High Performance Hypercube for Better Resource Discovery and Allocation in Simulated Grid based Environment

M.Shanthi
Velammal Engineering College
 Chennai-66

A. Anthony Irudhiyaraj, Ph.D
AVIT,
Kachipuram

ABSTRACT

Grid is basically comes under the distributed computing where number of nodes together to share resources, data, services, computational power to solve the problem. The main objective of our research work is to identify the factors of the grid resource discovery and allocations strategies like resource availability, network bandwidth, pending users, thread capability and storage capability. Based on the factors the weight age of the nodes will be evaluated. The main aim of this research to find out the resource availability in hypercube based clusters and allocation of request to the appropriate alive node using hypercube algorithm and then compare with least cost method.

General Terms

Distributed computing- Grid Computing

Keywords

Allocation, Cluster, ,Least cost method, Resource Discovery,

1. INTRODUCTION

 Grids have become an appealing environment for developing large scale parallel applications. The Main advantage of Grid computing system is users can empower huge computational power that's not feasible from a standalone machine. Resource heterogeneity, Supporting adaptability and scalability holds major impact. These areas provide us a vision for our future research directions. In a grid computing environment, resource location and allocation are always a very crucial problem, which impacts the effectiveness of this new technical jargon. To improve the efficiency of grid resource management numerous works have been done by researches in this area[2,3,5,6.14,15,20,21].
Resource discovery is one of the basic and key aspects in Grid Computing to satisfy the requirement of user applications. Resource is any real or conceptual object that can be identified. According to the author in paper [10] resources are computers, cluster of computers, online instruments, storage space, data, application and a resource discovery mechanism returns the identity of matching resources for a given description of desired resources. Discovery involves the finding and retrieving of resources that are relevant to the user of the system. The fundamental aspects of Resource Discovery include: Global scope, distributed in nature, non-structured, no single protocol standard and scalability. Clustering is the problem of discovering natural groups in data sets by partitioning given data points into a number of groups, such that points within a group are more similar to each other than to points from different group. The aim of this clustering is to reduce the inter-job communication and thus, decrease the time required for parallel execution [11].

In the grid construction process the resources may be arranged in NW (Newman-Watts) model [13] that is Manhattan distance between points (x_1, y_1) and (x_2, y_2) is the distance that must be travelled while moving only horizontally or vertically, namely $|x_1 - x_2| + |y_1 - y_{21}|$.

The Watts-Strogatz[19] model is a random graph generation model that produces graphs with small-world properties, including short average path length and high clustering. Our research work focus on hypercube interconnection networks. Among bus, ring, star, tree and hypercube interconnections, hypercube structure is chosen to locate the resources because it does not require the information of all the nodes in the network it is enough to search only the neighboring connected. The distance between two nodes is the length of the shortest path connecting the nodes and is given by the following equation[4].

$$D(x,y) = \sum_{K=0}^{d-1} |x_k - y_k|$$

Where x and y are nodes and k is dimension value of them. For example if the node x=(x0x1x2) and node y=(y0y1y2) then distance between the nodes can be calculated as d(x,y) =x_0-y_0+ x_1-y_1+ x_2-y_2,which is nothing but the number of places at which the binary representation of the nodes x and y differ. Both x and y are neighbors' only when d (x,y) =0.

The Flooding ,Swamping ,Random walk, Random pointer jump , Name Dropper Algorithm and Ant Colony search algorithm are some existing resource discovery algorithms. The challenges based in the existing algorithms are

- Supporting adaptability and scalability was difficult
- Heterogeneity
- Meeting computational cost
- Single point of failure
- Security

The remainder of this paper is organized as follows. Related work in grid resource discovery and allocation are presented in Section II. The architecture and construction of hypercube and algorithm of grid based on small-world cluster is provided in Section III. Section IV presents small-world resource discovery algorithm in detail. Section V gives the experiments and performance analysis. Conclusions and future work are described in Section VI.
.

