INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATIONS AND ROBOTICS
ISSN 2320-7345

A SURVEY ON JOB SCHEDULING MECHANISMS IN PEER TO PEER COMPUTING SYSTEMS

Ashwathy M.G

Post-Graduate Student, Department of Computer Science and Engineering, Karunya University, India.

Abstract

Peer to peer computing is a kind of computing which uses the idle cycle of large number of computers that are connected through internet and their resources are utilized to solve large tasks. It uses the idle cycles of all the computers connected to the network for solving huge parallel applications. An efficient scheduling mechanism is the most important challenge taken by the peer to peer network for utilizing the resources properly and also to assign the tasks to the resources effectively. The effective mechanism is needed for decent resource sharing, get adapted to changing dynamic nature, preventing the various internet attacks from participating peers, and coordination of the different policies. Cooperation is appreciated because users in P2P networks use maximum resources they get but offer only less number of resources in response. This study provides a comparative analysis on various job scheduling mechanisms in peer to peer systems that improve the poor resource utilization.

.

Keywords: Grid computing; P2P computing; Grid scheduling

I. Introduction

Grid computing consists of geographically distributed heterogeneous resources. Grid system has a dynamic nature in which resources may join and leave the system at any time. Because of this dynamic nature of Grid the system face the difficulty to schedule resources which results in poor resource utilization and its management. The scheduling is the process of allocating jobs to available resources. In order to perform proper scheduling of jobs to the resources several techniques has been developed but still it is impossible to achieve proper resource utilization. When resources are allocated to jobs, resources with good performance may get assigned to jobs thus overloading the resource and reducing the proper utilization of resources of the system.

P2P computing systems uses the CPU idle cycles of computers connected through Internet. They use the resources to execute large parallel distributed applications. Such systems require an efficient scheduling mechanism so that the tasks can be assigned to the heterogeneous computing resources. However, since scheduling mechanisms in decentralized environments require long processing times, the centralized scheduling methods seems to be inefficient. This is because they may cause bottleneck and a single point of failure.

The development of incentive techniques is necessary to bring cooperation and resource sharing among peers to improve system performance. Cooperation is appreciated because users in P2P networks use maximum resources they get but offer only less number of resources. A two-level scheduler mechanism is used for distributing parallel applications in P2P computing systems in an efficient way.

This paper is organized as follows. In Section II, we do the discussion on the key concepts in this paper. In Section III, we present the comparative study. In Section IV, we perform the parametric comparison of various scheduling mechanism. We conclude in Section V.

II. Key Concepts

A. Grid computing

Grid computing is a term referring to combining various resources together which is widely dispersed for executing large tasks. It provides mechanisms that help to use and access large number of heterogeneous resources. A grid system is said to be dynamic because the resources and tasks changes rapidly. The grid is heterogeneous because various kinds of resources and tasks which are widely distributed are gathered together from different sources to solve a common goal. The grid system provides the user the capabilities with large scale, i.e. the system should be capable to collect large number of resources and tasks and provide the users a secure way to access them. The grids are more loosely coupled, heterogeneous, and geographically dispersed which distinguishes the grid computing from conventional computing systems. Grid computing involves computation in a distributed fashion, which may also involve the collection of large number of areas.

B. Peer-to-Peer computing

Peer-to-Peer (P2P) computing is widely used as a distributed paradigm for developing applications, ranging from large scale scientific applications to small ad hoc information sharing, by combining the resources from various sources to solve tasks. However, the heterogeneous and dynamic nature of the peers across various domains brings the challenge for resource sharing in the grid environment. However, resource management in these environments is a complex one. These systems need effective mechanism for decent resource sharing, get adapted to changing dynamic nature, preventing from various internet attacks from participating peers, and coordination of the different policies, different cost models and varying loads in different peers. P2P computing systems uses the CPU idle cycles of computers connected through Internet. They use the resources to execute large parallel distributed applications. Such systems require an efficient scheduling mechanism so that the tasks can be assigned to the heterogeneous computing resources.

