Novel Resource Discovery in Grid Computing using ACO

Mr. Laxman J. Deokate, Mr. V. V. Puri.

***Abstract*— Considerable efforts are going on towards the decentralization of Grid Information Service (GIS). In order to facilitate resource sharing it is important that Grid Resource Management System (GRMS) finds the location of desired service and timely fashion. Current approaches to resource discovery in GIS are based on centralized or hierarchical client/server models. In this model one or more index servers are used as centralized resource discovery. This centralized index server cannot cope up with scalability and fault tolerance. The framework proposed in this paper organizes grid network in two layers. Gateway nodes are responsible for routing and resource discovery. Gateway nodes utilize Ant colony optimization technique to discover resources in Grid Environment. Result shows that proposed algorithm is better than peer to peer resource discovery.**

***Index Terms*— Computing, Resource Discovery, GRMS, Ant Colony Optimization, GIS, peer to peer.**

I. Introduction

Grid computing has emerged as the next-generation parallel and distributed computing methodology. Its goal is to provide a service-oriented infrastructure to enable easy access to and coordinated sharing of geographically distributed resources for solving various kinds of large-scale parallel applications in the wide area network. Nowadays, grid computing has been widely accepted, researched, and given attention to by researchers [1].

Unlike the traditional file exchange, as supported by the Web or peer-to-peer systems, users in the grid can access the required resource or service in a transparent way as if they were to use local resources or services. However, it gives rise to an irreconcilable conflict between grid users and resource providers in usage policy of the local resources. For users, in addition to simplicity and easiness, to get desirable service functionalities, some quality of service (QoS) targets associated with the service, such as grid service reliability [2], the financial cost of the resource, and the efficiency of grid service, may be specified when a service is submitted. On the other hand, resource providers receive the compensation from grid users for the consumed resources at the price of sacrificing local task executions [3]. Meanwhile, resource providers may not participate in the grid unconditionally, and they may specify different policies that govern how the

Manuscript received Dec, 2014.

Mr. LaxmanJ. Deokate is currently pursuing Master degree in Information Technology,Pune University, Maharashtra, India.

Mr.V.V.Puri has completed Master degree in Computer Science and Engineering, Pune University, Maharashtra, India.

II. Releted work

Grid computing is a form of distributed computing whereby same time to solve a single problem. Grid systems are designed for collaborative sharing of resources. It can be thought of as distributed and large scale cluster computing. Grid computing is making big contributions to scientific research, helping scientists around the world to analyze and store massive amounts of data by sharing computing resources. Grid computing is a new model of supercomputing where the jobs are scheduled and distributed in a network of normal personal computers. Rather than configuring and developing a super computer, the system takes the help of the individual computers in the network and performs the jobs. The performance depends upon many factors such as the processing and the memory capability of the nodes, number of jobs already running in the nodes, cost to route a job from one system to another in a grid computing environment and so on. Therefore merely scheduling a job parallel to other computers does not achieve the performance gain desired. Hence we take the help of grid computing system as processing power, disk storage or network bandwidth, directly available to other network participants, without the need for central coordination by servers or stakeholders

A. Towards Problem

Resources comprising the grid are heterogeneous (e.g. computer, data sources, applications etc) and the main goal of the Grid is to efficiently enable users to access and integrate these resources through a single interface. The main subsystem responsible for achieving this is the Grid Resource Management System (GRMS). In order to facilitate resource sharing, it would first be necessary that the GRMS finds the location of desired resources in an efficient and timely fashion. This problem is known as resource discovery. Resource discovery, a basic service of any GRMS, is concerned with the problem of matching a query for resources to a set that meets the query's conditions. These conditions (constraints) describe required characteristics (or attributes) of the resources (e.g., having certain CPU architecture).

B. How the Existing Problem Works

Current approaches to resource discovery on grids are based on centralized and hierarchical client/server models. In such models, there exist one or more index servers which act as a centralized resource discovery directory. These servers maintain index information about available resources, and accept requests from resource consumers. As Grid systems increase in scale, they begin to create issues of self-configuration, fault-tolerance, and scalability. Even though most of the researches apply this concept of resource

ISSN: 2278 – 1323

discovery on p2p network they fail to achieve these three points: i)The ability to perform multiattribute exact match and range queries to discover resources matching a given criteria, ii) The ability to scale to a large no of nodes and organizations scattered over geographically distributed locations. iii) Maintaining administrative control over the local resources and their states

C. Scalable Autonomous Resource Discovery for the Grid(SAPRDG)

SAPRDG (Scalable Autonomous Resource Discovery for the Grid) a framework for scalable resource discovery on the Grid based on structured P2Poverlay networks is proposed in [7]. The framework attempts to solve many of the difficulties encountered in P2P resource discovery while, at the same time, remaining scalable to a large number of distributed nodes. In particular, the contributions of the proposed framework are: (i) the ability to perform multi attribute exact match and range queries to discover resources matching a given criteria, (ii) the ability to scale to a large number of nodes and organizations scattered over geographically distributed locations, and (iii) maintaining administrative control over the local resources and their states. The remainder of this paper is organized as follows. Section 2 give an overview on the most recent efforts in utilizing Structured P2P overlay networks for decentralizing the Grid Information Service (GIS).