2. RELATED WORK

Existing methodologies to resource discovery and allocation in a Grid environment are non coordinated. Generally resource discovery methods are classified into three main categories in the literature. They are centralized and hierarchical systems, peer to peer (P2P) based systems, and agent based systems. Alchemi , Gridsim, Globus Toolkit, Unicore and gLite are the most popular examples of such systems. Globus provides many grid services such as resource discovery, allocation and monitoring by extending standard service interfaces. Schedulers such as Nimrod- Grid bus broker, and Condor-G perform scheduling related activities independent of the other schedulers in the system. In this paper, they discussed performance issues of stretched hypercube graphs by introducing a reasonably accurate mathematical model to predict the average message latency in wormhole stretched hypercube using a high performance routing algorithm.[1].Anju Sharma and Seema Bawa [5] analyzed various resource discovery algorithm like Peer-to-Peer approach, Request forwarding approach and compared performance factors like scalability, reliability and adaptability.In research paper [8] the author proposed three steps for resource discovery, which includes of Global Resource Record Table, Grid Management Service and hierarchy discovery mechanism to offer higher hit rate and lower overhead. During the process of execution, it may face the situation of resource provider crash. They also proposed a checkpoint mechanism by monitor the state of executing tasks for fault tolerances. Adriana Iamnitchi and Ian Foster[9] proposed the set of request forwarding algorithms in a fully decentralized architecture, designed to accommodate heterogeneity and dynamism.

In paper[12] they mainly focus on allocation of jobs and resources that allow scheduling across wide area distributed clusters the algorithm considered for their research is cost-time optimization scheduling. The simulator was implemented using GridSim toolkit that allows modeling and simulation of distributed system entities for evaluation of scheduling algorithms. They analyzed how the resource supply and demand pattern affects the systems scalability/performance in terms of total message complexity. The system proposed in paper[16] is a fully decentralized economic allocation scheme, which aims to achieve a high degree of scalability and reliability, and easily allows resources to retain their autonomy. In paper[18] they proposed a new Grid resource allocation method (Modified Least Cost Method) that assigns tasks to computing nodes in a way that is close to the optimum. The proposed method is an advanced form of the Least Cost Method. This paper also includes a comparative performance analysis of their proposed Resource Allocation method with the existing ones. Their aim is to find the portion of each job that is allocated to each processor, so as to minimize the overall computational cost.

This paper[21] presents a new formulation of the resource selection problem and a solution to the resource selection and binding problem called integrated selection and binding. The authors evaluated the integrated selection and binding approach in terms of success rate, selection quality and cost, and binding quality and cost. Several researchers have recently proposed analytical models of popular interconnection networks, In this paper, they discussed performance issues of stretched hypercube graphs by introducing a reasonably accurate mathematical model to predict the average message latency in wormhole stretched hypercubes using a high performance routing algorithm.[1]

3.GRID BASED ON SMALL-WORLD CLUSTER

In computing grid environment, resource discovery efficiency, resource allocation and fault tolerance are challenging problems when developing grid applications. One of the key motivations for constructing Grids is to provide application-level connectivity among the various machines so that resources and services supported by the individual systems can be shared in a Global fashion. Parameters such as CPU speed and memory size may accompany with the task submission to the agent. This paper introduces small-world cluster into grid in which cluster nodes are arranged in hypercube structure to form small-world network. The small world phenomenon was first observed by Milgram[13], who discovered the interesting "six degrees of separation" in a social network. Although the notion of small world phenomenon originates from social science research, it has been observed that the small world phenomenon is pervasive in a wide range of settings such as social communities, biological environments, and data communication networks. a small world network can be viewed as a connected graph characterized by low characteristic path length(i.e., similar to the average path length in random networks) and high clustering coefficient (i.e., much greater than that of random networks). Previous researches mostly focus on understanding the navigability of small-world networks by analyzing the diameter and the routing efficiency.

The characteristic path length, L, is the path length averaged over all pairs of nodes. The path length d(i, j) is the number of edges in the shortest path between nodes i and j. The clustering coefficient a measure of the cliqueness of the local neighborhoods. For a node with k neighbors, then at most $kC_2 = k(k − 1)/2$ edges can exist between them. The clustering of a node is the fraction of these allowable edges that occur. The clustering coefficient, C, is the average clustering over all the nodes in the graph.(19). For hypercube the clustering coefficient is $n/2^{n-1}(2^n-1)/2$.

