Job Scheduling for High Performance Computing and Cloud Environment

Lett Yi Kyaw, Sabai Phyu
University of Computer Studies, Yangon, Myanmar
lettyikyaw@ucsy.edu.mm, sabaiphyu72@gmail.com

Abstract

High performance computing (HPC) is famous and latest trend in current research era. User request many jobs and they want to get accurate results and quick time within eye blink. There are so many processors running large data centers and large systems. So many processes will work large data center and large system. Many jobs within high performance scheduling environment is preparing the job to understand and calculate to get the improve result and to reduce the processing time in a system. Requests for greater capacity were met by building more powerful systems with more compute nodes, scalability and higher processor operating frequencies for business applications, education, and scientific application. There are various ways to schedule parallel jobs but in this paper discusses some of the technical challenges ,methods of scheduling and one of proposed system in cloud (HPC) environment.

Keywords: *HPC, preparing jobs, technical challenges, scheduling methods*

1. Introduction

The recent trend of increasing performance in a system is very interested field in current research areas. Many researchers are searching methods on both hardware and software. There are many challenges in these. Demands for greater capacity were met by building more powerful systems with more compute nodes, greater transistor densities, and higher processor operating frequencies. Unfortunately, the scope for further increases in processor frequency is restricted by the limitations of semiconductor techno- logy. Instead, parallelism within processors and in numbers of compute nodes is increasing, while the capacity of single processing units remains unchanged. The volume of data generated by scientific applications grows and the force on the input/output system of high performance computing (HPC) systems also increases.

The HPC is classified into three parts: Cloud Computing, Grid Computing and Cluster Computing. Grid Computing is similar to distributed computing as information is stored and accessed from/in different multiple locations and use of system resources, decentralized the system and interconnected the systems. Cluster computing is to solve cost of hardware which consists multiple low cost computers and interconnected to each other and synchronous their work and very popular method for economic, performance, and flexibility. Cloud Computing is combination of virtualization and different resources that can be used computing resources.

High Performance Computing applications are ravenous. They will, at times, consume all of the computer resources in a given system. Machine learning applications also require significant computing resources in order to power their algorithms. While domains that include manufacturing, energy exploration, life sciences and basic research can often scale their applications, either within an SMP server or across systems using communication protocols such as MPI, that scaling is often limited to a few thousand cores. But the benefits of getting results faster cannot be overlooked. As more applications have been re-coded to take advantage of GPUs, performance has increased significantly compared to traditional CPU-based implementations.

HPC platforms are cluster-based computing systems consisting of loosely coupled computers connected through fast networks for an efficient, reliable and quick execution of advanced applications. The term “supercomputer” is used to refer to a more powerful subset of HPC platforms, systems that perform at or near the currently highest operational rate for computers. While there are few supercomputing centers, HPC platforms can be found in any research.

Grid Computing has the following issues and challenges [10]:

- Sharing of Applications and Data
- User-friendly Environment
- Centralized Management
- Online Control
- Standard Protocols

And then cluster computing also has the following issue and challenges [11]:

- Middleware
- Program
- Elasticity
- Scalability

Cloud computing has the following issues and challenges [12]:

- Automated service provisioning
- Virtual machine migration process
- Server combination to single unit
- Energy Management
- Traffic controlling system
- Data Security and forensics
- Software frameworks

The most efficient scheduling is a very complicated when realistic models are observed. Unlike many anterior jobs that only observe scheduling algorithms, many researcher need to search a detailed insight into the complexity of the problem, using many information collect and analysis from real life. Especially, how have we stressed out several component of the system interact together and to affect the resulting performance. In many different factors to study and check when observing the "best" scheduling algorithm, some of them are:

CPU Utilization

Despite the development in parallel computing architecture, an alternative way to optimize the utilization of the heterogeneous resources is to create a unified abstraction and hide the heterogeneity from workload. Several high level programming libraries are thus developed to unify CPU/GPU computing and to make the workload partition transparent to the programmer. An important factor in all scientific applications is the starvation of resources, just because the large amount of data and processing power required is the use of the CPU. We must therefore search for a new mathematical model for future research areas in order to reduce costs, carbon emissions, HPC workload and CPU energy efficiency [9].

