

Chapter 60

A High Performance Model for Task Allocation in Distributed Computing System Using K-Means Clustering Technique

Harendra Kumar

Department of Mathematics and Statistics, Gurukula Kangri University, Haridwar, India

Nutan Kumari Chauhan

Department of Mathematics and Statistics, Gurukula Kangri University, Haridwar, India

Pradeep Kumar Yadav

Department of Research Planning and Business Development, Central Building Research Institute, Roorkee, India

ABSTRACT

Tasks allocation is an important step for obtaining high performance in distributed computing system (DCS). This article attempts to develop a mathematical model for allocating the tasks to the processors in order to achieve optimal cost and optimal reliability of the system. The proposed model has been divided into two stages. Stage-I, makes the 'n' clusters of set of 'm' tasks by using k-means clustering technique. To use the k-means clustering techniques, the inter-task communication costs have been modified in such a way that highly communicated tasks are clustered together to minimize the communication costs between tasks. Stage-II, allocates the 'n' clusters of tasks onto 'n' processors to minimize the system cost. To design the mathematical model, executions costs and inter tasks communication costs have been taken in the form of matrices. To test the performance of the proposed model, many examples are considered from different research papers and results of examples have compared with some existing models.

DOI: 10.4018/978-1-7998-5339-8.ch060

1. INTRODUCTION

Distributed computing system (DCS) consists of multiple processors, which are distributed at different sites and interconnected by a communication network. Task allocation is very important step in DCS. The purpose of task allocation is to maintain maximum utilization of computing resources. Task allocation approaches can be classified into two categories: statistical and dynamical. Statistical approaches allow a stochastic relaxation that reduces the cost function in search of a global minimum to take the task allocation decision. The statistical approaches get near to optimal solutions. In statistical approaches, when a task is assigned to the processor, it remains there until the execution of program. Due to this reason, it required more convergence times. While dynamical approaches are employed more frequently because they offer a better efficiency, having the flexibility of tasks migration.

The performance of the system depends upon several components such as load distribution, system reliability and system cost. Load distribution improves the performance of the system by transferring tasks from heavily loaded processors to lightly loaded processors. If workloads at some processors are typically heavier than at others, or if some processors execute tasks more slowly than others, the situation of lightly loaded or heavily loaded processors are likely to occur often. It is shown that even in such D; at least one processor is likely to be idle while other processors are heavily loaded due to the statistical fluctuations in the arrival of tasks computing system to the system and task service time requirements. Therefore, even in a homogeneous distributed, system performance can be potentially improved by the appropriate transfer of the workload from heavily loaded processors to lightly loaded processors. System reliability also plays an important role to improve the performance of the system. The system reliability can be defined as that all involved processors and communication links are operational during the execution of program.

In this manuscript, we have used the clustering technique to reduce the inter-task communication costs. Clustering is the act of grouping a set of objects in such a way that objects in the same groups (clusters) are more similar (in some senses) to each other to those in other group (cluster). The result of a cluster analysis can contribute directly to development of classification schemes. Indeed, clustering is used to reduce a very large data to a relatively compact analysis. Clustering was originated by Driver and Kroeber in 1932 and introduced to psychology by Zubin in 1938. Clustering is used in many fields such as machine learning, pattern recognition, image analysis, information retrieval, bioinformatics, data compression and computer graphics.

2. RELATED RESEARCH

Task allocation problem attracted the attention of many researchers due to its importance in DCS. Researchers have done work on the problem of task allocation according to their interest. Several researchers have been studied statistical task allocations. Sarje and Sagar (1991) proposed a heuristic model task allocation in DCS. Lee et al. (1992) discussed the problem of assigning tasks to the processors of DCS such that the sum of execution and communication costs is minimized. A task allocation model has been discussed by Rewini and Ali (1995) in multiprocessing system. Peng et al. (1997) present an optimal solution to the problem of allocating communicating periodic tasks to heterogeneous processing node in a distributed real time system. Elsedek and Well (1999) have developed a heuristic model for task allocation in heterogeneous DCS. Topcuoglu et al. (2002) discussed a critical path on a processor

23 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/a-high-performance-model-for-task-allocation-in-distributed-computing-system-using-k-means-clustering-technique/275337

Related Content

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

Detection of Worms Over Cloud Environment: A Literature Survey

Thangavel M., Jeyapriya B. and Suriya K. S. (2021). *Research Anthology on Architectures, Frameworks, and Integration Strategies for Distributed and Cloud Computing* (pp. 2472-2495).

www.irma-international.org/chapter/detection-of-worms-over-cloud-environment/275400

Access Control Framework Using Multi-Factor Authentication in Cloud Computing

Subhash Chandra Patel, Sumit Jaiswal, Ravi Shankar Singhand Jyoti Chauhan (2021). *Research Anthology on Architectures, Frameworks, and Integration Strategies for Distributed and Cloud Computing* (pp. 1130-1146).

www.irma-international.org/chapter/access-control-framework-using-multi-factor-authentication-in-cloud-computing/275330

Study on Secure Dynamic Covering Algorithm for E-Logistics Information in a Cloud Computing Platform

Yan He (2021). *Research Anthology on Architectures, Frameworks, and Integration Strategies for Distributed and Cloud Computing* (pp. 1097-1111).

www.irma-international.org/chapter/study-on-secure-dynamic-covering-algorithm-for-e-logistics-information-in-a-cloud-computing-platform/275328

Cloud Computing, Green Computing, and Green ICT

Kijpokin Kasemsap (2021). *Research Anthology on Architectures, Frameworks, and Integration Strategies for Distributed and Cloud Computing* (pp. 2083-2099).

www.irma-international.org/chapter/cloud-computing-green-computing-and-green-ict/275380