We assume that requests are produced by the application periodically and are mapped and allocated to the node. The amount of request processed by these applications poses a great challenges on the grid structure A well known strategy efficient execution of huge application on a heterogeneous grid is to partition the grid in to clusters based on small world theory concept. Small-world network has similar clustering coefficient to regular network and low average path length like random network. Any search can be performed in $O(\log 2\ n)$ steps in small-world networks, where *n* is the number of nodes. In the process of constructing small-world network, a new construction method on the basis of Hypercube interconnection model is presented, that is, Two nodes in the hypercube are adjacent if and only is binary representation differs in exactly one bit. Here, we have designed grid an application in which all the nodes share census database and provide allocation services on thread basis to all the requests from censes application.

4.PROPOSED SYSTEM

4.1 PROPOSED PROBLEM

The main aim of this research to find out the resource availability in hypercube based clusters and allocation of request to the appropriate alive node using hypercube algorithm and then compare with Least cost method. Existing system takes care of the cost. But our proposed algorithm takes care of cost, priority and resource intensity

.Application is designed in such a way that the request is coming from censes application which is retrieved from SQL database. Threading concept is used for handling the request efficiently. In the clusters, nodes are arranged in hypercube format based on the concept of small world theory.

4.2 PROPOSED ARCHITECTURE
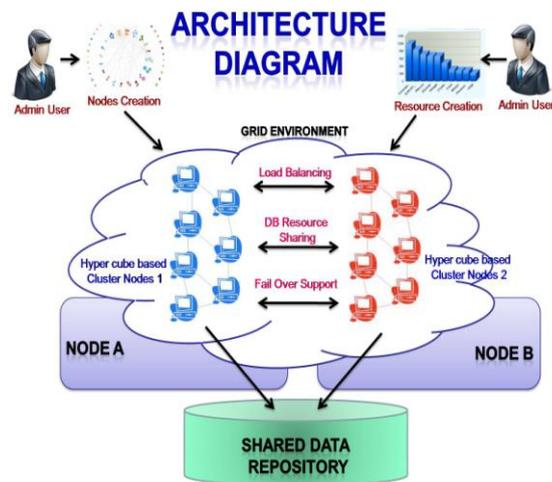
Fig 1: Proposed Architecture Diagram

The admin user create nodes based on the user request and send request to the Grid scheduler. The scope of the grid scheduler includes resource discovery and dynamically allocating the request to free node. In case any node failure occurs grid scheduler will find alternate working node and task is routed to corresponding working node. For example if 10 requests are coming inside, the admin user forward all the requests to grid scheduler. The scheduler created 2^4 nodes and requests are allocated dynamically based on the load of the each node..

4.2 PROPOSED ALGORITHM

Step 1: From each node listen for the incoming request or message.
Step 2: Search through routing table for a path
Step 3: If connection is direct
       If node is active{ set timeout limit }
WeightageR = function(Rdelta, Compute Resource Intensiveness)
WeightageU = function(Udelta, Compute Total User Intensiveness)
WeightagePU=function(Pudelta,ComputePending User Intensiveness)
WeightageN=function(Ndelta,ComputeNetworkCapacity i intensiveness)
WeightageM = function(Mdelta, Compute Memory Intensiveness)
Considering the above factors, The final weight of a computer node is
    Weffective = WeightageR + (WeightageU - WeightagePU) +
                  WeightageN + WeightageM
   If resource is not available go to step 2
     else node is not active{ set timeout limit }
    set flag search for the alternative path (Back track
   to the nearest node and go to step 2)
    Continue this procedure till the resource is find.
Step 4.
Stop.

An algorithm is designed using .net to handle the request efficiently in hypercube based clusters.
The algorithm consists of four steps
- Hyper cube creation and arrangement

- Resource allocation
- Evaluation of Threading performance
- Overall performance
- Comparison with least cost method

Based on the given dimension Hypercube is generated For example if 8 is chosen 28 = 256 node hypercube is generated. All generated nodes are arranged in clusters. Each cluster contains 64 nodes .Each node is capable of handling requests. These nodes arranged using small world theory.

