Chapter 63
A Novel Task Scheduling Algorithm in Heterogeneous Cloud Environment Using Equi-Depth Binning Method

Roshni Pradhan
KIIT University, India

Amiya Kumar Dash
KIIT University, India

ABSTRACT

Cloud computing is modern tool for large-scale distributed computing and parallel processing. It has become a growing technology to deliver highly scalable service to the user. Task scheduling is one of the essential strategies to expeditiously utilize the potential of heterogeneous computing systems. In heterogeneous framework mapping, a task to a machine is a NP complete problem. This issue can be comprehended just utilizing heuristic approach. There are various heuristic approaches that were proposed to deal with scheduling of independent tasks. Different scheduling measures can be utilized for measuring the potency of scheduling algorithms. The most essential of them are makespan, flow-time, and overall resource utilization. Cloud generally is a single machine or combination of machines. Applications in the form of set of tasks are processed by the cloud.

INTRODUCTION

Cloud computing has gained a tremendous popularity nowadays. It delivers highly scalable services to the user and provided many services. The cloud services include Infrastructure as a service (IaaS), Platform as a service (PaaS), Software as a service (SaaS) etc. Resources are available to the user through IaaS on demand basis (Buyya, 2009). These resources are available to the users as virtual machines (VMs)in the heterogeneous cloud domain (Nathani, Chaudhary, & Somani, 2012). The mapping of tasks to the

DOI: 10.4018/978-1-7998-5339-8.ch063

VMs is a popular NP Complete problem. The goal is to minimize the overall processing time which is known as makespan (Zhang, Wang, Burgos, & Boroyevich, 2009). Heterogeneous cloud environment consists of machines having different properties. The computing power of all the machines also varies according to the properties. There are many techniques and application of Datamining are there in the area of cloud computing. Information mining in distributed computing is the method of obtaining structured and valid data from unstructured or semi-structured resources. Datamining provides methods to generate a smoothed value from a large volume of data. Smoothing technique include, smoothing by bin method which smooth a sorted data (Maheswaran, Ali, Siegal, Hensgen, & Freund, n.d.). Sorted values are distributed a number of buckets or bins. After smoothing, scheduling techniques can be applied to the smoothed data to get a better makespan and average cloud utilization. In the past decades many new scheduling heuristics has been proposed with various evaluating criteria (Han, & Kamber, 2001; Li et al., 2012; Ming, & Li, 2012). These scheduling techniques are either in homogeneous environment or heterogeneous environments. Researchers have proposed many tasks scheduling algorithms based on many criteria. Tasks scheduling in cloud can be broadly classified in to two types. One is batch mode and other one is online mode (Han, & Kamber, 2001). In the batch mode of scheduling, tasks are grouped in to batches. The batches are processed with an appropriate interval of time. Scheduling techniques like First Come First Service (FCFS), Round-Robin (RR), Max-Min and Min-Min are examples of batch mode scheduling. In FCFS, task which comes early is served. This algorithm consider the arrival time of a task. It is a simple and fast (Chen, Wang, Helian, & Akanmu, 2013; Braun et al., 2001; Cao, Spooner, Jarvis, & Nudd, 2005). In RR, tasks are served in a FIFO manner. A limited time sliced or quantum is given to each task. If a task does not complete before CPU time expires, the CPU is given to another task waiting in a queue. The preempted task is placed in the ready list and served later (Izakian, Abraham, & Snasel, 2009; Mantri, 2011). In Min-Min scheduling algorithm in cloud environment, a task is chosen from a bag of unmapped tasks and assign to the machine where the respective task has minimum completion time. This algorithm takes all the tasks at a time and gives an effective makespan (Nagadevil, Satyapriya, & Malathy, 2013; Freund et al., n.d.). Max-Min algorithm is close to Min-Min algorithm. The tasks having maximum completion time is selected. The task is mapped to the resource that has minimum completion time for that task. The ready time of the machines are reformed consequently. Until the unmapped tasks are assigned, the tasks allocation to the machine is repeated. Here the aim is to reduce the wait time of large jobs (Ibarra, & Kim, 1977). In Minimum Completion Time (MCT), tasks or jobs are assigned to the resources based on their completion time. Completion time is the sum of ready time and expected execution time of a particular task on particular machine. Task with minimum completion time is chosen (Mantri, 2011). Scheduling algorithm based on data-mining techniques are also consider the criteria such as makespan and machine utilization (Potluri, & Rao, 2017). A scheduling algorithm based on quality of services results a good makespan. In that method factors like makespan, QoS, accepted rate, cost, fairness completion time and minimum completion time are taken (Braun et al., 2001; Panda, Pradhan, Neha, & Sathua, n.d.). Tasks are arranged in a fair ratio using certain algorithm (Pradhan, Panda, & Sathua, 2015), which results a good makespan and resource utilization. In Taguchi Orthogonal approach based scheduling algorithm, was incorporated at CSO tracing mode for task mapping on virtual machines. It is implemented using cloudsim . A task scheduling algorithm is implemented for conservation of energy using broadcasting. It is basically describes about the cloud server maintenance and scheduling of processes (Gabi, Ismail, Zainal, & Zakaria, 2017; Sultanpure, Gupta, &, Reddy, 2018). The reminder of this paper describe as follows. We describe the smoothing technique in datamining in the Section 2. We discussed a cloud model and features of the task scheduling

Related Content

The Cloud in Education: Policy, Leadership, and Management Issues
Karl Donert (2021). *Research Anthology on Architectures, Frameworks, and Integration Strategies for Distributed and Cloud Computing (pp. 2371-2393).*
www.irma-international.org/chapter/the-cloud-in-education/275395

“Saksham Model” Performance Improvisation Using Node Capability Evaluation in Apache Hadoop
Ankit Shahand Mamta C. Padole (2021). *Research Anthology on Architectures, Frameworks, and Integration Strategies for Distributed and Cloud Computing (pp. 1282-1302).*
www.irma-international.org/chapter/saksham-model-performance-improvisation-using-node-capability-evaluation-in-apache-hadoop/275339

Domain Knowledge Embedding Regularization Neural Networks for Workload Prediction and Analysis in Cloud Computing
Lei Li, Min Feng, Lianwen Jin, Shenjin Chen, Lihong Maand Jiakai Gao (2021). *Research Anthology on Architectures, Frameworks, and Integration Strategies for Distributed and Cloud Computing (pp. 1158-1176).*
www.irma-international.org/chapter/domain-knowledge-embedding-regularization-neural-networks-for-workload-prediction-and-analysis-in-cloud-computing/275332

Fog Computing Qos Review and Open Challenges
R. Babu, K. Jayashreeand R. Abirami (2021). *Research Anthology on Architectures, Frameworks, and Integration Strategies for Distributed and Cloud Computing (pp. 1147-1157).*
www.irma-international.org/chapter/fog-computing-qos-review-and-open-challenges/275331

Novel Taxonomy to Select Fog Products and Challenges Faced in Fog Environments
Akashdeep Bhardwaj (2021). *Research Anthology on Architectures, Frameworks, and Integration Strategies for Distributed and Cloud Computing (pp. 2558-2571).*
www.irma-international.org/chapter/novel-taxonomy-to-select-fog-products-and-challenges-faced-in-fog-environments/275404