest level of control and reduces the cost savings. Private clouds are used when the company is large enough to run a next generation cloud data center efficiently and effectively on its own.
3. Hybrid cloud: It includes both public and private clouds with multiple providers. These clouds are used when the company is concerned more about security and wants to use Saas application.

C. Cloud service models

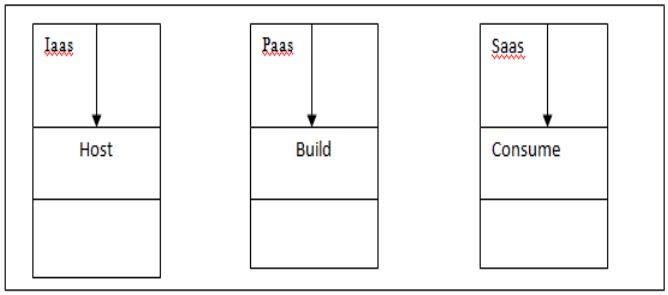

Figure 1.Cloud service models

Figure 1 explains the following:

1. Infrastructure as a service (Iaas): It is the first layer of loud computing. This model helps in managing our applications, data operating system, middleware and runtime. This service provider manages virtualisation, servers, networking and storage. The basic advantage which this model provides is that extra data processing space is available whenever we need. Some of the Iaas users are Amazon, Microsoft, rack space and red hat.
2. Platform as a service (Paas): This is the second layer of cloud computing . We manage our own applications and data and the cloud vendor manages everything else. Its basic advantage includes streamlined version deployment and the ability to change or upgrade and minimize expenses. One of the popular Paas is Google app engine.
3. Software as a service (Saas) : this is the final and the last layer of cloud services models. Here ,almost all the services are managed by cloud vendor. There is no need to pay extra licensing fees and new users can also be added. Its examples are online banking. Gmail and hotmail. By using Saas, we are easily able to access the software from a variety of devices in the office or on go[2].

D. Actors that participate in the transaction process and performs tasks in the cloud computing

CLOUD CONSUMER	Person/organisation that maintains business relationship and uses services from cloud provider
CLOUD PROVIDER	Entity responsible for making a service available to interested parties
CLOUD AUDITOR	Party that conducts independent assessment of cloud services and security of cloud implementation
CLOUD BROKER	Entity that manages the use, performance and delivery of cloud services.
CLOUD CARRIER	Provides transport of cloud services from cloud provider to cloud consumer

E. Green computing

In simple words, green computing refers to environmentally sustainable computing or IT. It is basically environmentally responsible and eco-friendly use of computers and their resources. In broader terms, it is defined as the study of designing , manufacturing , using and disposing of computing devices in a way that reduce their environment impact. It is also the development of environmentally sustainable production particles, energy efficient computers and improved disposable ad recycling products. Increased energy use has increased IT costs. Growing appreciation of risks of climate change and increasing concerns about energy security has elevated green computing to global issues. The desire to lessen energy use has also given rise to green computing.

F. Resource Scheduling

The allocation of system resources to various tasks is known as job/ resource scheduling. The system maintains the prioritizes queues of jobs waiting for CPU time and must decide which job to take from the queue and how much time to allocate to it so that all jobs are completed in a fair and timely manner[3].

2. Existing Algorithms For Resource Scheduling

There are Scheduling Strategies which divide the scheduling algorithms into following three parts:

- Load based scheduling algorithm
- Temperature based scheduling algorithms
- Other types of scheduling algorithms

Here we will study all the types of algorithm starting with the load based algorithms.

2.1 Load Based Scheduling Algorithms

A. First Come First Serve (FCFS)

The most intuitive and simplest technique is to allow the first process submitted to run first. This approach is called as first-come, first-served(FCFS) scheduling. In effect, processes are inserted into the tail of a queue when they are submitted. The next process is taken from the head of the queue when each finishes running. The process is allocated to the CPU which has least burst time. A scheduler arranges the processes with the least burst time in head of the queue and longest burst time in tail of the queue. This requires advanced knowledge or estimations about the time required for a process to complete[4].The idea is illustrated in Figure 2

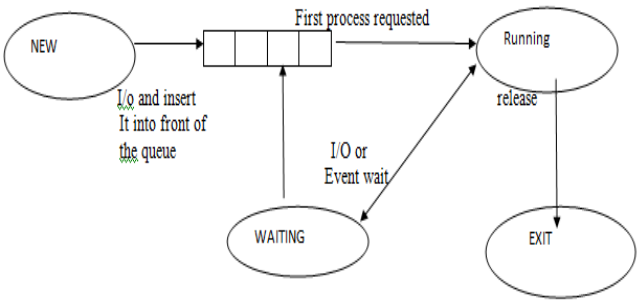

Figure 2. First come first serve Schdeluing

Characteristics

1. The lack of prioritization does permit every process to eventually complete, hence no starvation.
2. Turnaround time, waiting time and response time is high.
3. One, Process with longest burst time can monopolize CPU, even if other process burst time is too short. Hence throughput is low.