First step is to determine the capacity of each node in the hypercube based clusters. For example , if 6 is selected is responsible to handle maximum of six . In Case any failure occurs , failover option can be invoked, so that the entire remaining process is re rooted to another cluster, without affecting processor time. Requests are loaded from censes application. When the requests hit the node in the hypercube the process is initiated. By giving different values that is 3,4,5,6&7 we can evaluate thread performance. Now the input from censes application is stopped and request from senses application cost is initiated performance is evaluated by running hypercube cost program. The parameters measured between the two algorithm are Memory, Processor time While handling the request memory ,I/O deadlock latency time of each request is recorded .
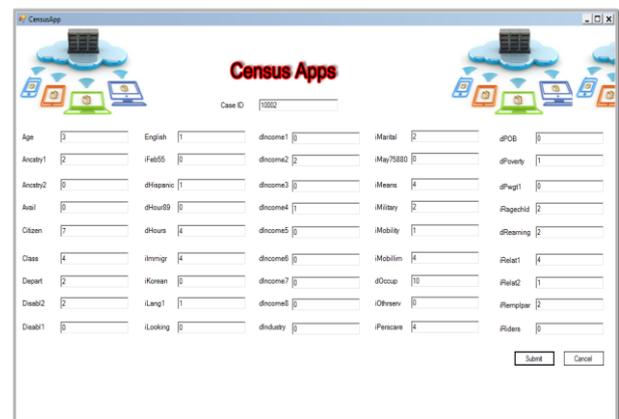

Fig 2.Censes application
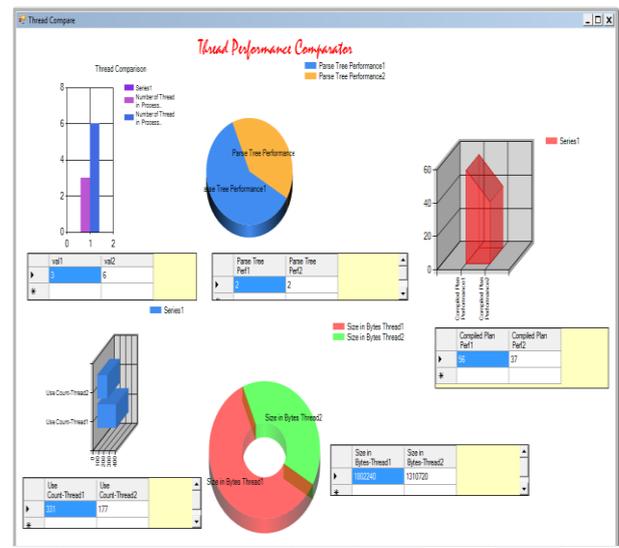
Fig3.Thread performance comparison
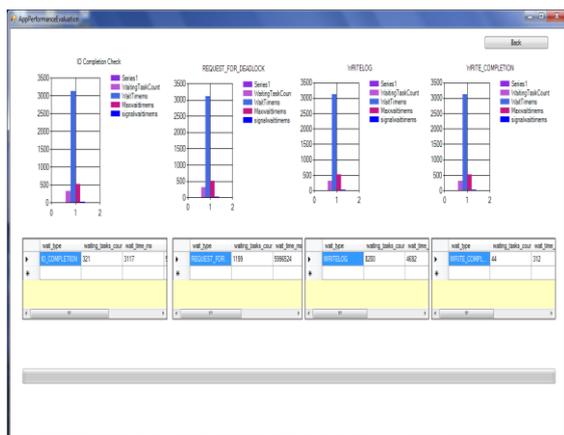
**Fig4.Overall performance**
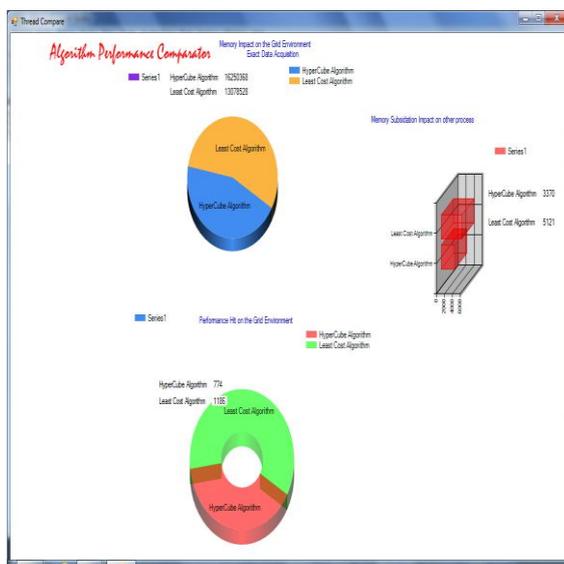
Fig5.Comparison with Leastcost method

5. CONCLUSION

One of the primary challenges for grid computing is the ability to discover the resource on time, tolerate failure and recover from them. A Resource discovery should be decentralized to avoid bottlenecks and guarantee scalability and adaptability An application is designed based on small world network to handle continuous request from census application. Threaded approach is used to increase the performance of the discovery and allocation process. A simulated grid based system would be adopted on a large scale by commercial or research applications only if it proves to be efficient, in terms of cost, time and effort. Our approach is best in terms of cost, memory and processing time.