C. Grid scheduling

Grid scheduling is defined as the process of making scheduling decisions involving resources over multiple domains. Local scheduling is used at the intra-level, usually to balance load. Global schedulers are used to map the tasks to the resources. This process includes scheduling the user jobs to multiple resources at a single area or multiple areas. In general we can differentiate between a Local scheduler and global scheduler, that is, a scheduler that is responsible for scheduling and managing resources at a single area, or multiple areas. Grid computing is becoming a generic platform for high performance and distributed computing due to the variety of services it offers such as computation services, application services, data services, information services, and knowledge services. These services are provided by the processing elements in the grid computing system. The servers and the processing elements are typically heterogeneous in nature.

III. Comparative Analysis among various scheduling mechanisms

This section includes a study on various scheduling mechanisms in improving scalability and the system performance of the peer to peer systems.

A. Self-Policing mechanism

A self-policing and distributed approach [3] was developed, to make the peer nodes profitable under untrusted environment. The mechanism is the combination of two models: PET, and M-CUBE, that form the foundation for resource sharing in an untrusted P2P computing environment. The mechanism is flexible which gets adapted to various kinds of requirements. The uniqueness of this approach is in its ability to integrate the trustworthiness and dependability of peers for resource trading. The self-policing mechanism helps providing cooperation among the resources of peer to peer network. The mechanism controls the flow of free riders in the system and provides an efficient resource sharing among the peers. This helps the system to be more secure and reliable. The resource sharing sometimes becomes inefficient because of the selfish peers. The cooperation among the peers of different domains is considered more challenging.

B. Balanced multiway tree based mechanism

An efficient resource discovery mechanism [1] based on balanced multiway tree structure, was developed to support all kinds of exact and range queries efficiently. This approach uses a structured architecture. The peers in this mechanism are based on a tree topology. The rebalancing algorithm provides an efficient way to maintain the tree in a balanced manner and also relink the isolated node if any. The mechanism follows effective searching techniques and is highly robust in nature. The unsuccessful search rate is improved because of the good performance of the rebalancing mechanism. The response time is also acceptable when compared to others. In this mechanism the congestion can take place due to updating algorithm and so it is not so scalable in large scale network. The increase in communication time and message overhead are other demerits of this mechanism.

C. Resource bundle based mechanism

A resource bundle based mechanism [4], was developed to overcome migration overhead resulted by poor deployment decisions. The decisions were made by selecting nodes based on observed resource capacities. The resource information of every node is collected and maintained under schedulers. In this mechanism all the information is collected together as a group. The resource bundle information is used to schedule resources for the task. The scheduling scheme distributes the load to all the bundles. The scheme use two models, which are statistical model and cluster based model. In statistical model, the information is grouped up using aggregate functions whereas in cluster model, the grouping is done based on similarity values. It is suitable for large scale systems. The resource bundles are used to classify and locate the group level features and characteristics such as the group capacity. The resources of every node are monitored regularly and it is done in simplified manner. The information collected by the scheduler is updated frequently. This result in resource update overhead in the grid environment. The scheduling mechanism is not so cost effective and also there are not enough security features.

D. Coordinated load management in peer to peer systems

The Coordinated load management in peer to peer grid systems [5], was developed for resource allocation and decentralized coordination. Resource brokering services control the whole management of the applications in the distributed Grid computing environments. The coordinated load management mechanism uses a Distributed Hash Table for efficient management of the coordination objects. Specifically, in this mechanism the resource brokers and resources are organized based on a DHT overlay. In this approach, resource brokers post their resource needs by using a Resource Claim object into the coordination space, while resource providers notifies their resource by using a Resource Ticket object. These objects are mapped to coordination services using a hashing technique.

E. An Exchange-based incentive mechanisms

The performance of peer to peer resource sharing networks depends on the cooperation of participants. The exchange based incentive mechanism [2], provides incentive for cooperation in peer to peer system. The approach provides strong incentives for file sharing, offers increase in service time to users when compared to free riders, without the complexity of cash based and credit based systems. The basic idea is that peers give high priority to the peers that can provide a service in return. The mechanism is suitable for systems with more load or overloaded systems. It provides strong incentives for users to share resources.