Throughput

One of the common types of HPC workloads is throughput. In throughput workloads, multiple individual jobs run simultaneously to complete a given task, each running independently with no communication between jobs. HPC throughput workloads are computing intensive and require scalable and highly performance infrastructure.

Turnaround Time

The various scheduling techniques, the effective task scheduling algorithms are implemented with the help of simulation tool and the result obtained reduces the total turnaround time and also increase the performance.

Waiting Time

The waiting time of a job based on the number of compute-nodes it needs that request more nodes wait longer than jobs that request fewer nodes.

Load Average

The numbers which calculate the jobs processing in the queue waiting for their playing time to get into the system of the CPU is defined as load average.

Response Time

The total time required to respond to a service request (a job to be serviced by the user) is the sum of the service time and waiting time.

2. Background Theory

There are three methods in scheduling: resource scheduling, task scheduling and workflow scheduling. Resource scheduling is managing the resource to get the user request by mapping of virtual resources relies on physical machines. Choosing which hardware resource to use or not for some application workload and spatial work mapping are resource mapping process.

Task scheduling is divided into two methods which are distributed scheduling and centralized scheduling, worked in homogeneous environment upon dependent or independent tasks.

Workflows are represented as Directed Acyclic Graphs (DAG) i.e., each vertex represents one or more work tasks and the edges express control or data dependencies between the (vertices) tasks. Users usually follow one of two strategies to submit a workflow. The strategies balance between lower resource consumption (as chained jobs) and shorter turnaround time (as a pilot job). The metrics used to compare are compare experiment results and method to calculate them are wait time, runtime, turnaround time, actual utilization and job's slow down.

Many clients send various jobs and the system to be processed on computational resources consist of processing portions. Efficient job scheduling is important and hard to manage issues in homogeneous and heterogeneous computing fields.

We studied the process that wide variety of scheduling algorithms are available, we need a methods to select among them. In figure 1 described the component of scheduler.

2.1 Job Scheduling Algorithms

There are different types of job scheduling algorithms, some of which are there:

First Come First Serve (FCFS): First Come First like normal queue, simple, fair, but poor performance. Average queuing time may be long. In this system, if one job is not complete data processing this will have to wait to go back through.

First Come First Serve with Best Fit (FCFS/BF): When the first arrival job is processed first, but it does not block. When a process enters the ready queue, its PCB is connected to the tail of the queue and the algorithm selects the next eligible job.

Shortest Job First (SJF): SJF is the best method and resize the waiting time. Definite period brought with the process needs to be previously acknowledged to the processor. However this is an idealization, although it can be implemented in an approximate manner, the actual SJF according to the definition cannot be implemented.

Longest Job First (LJF): Choose the longest job and this algorithm is another opposite method of SJF. If more than one jobs the same highest, the algorithm chooses the first largest job to wait.

Score-Based Priority (SBP): Score based deadline constraint workflow scheduling algorithm. The algorithm is foundational to the concept of score and distribution of deadlines. The resources or virtual machines are assigned the score depending on the MIPS, MOPS and space available and the task, subtasks are assigned the sub-deadline because then the processing never regains the deadline set by the end user.

Multi-Queue Priority (MPP): Each queue has its algorithm both for scheduling. Then one or more of these (conceivably prioritized) algorithms arbitrate between queues. Method is complex but flexible for scheduling process. So this process could separate system processes, interactive, batch, favored, unflavored process.

2.2 Task Mapping Algorithm

Scheduling refers to a group of policies and mechanisms for controlling the arrangement or order of the task that a computer system should handle. The communication on their allocated resources ensures that tasks on the same resource are mutually excluded. Depending on the execution model, task mapping is divided into two: static mapping and dynamic mapping. Mapping the workflow onto available resources needs a broker to schedule jobs or tasks to the resources. A quality mapping and scheduling is one that optimizes the design goals, including minimizing energy consumption, maximizing timing performance or balancing the use of memory and other factors, in addition to meeting the design constraints.

Random: It is a method that selects and works the state/input sample, a random set of unoccupied calculation nodes.

Round-Robin: The performance of the Round Robin (RR) depends heavily on the size of the time quantum. If the time quantum is profoundly and astronomically immense, the RR policy is equilibrium to the FCFS policy. The length of the quantum is the critical issue with the RR policy. If it is too short, the CPU will spend more time switching the context. Otherwise, there will be interactive processes.