6.REFERENCES

[1] Abdel Aziz Farrag and Shituo Lou, Designing and Reconfiguring Fault-Tolerant Hypercubes .

[2].Hameurlain, D. Cokuslu, and K. Erciyes, "Resource discovery in grid systems: a survey," International Journal of Metadata, semantics and Ontologies, vol. 5, , pp. 251-63. 2010.

[3].Alan Bradley,Kevin Curran, Gerard Parr, Discovering Resources in Computational GRID Environments. The Journal of Super Computing Vol 35 Issue 1 ,January 2006.

[4] Alan Parker Algorithms and Data Structures in C++

[5] Anju Sharma and Seema Bawa , Comparative Analysis of Resource Discovery. Approaches in Grid Computing Computer Science & Engineering, Journal of Computers, Vol. 3, No. 5,pp.60-64. 2008.

[6] Anand Padmanabhan, Shaowen Wang. Sukumar Ghosh, and Ransom Briggs.A Self-Organized Grouping (SOG) Method for Efficient Grid Resource Discovery.

[7] Bucur, A.I.D., Epema, D.H.J., Schedulling Policies for Processor Co allocation in Multicluster Systems, IEEE TPDS, 18(7),958- 972, 2007

[8] Chun-Fu Lin and Ruay-Shiung Chang A Resource Discovery and Allocation Mechanism in Large Computational Grids for Media Applications.

[9] Iamnitchi, A., Ripeanu, M., Foster, I.: Small-world file-sharing communities. In: Proc. IEEE INFOCOM. (2004)

[10] Iamnitchi A, Foster I, "On Fully Decentralized Resource Discovery in Grid Environments,"Proc. The 2nd IEEE/ACM International Workshop on Grid Computing 2001, Denver.

[11] Chris Allick, André Pinter, Damon Tymon Tyman Grid Computing Communication Strategies for Cross Cluster Job Execution. 2006, pp 33-38

[12] Manavalasundaram V.K, Duraiswamy.K Association Based Grid Resource Allocation Algorithm European Journal of Scientific ResearchISSN 1450-216X Vol.78 No.2 (2012),pp.248-258.

[13] Milgram, S.: The small world problem. Psychology Today(1967) 60–67

[14] Muthucumaru Maheswaran and Klaus Krauter A Parameter-based Approach to Resource Discovery in Grid Computing Systems.

[15] Raman, N.Mohanram and S.Varun, "Semantic based Grid Resource Discovery and its Integration with the Grid Service Broker", ADCOM 2006: Proceedings of 14th International Conference on Advanced Computing & Communications (2006), pp. 84 – 89.

[16]Simon Mark Davy Decentralised Economic Resource Allocation For Computational Grids

[17] Sina Meraji and Hamid Sarbazi-Azad Mathematical Performance Modelling of Stretched Hypercubes.

[18] Syed Nasir Mehmood Shah, Ahmad Kamil Bin Mahmood and Alan OxleyModified. Least Cost Method for Grid Resource Allocation 2010 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery.

[19]Watts, D., Strogatz, S.: Collective dynamics of small-world networks. Nature 393 (1998) 440–442.

[20] Yan Ma,Bin Gong and Lida Zou 'Resource Discovery Algorithm Based on Small orld Cluster in Hierarchical Grid Computing Environment', Seventh International Conference on Grid and Cooperative Computing, 2008. [21] Yang-Suk Kee, Ken Yocum, Andrew A. Chien Improving Grid Resource Allocation via Integrated Selection and Binding

[21] Yang-Suk Kee, Ken Yocum, Andrew A. Chien Improving Grid Resource Allocation via Integrated Selection and Binding