Algorithm FCFS:

- Initialize Tasks.
- First task assigned to the queue and add tasks up to n numbers.

Add next task 'I' at last position in the main queue

B. Shortest Job First

The process is allocated to the CPU which has least burst time. A scheduler arranges the processes with the least burst time in head of the queue and longest burst time in tail of the queue. This requires advanced knowledge or estimations about the time required for a process to complete. This algorithm is designed for maximum throughput in most scenarios. The process is allocated to the CPU which has least burst time. A scheduler arranges the processes with the least burst time in head of the queue and longest burst time in tail of the queue. This algorithm is designed for maximum throughput in most scenarios This idea is illustrated in the figure 3.

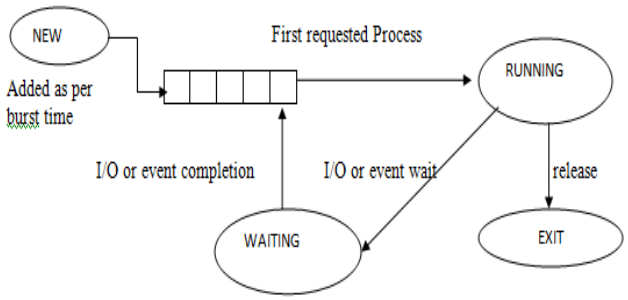

Figure 3. Shortest Job First Scheduling

Characteristics

1. The real difficulty with the SJF algorithm is, to know the length of the next CPU request.
2. SJF minimizes the average waiting time because it services small processes before it services large ones. While it minimizes average wait time, it may processes with high service time requests. If the ready list is saturated, then processes with large service times tend to be left in the ready list while small processes receive service. In extreme case, when the system has little idle time, processes with large service time will never be served[5].

Algorithm SJF:

- for i = 0 to i < main queue-size
  - if task $_{i+1}$ length < task $_{i}$ length then
    - add task $_{i+1}$ in front of task $_{i}$ in the queue
  - end if
  - if main queue-size = 0 then
    - task $_{i}$ last in the main queue
  - end if
- end for

C. ROUND ROBIN SCHEDULING

In this algorithm, a time slot gets decided for every job. A job would execute only that much time and after then next job would be executed. The ready processes are kept in a queue. The scheduler goes around this queue, allocating the CPU to each process for a time interval of assigned quantum. New processes are added to the tail of the queue [3].Figure 4 explains about this algorithm more clearly.

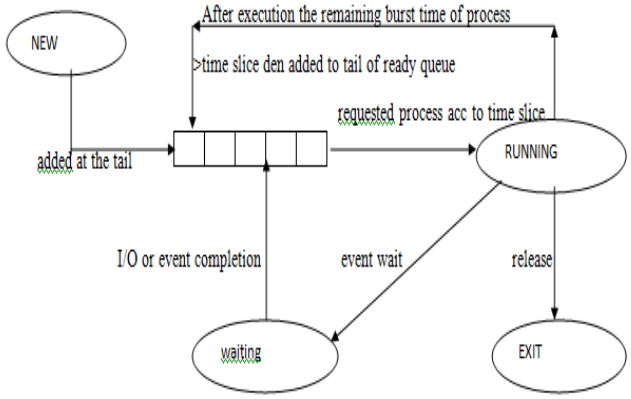

Figure 4. Round Robin Algorithm

Characteristics

1. Setting the quantum too short causes too many context switches and lower the CPU efficiency.
2. Setting the quantum too long may cause poor response time and approximates FCFS.
3. Because of high waiting times, deadlines are rarely met in a pure RR system.