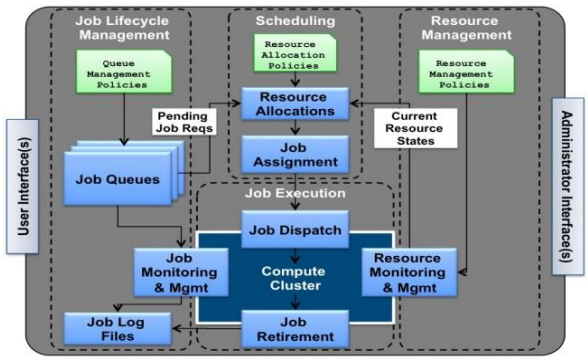

Figure 1. Key components of schedulers

The main advantage of job scheduling is the systemic management of processes, high performance computing and the best system performance. Traditional job scheduling algorithms cannot systematically and completely control the scheduling in the cloud environments. In figure 2 described the scheduling methods in cloud computing.

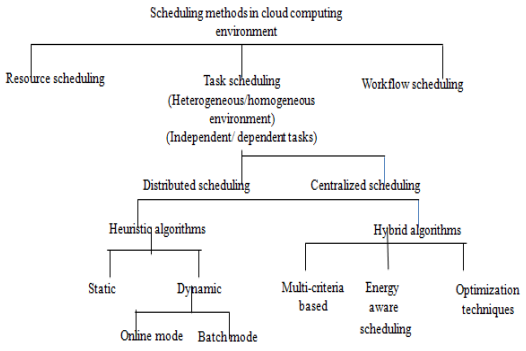

Figure 2. Types of scheduling

Table 1. Comparison of existing scheduling algorithm

Scheduling Method	Parameters Considered	Advantages	Disadvantages
Min-Min algorithm	Makespan,Expected completion time	Better makespan compared to other algorithms	Poor load balancing and QoS factors are not considered
Priority based job scheduling algorithm	Priority of tasks ,Expected completion time	Priority for scheduling is considered. Designed based on multiple decision-making criteria.	Improvement in performance can be considered for the consistency and complexity of the proposed method. Job having low priority will be lost when the system crashes. Starvation for resources they need.

3. Related Works

Many researchers try to reduce energy savings and to create and research green cloud computing. The one challenge in cloud computing is not only high energy consumption, but also high carbon emissions, which are not sustainable in the environment. These [1] scheduling policies are able to reach and reduce cost to the economic target and satisfied into their percent of energy saving. A new priority based job scheduling algorithm to solve issues related to complexity, consistency and make span [2]. Some previous works consider how to save energy for the execution of HPC applications worldwide.

D.J.Bradley et al. [3] proposed algorithms to minimize the use of workload history power and to predict future workloads in an acceptable and reliable manner. B.Lawson et al.[4] proposed an energy saving scheme as a novel with a dynamic adjustment of the number of CPUs in sleep mode when the utilization is low. Performance, power and availability behavior in data center is resized this center by using this application of nonlinear function approximation [5]. Ismali Ari et al. explore the models and problems as well as the performance benefits of hybrid job scheduling over working together and sharing physical clusters. The currently developed and supported clustering technologies include MPI, Hadoop-MapReduce, and NoSQL systems [6].

Zhou, S. eta l. presented a method using queuing network facilities to evaluate the performance of load allocating algorithms [7]. The system adjusts the order of execution of sent MPI jobs according to their key requirements and priorities. The authors [9] described the system that additional savings with task-aware cluster scheduling are possible above the MPI level. I. Ari et al. presented the implementation of a new Finite Element Analysis (FEA), Cloud Computing Service and an end-to-end discussion on technical design issues. The direction describes the performance of linear and non-linear mechanical analysis workloads on multi-core and multi-node computing resources [8].

4. Proposed System

Many existing scheduling algorithms are using standard and compare with improve new scheduling algorithm according to many parameters. In this section, improve new scheduling algorithm from min-min and priority based scheduling algorithm is proposed and described in figure 3.

Initialized the scheduler:

Step 1: Incoming jobs are classified their categories (size, type)

Process tasks:

Step 2: Sorting their categories (Calculate minimum load capacity of each vm in vm_list and sorting ascending order) and calculate minimum execution time process with min-min algorithm and assign job to vm .