F. Credit based incentive mechanisms

The incentive and scheduling mechanism [7], was developed based on credits with a two level topology. The low level implements a non-negative credit function. The reinvestment of credits increases peer cooperation. The upper level is introduced to improve scalability and scheduling is achieved irrespective of number of peers with improved scalability. The global incentives mechanism can be improved if it considers the replicated and multi task execution on each peer. The mechanism limits the free riders into the system and improves the system throughput

.

G. Two level scheduling mechanisms

The two level scheduling mechanisms [6], was developed for distributed and parallel applications in peer to peer environment. The low level scheduling is performed by each manager at the individual area. The information required about the resources is assumed to be known by the scheduler. The low level scheduler uses scheduling policy to perform near optimal mapping of tasks. Due to the scalability issues the high level scheduling is introduced. The upper level requires more scheduling information. To reduce the information required in this level of scheduling, scheduling criteria is used to schedule task among the areas. The distributed scheduler is used for assigning task among areas. The tasks are distributed in a balanced and efficient way. The scheduling cost is also reduced and managed by the use of the mechanism. The cooperative scheduling mechanism is designed to work on huge, distributed and heterogeneous environment.

IV. Comparison

The comparison of various scheduling mechanisms is done using different parameters and is shown in Table 1.

Table 1: Parametric comparison

Papers/ Parameters	Distributed	Cooperative	Scope	Message Overhead	Scalability
Self-Policing mechanism	Yes	No	Large	Yes	Medium
Balanced multiway tree based mechanism	No	No	Small	Yes	Low
Resource bundle based mechanism	No	No	Small	Yes	Low
Coordinated load management in peer to peer systems	No	Medium	Large	Yes	Low
Exchange-based incentive mechanisms	Yes	No	Large	No	Low
Credit based incentive mechanisms	Yes	Medium	Large	Yes	Medium
Two level scheduling mechanisms	Yes	Yes	Large	No	High

V. Conclusion

The study on various job scheduling mechanisms has been performed. Their merits and demerits have also been discussed. The comparative study has been performed based on various parameters. The goal of this study is to provide comparison on different mechanisms and insight into the characteristics of each mechanism. The study will help to develop more capable and competent scheduling mechanisms. The future work can be done by giving importance on eliminating the demerits of mechanisms so that it can be made more efficient in dynamic and heterogeneous environment.

REFERENCES

[1] D. Castella, H. Blanco, F. Gine, F. Solsona,(2010),"A Computing Resource Discovery Mechanism Over a P2P Tree Topology", In proceeding of: High Performance Computing for Computational Science - VECPAR'10 - 9th International conference, Berkeley, CA, USA, pages 366-379

[2] Kostas G. Anagnostakis, Michael B. Greenwald,(2004),"Exchange-based Incentive Mechanisms for Peer-to-Peer File Sharing", In: Proceedings of ICDCS 04- International Conference on Distributed Computing Systems, pages 524 - 533

[3] Liang Z and Shi W (2005),"Enforce Cooperative Resource Sharing in Untrusted Peer-to-Peer Computing Environment", Journal of Mobile Networks and Applications, pages 971-983

[4] M. Brinda Kumar, Dr. K.P. Kaliyamurthie,(2013),"A Novel Resource Distributed Discovery and Management in Grid Computing", International Journal of Computer Science and Mobile Computing, pages 623 – 627

[5] Rajiv Ranjan, Aaron Harwood, Rajkumar Buyya,(2012),"Coordinated load management in Peer-to-Peer coupled federated grid systems", J Supercomput, pages 292–316

[6] Rius J, Cores F, Solsona F,(2013),"Cooperative scheduling mechanism for large-scale peer-to-peer computing systems", Journal of Network and Computer Applications

[7] Rius J, Estrada S, Cores F, Solsona F,(2012),"Incentive mechanism for scheduling jobs in a peer-to-peer computing system", Simulation Modeling Practice and Theory 25, pages 36–55