Algorithm RRS:

- Keep the ready queue as a FIFO queue of processes.
- New processes are added to the tail of the ready queue.
- The CPU scheduler picks the first process from the ready queue, sets a timer to interrupt after 1 time slot, and dispatches the process.
- The process may have a CPU burst of less than 1 time quantum.
    - In this case, the process itself will release the CPU voluntarily.
    - The scheduler will then proceed to the next process in the ready queue.
- Otherwise, if the CPU burst of the currently running process is longer than 1 time quantum,
    - the timer will go off and will cause an interrupt to the OS.
    - A context switch will be executed, and the process will be put at the tail of the ready queue.
    - The CPU scheduler will then select the next process in the ready queue.

D. PRIORITY BASED SCHEDULING

In this algorithm, every job is assigned with a priority. The high priority job would get executed first and then later on, low priority jobs would execute. The O/S assigns a fixed priority rank to each process. Lower priority processes get interrupted by incoming higher priority processe[6]. This idea is illustrated in figure 5.

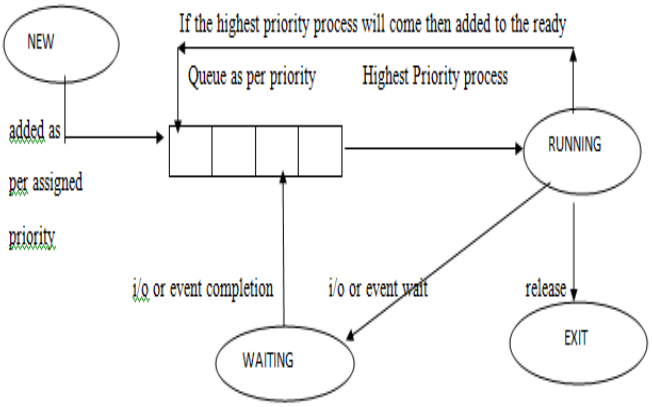

Figure 5. Priority Scheduling

Characteristics

1. Starvation can happen to the low priority process.
2. The waiting time gradually increases for the equal priority processes
3. Higher priority processes have smaller waiting time and response time.

Algorithm PSA

- for i = 0 to i <main queue-size
    - if priority (task $_{i+1}$) > priority (task $_{i}$) then
        - add task $_{i+1}$ in front of task $_{i}$ in the queue
    - end if

end for

2.2 Temperature Based Scheduling

Cloud computing has rapidly emerged as a widely accepted computing paradigm, but the research on Cloud computing is still at an early stage. Cloud computing provides many advanced features but it still has some shortcomings such as relatively high operating cost and environmental hazards like increasing carbon footprints. These hazards can be reduced up to some extent by efficient scheduling of Cloud resources. Working temperature on which a machine is currently running can be taken as a criterion for Virtual Machine (VM) scheduling. There is a new proactive technique that considers current and maximum threshold temperature of Server Machines (SMs) before making scheduling decisions with the help of a temperature predictor, so that maximum temperature is never reached. Different workload scenarios have been taken into consideration. The results obtained show that the proposed system is better than existing systems of VM scheduling, which does not consider current temperature of nodes before making scheduling decisions. Thus, a reduction in need of cooling systems for a Cloud environment has been obtained and validated[7].

A. Core level scheduling
At the core level, the scheduler determines the predicted temperature of each core and migrates the job from hot to cold ones.

B. Socket level scheduling
At the socket level, there is a scheduler which manages the job between sockets. It takes temperature, performance, and fan speed information as an input. The techniques used are spreading and consolidation

```
IF (Event(Application_Arrival) ==TRUE)
  Put new request in waiting queue
  FOREACH physical machine in the Unified list
    If (current temperature < threshold temperature)
       Estimate  for each node and put  in the predict queue
    END IF
  END FOREACH
  Sort predict Queue in the decreasing order of
  FOREACH VM request in waiting queue
    Check availability of VM in the first node in predict queue
    IF Available then allocate VM to this VMRequest
      Update Unified and Local lists.
    ELSE
      Check next node in the Predict Queue till last node is
reached
    END ELSEIF
    IF no node is available
      Event(Activate_SM) and allocate VMRequest to a new
VM on this SM.
    ENDIF
  END FOREACH
ENDIF
IF (Event(Application_Complete) ==TRUE)
    Event(Free_VM)
    Update Unified and local lists
  IF(Active_VM_Count == 0)
    Event(Hibernate(SM)).
    Update Unified and Local lists.
  ENDIF
ENDIF
```

Algorithm 1: Proactive thermal VM scheduling

2.3 Other types of scheduling algorithms

A. Bee algorithm

It is a nature inspired algorithm which tries to track the activities of bee to get their food. First they select scout bee to go and search a wide domain of areas, if a scout bee finds a potential food resource it returns to its hive and does a waggle dance which tells other bees the direction and the distance of the potential food resource. A set of selected bees goes to the food resource and starts bringing in the honey while other scout bees does the same work and sets of bees are sent to different location to bring in the food. After every identification of a food resource the scout bee informs others and sets its course for other new sites nearby the potential food resource.