Step 3: Calculate idle VM (starting time, executing time and completion time)

Step 4: Place priority job in appropriate VM using priority based scheduling algorithm.

Step 5: Calculate Throughput, Turnaround and Response Time.

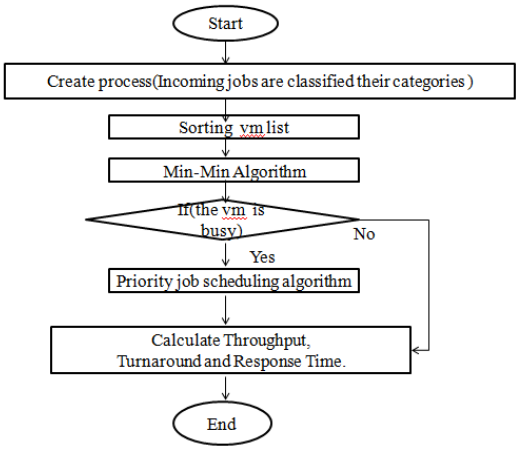

Figure 3. Flow chart of the proposed system

The proposed system will hope high performance compare existing system.

5. Conclusion and future work

Jobs can be divided into different categories according to the needs of users, and then set the best running time on the basis of different goals for each task. It will improve the QoS of task scheduling indirectly in a cloud environment.

Scheduling was optimized for trust and makespan, other factors like cost, bandwidths are considered. Relationship between resources and their impact on of different applications/use cases execution is investigated further to improve cloud scheduling. The proposed system will intend to not only study the performance of various job scheduling algorithms with sufficiently detailed models for resource contention, job interaction and interference on specific target HPC systems but also to get the improve performance of different parameters in HPC system. Many researchers are searching but until makespan is still remain in issues as [2].

References

[1] Saurabh Kumar Garg, Chee Shin Yeo, Arun Anandasivam, Rajkumar Buyya, " Energy-Efficient Scheduling of HPC Applications in Cloud Computing Environments ",Preprint submitted to Elsevier, Department of Computer Science and Software Engineering, The University of Melbourne, Australia,18 Septmber,2018.

[2] Shamsollah Ghanbari,Mohamed Othman,"A Priority based Job Scheduling Algorithm in Cloud Computing", International Conference on Advances Science and Contemporary Engineering 2012 (ICASCE 2012),Dept of Communication Tech and Network Faculty of Computer Science and Information Technology,Universiti Putra Malaysia,43400 UPM Serdang,Selangor D.E.,Malaysia,2012.

[3] D. Bradley, R. Harper, S. Hunter, Workload-based power management for parallel computer systems, IBM Journal of Research and Development 47 (5) (2003) 703–718.

[4] B. Lawson, E. Smirni, Power-aware resource allocation in high-end systems via online simulation, in: Proceedings of the 19th Annual International Conference on Supercomputing, Cambridge, USA, 2005.

[5] G. Tesauro, et al. Managing power consumption and performance of computing systems using reinforcement learning, in: Proceedings of the 21st Annual Conference on Neural Information Processing Systems, Vancouver, Canada, 2007.

[6] Ismail Ari,Ugur Kocak,"Hybrid Job Scheduling for Improved Cluster Utilization",Conference Paper,DOI: 10.1007/978-3-642-54420-0_3,January 2004.

[7] Zhou, S., Zheng, X., Wang, J., and Delisle, P. (1993). "Utopia: a load sharing facility for large, heterogeneous distributed computer systems. Software: Practice and Experience", 1993,23(12), 1305-1336.

[8] I. Ari, N. Muhtaroglu, "Design and implementation of a finite element service in the cloud", Advances in Engineering Software, Nov, 2013.

[9] https://www.studytonight.com/operating-system/cpu-scheduling.

[10] Manoj Kumar Mishra, *et al*., "A Review and Classification of Grid Computing Systems", *International Journal of Computational Intelligence Research*, ISSN: 0973-1873, vol. 13, no. 3, pp. 369-402, 2017

[11] Mohammad Alamgeer and Iqrar Ahmad, "Cloud Computing: Is a Step Ahead?", *International Journal of Computer Applications (0975 – 8887)*, vol. 159, no. 5, February 2017.

[12] Qi Zhegen, "Cloud Computing: State-of-the-art and Research Challenges", *Journal of Internet Service Applications*, Vol. 1, pp. 7-18, 2010.