It has many Applications which are as follows :

- Training neural networks for pattern recognition.
- Forming manufacturing cells.
- Scheduling jobs for a production machine.
- Solving continuous problems and engineering optimization.
- Finding multiple feasible solutions to a preliminary design problems.
- Data clustering
- Optimising the design of mechanical components.
- Multi-Objective Optimisation.
- A fuzzy logic controller for a robot gymnast.
- Vision and Image Analysis

B. Ant Colony Algorithm

It is same as bee algorithm. It is work done as bee algorithm the first algorithm was aiming to search for an optimal path in a graph, based on the behaviour of ants seeking a path between their colony and a source of food.

Its Applications are as follows:

- Ant Colony Optimization with Fuzzy Logic
- This method introduces fuzzy intelligence into ants to accelerate searching ability

C. Work Flow Algorithm

Workow scheduling is the problem of mapping each task to appropriate resource and allowing the tasks to satisfy some performance criterion. Work Flow is processes that consist of a series of steps which simplifies the complexity of executions and management of applications.

Work flow algorithm has many Advantages .With the emerging of cloud computing, cloud work flow systems are designed to facilitate the cloud infrastructure to support large scale distributed collaborative e-business and e-science applications. The management and scheduling of resources in Cloud environment is complex, and therefore demands sophisticated tools for analysis the algorithm before applying them to the real system.

It also has certain Disadvantages .Existing work flow algorithms does not consider the execution time. Therefore there is a need to implement a new scheduling algorithm that can minimize the execution time in cloud environment. Moving work Flows to a cloud computing environment enables the use of various cloud services to facilitate work Flow execution.

D. Genetic algorithm

This work describes a genetic algorithm approach to resource-constrained scheduling using a direct,time-based representation .This document describes a genetic algorithm for finding optimal solutions to dynamic resource constrained scheduling problems. The genetic algorithm was applied to over 1000 small job shop and project scheduling problems (10-300 activities, 3-10 resource types).The algorithm performed fairly well on a wide variety of problems. Genetic algorithms operate on a population of solutions rather than a single solution and employ heuristics such as selection, crossover, and mutation to evolve better solutions[10]

This algorithm is very beneficial also. It is Used to generate useful solutions to optimization and search problems. Whereas it has certain Disadvantages also. The genetic algorithm performed well on some problems that were very difficult for the branch and bound techniques (i.e. the

branch and bound method took a long time to and the optimal solution the genetic algorithm did not perform well on problems in which the resources were tightly constrained. This comes as little surprise since there presentation forces the genetic algorithm to search for resource- feasibility, and tightly constrained resources mean fewer resource-feasible solutions the genetic algorithm did not perform well on the job shop problem

Its Applications are as follows :

- Bioinformatics
- Phylogenetics
- Computational science
- Engineering
- Economics
- Chemistry
- Manufacturing
- Mathematics

Two basic metrics were collected to compare the performance of the algorithms:

1. Average Response Time: This the sum of the wait time plus the execution time of a job, averaged over all jobs. This metric captures the ability of the scheduling algorithm to service the smaller, shorter jobs. Since the job stream contains a large number of smaller short jobs, the metric provides a good measure of the quality of the scheduling algorithm.
2. Weighted Average Response Time: The weight of a job is denned as the product of its resource requirement and execution time. The weighted average response time is then the product of the job weight and the job response time, averaged over all jobs. This is basically a measure of how well the scheduling system provides progress to the large jobs. It is also a measure of overall resource utilization over time.

3. CONCLUSION AND FUTURE WORK

A. Conclusion

After studying about all the above algorithms, we find that Round Robin Algorithm has the largest waiting time whereas Priority algorithm has the least waiting time than FCFS Algorithm and SJF Algorithm. When we talk about complexity which defines which type of algorithm is simple or easy to use in processing, FCFS Algorithm is the simplest Scheduling Algorithm. Both SJF and Priority Algorithm become difficult to understand and code whereas the performance of Round Robin Algorithm depends heavily on size of time quantum. In case of FCFS Algorithm, CPU is allocated in the order in which the processes arrive whereas in SJF Algorithm CPU is allocated to the process with least CPU burst time. Based on priority in the Priority Algorithm, the higher priority job can run first. In case of Round Robin Algorithm ,the preemption take place after a fixed interval of time.

B. Future work

Future work can explore designing of an efficient scheduler and making group of the jobs according to the energy consumption of the jobs as a high priority and low priority job. Also along with this making a dependency mapping concept so that the dependent jobs gets executed first and evaluating the results accordingly and to compare it with previous algorithms.

REFRENCES

[1] Savitha. P, J Geetha Reddy ,"A Review Work On Task Scheduling In Cloud Computing Using Genetic Algorithm", International Journal of Scientific & Technology Research Vol. 2, Issue 8, August 2013.

[2] Masnida Hussin, Young Choon Lee, and Albert Y. Zomaya, "Priority-based scheduling for Large-Scale Distribute Systems with Energy Awareness",Ninth IEEE International Conference on Dependable, Autonomic and Secure Computing, 2011, pp.503- 509.

[3] Tejinder Sharma, and Vijay Kumar Banga, "Efficient and Enhanced Algorithm in Cloud Computing", International Journal of Soft Computing and Engineering (IJSCE), Volume-3, Issue-1, March 2013, pp. 385-390.

[4] Lu Huang, Hai-shan Chen and Ting-ting Hu, "Survey on Resource Allocation Policy and Job Scheduling Algorithms of Cloud Computing" ,Journal of Software, Vol. 8, No. 2, February 2013, pp. 480-487.

[5] Van den Bossche, R., Vanmechelen, K., Broeckhove, J, "Cost Optimal Scheduling in Hybrid IaaS Clouds for Deadline Constrained Workloads", in 3rd IEEE International Conference on Cloud Computing, Miami (July 2010)

[6] Upendra Bhoi, and Purvi N. Ramanuj, "Enhanced Max-min Task Scheduling Algorithm in Cloud Computing",International Journal of Application or Innovation in Engineering & Management(IJAIEM), Volume 2, Issue 4, April 2013, pp. 259-264.

[7] Supriya Kinger,Rajesh Kumar,Anju Sharma,"Prediction based proactive thermal virtual machine scheduling in green clouds" The Scientific World Journal,Vol 2014,Article ID 208983,12 pages,2014,doi:10.1155/2014/208983

[8] Ram Kumar Sharma, Nagesh Sharma, "A Dynamic Optimization Algorithm for Task Scheduling in Cloud Computing With Resource Utilization", International Journal of Scientific Engineering and Technology, Volume No.2, Issue No.10, pp. 1062-1068

[9] Hong Sun, Shi-ping Chen, Chen Jin, and Kai Guo1,"Research and Simulation of Task Scheduling Algorithm in Cloud Computing", TELKOMNIKA, Vol.11, No.11, November 2013, pp. 6664-6672.

[10] Philippe Baptiste, Marek Chrobak ,Christoph D¨urr1," Polynomial Time Algorithms for Minimum Energy Scheduling", *ieee* 2010

[11] Neetu Goel, R.B. Garg, " A Comparative Study of CPU Scheduling Algorithms", *International Journal of Graphics & Image Processing* ,Vol. 2 , Issue 4,November 2012

[12] Ankur Bhardwaj " Comparative Study of Scheduling Algorithms in Operating System", International Journal of Computers and Distributed Systems, Vol. No.3, Issue I, April-May 2013

[13] V. Krishna Reddy, L.S.S.Reddy, "A Survey of Various task Scheduling Algorithms in Cloud Environment", Global Journal Engineering and Applied Sciences (GJEAS),Volume 2. No 1 , January 2012.

[14] W. Kubiak, B. Penz, D. Trystram, "Scheduling chains on uniform processors with communication delays," *Journal of Scheduling*, Vol. 5, No. 6, pp. 459-476, 2002

[15] Jasbir Singh, Gurvinder Singh, "Task Scheduling using Performance Effective Genetic Algorithm for Parallel Heterogeneous System", *International Journal of Computer Science and Telecommunications*,Vol 3, Issue 3, March 2012.