By studying current Grid systems based on P2P resource discovery techniques it is clear that utilizing such techniques in general, and structured P2P techniques in particular, for decentralizing the Grid Information Service (GIS) faces a number of challenges. The three main challenges are:

- .Nature of the query (Number of dimensions and constraint type): Because they use hashing to distribute keys uniformly in order to achieve load balance DHTs only support exact match queries. Extending DHTs to support range queries is a complex problem by itself and for supporting multidimensional range queries the problem is complicated even more.
- Type of the Resource attribute: In contrast to static attributes dynamic attributes change very rapidly within shorter periods of time. Resource state information inserted into the GIS containing such attributes must be periodically refreshed to reflect the current state of resource. This however may lead to high network traffic since a large no of messages get overlay between routed nodes.

In order to reduce the response time, recent approaches to resolve resource discovery queries in a decentralized fashion attempt to cluster the data being inserted into a flat overlay network so that points close together are mapped to nodes that are close together in the Id space. In contrast, we propose a framework that clusters the overlay nodes (not the data) depending on the autonomous system(AS) and, consequently, the response time.

Queries are resolved at the target cluster. For gloal queries, those spanning all the domains participating in the system, we parallelize the search process. In addition, due the difference in nature between static attributes and dynamic attributes, different techniques are used to resolve queries on dynamic attributes that take into account the volatile nature of their measured values.

D. Problems with SAPRDG

If number of gateways increases in Grid Information System then size of latency matrix also increases. In the SAPRDG system overlays are using only for two level, For complicated GIS this may not be sufficient. Also any changes in latency should be announced to every gateway node in GIS, which in turn increases network traffic. Failure of one gateway node may lead to count to infinite problems (Bad new propagate slowly) in routing.

III. PROPOSED FRAMEWORK

In the proposed framework we are using ant colony optimization as proposed in [6] to discover global resources, whereas to discover local resources we propose to use gateway nodes[7]. Following Fig 1 shows structure of grid.

If you are using *Word,* use either the Microsoft Equation Editor or the *MathType* add-on (http://www.mathtype.com) for equations in your paper (Insert | Object | Create New | Microsoft Equation *or* MathType Equation). “Float over text” should *not* be selected.

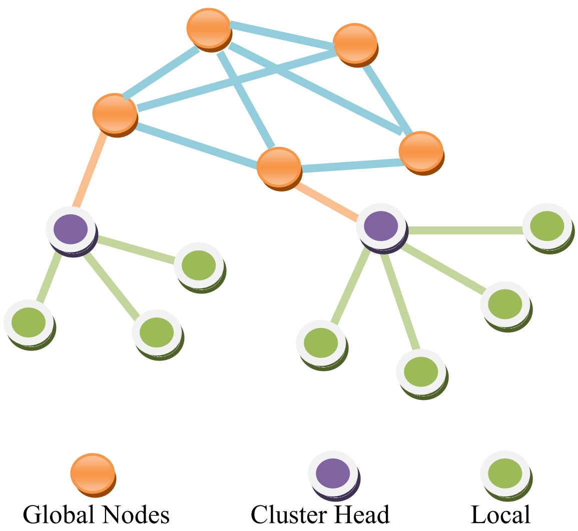

Fig:1 Grid Structure

IV. ANT COLONY OPTIMIZATION

Ant Colony Optimization (ACO) is a meta-heuristic optimization technique inspired by research on real ant foraging behavior. The ACO paradigm was first proposed by Dorigo and Blum, and has been successfully applied to diverse combinatorial optimization problems including traveling salesman, telecommunication networks and scheduling.

A. Construction of Solutions

A permutation of subtask scheduling subject to the constraint is termed a feasible solution. In the algorithm, the process of constructing a feasible solution can be divided into m steps. In the ith (i<0≤m)step , ant j finds one accessible node for subtask i , and then moves to the i+1st step for subtask i+1 scheduling until all the subtasks have been assigned, which can guarantee all the subtasks are scheduled. In the scheduling, all infeasible moves of ant j must be stored into a Tabu list denoted by Tabu(j). This list is the memory of ant j saving the index of infeasible allocations, and is initialized according to the hard constraint. Once node k is selected for subtask i by ant j , node k will be added to the Tabu(j) to avoid being selected again by ant j in the scheduling. In addition to the list Tabu(j) , all the nodes selected by ant j in m steps are stored in a memory represented by Path(j). A Path(j) is one feasible solution of subtask scheduling, and will be used to update the pheromones of moves selected by ants at the end of iteration

B. Selection Probability

In each step, an ant in a subtask chooses a node as the location of the corresponding subtask. This move can be represented by edge, which is the assignment of subtask to node . Ant chooses a node based on a combination of two factors, namely the desirability of that move, and the quantity of pheromone on the edge which is to be traversed.

C. Pheromone Updating Rule

Pheromone updating is a process of changing the quantity of pheromone over time on each edge. Before activating the next iteration, the quantity of pheromone on each edge is updated by the pheromone updating rule so as to avoid local convergence, and explore more search space as well.
The pseudo code of the ant colony optimization algorithm t solve the subtask scheduling in the grid is as follows.

1) Intialization (Initialize number of iteration).
2) Construction(For each ant in each step, choose t node to move into in probability. Update Tabu(j))
3) Pheromone update
4) Terminating condition

V. RESULTS

A. Simulation parameters

Table 5.1 shows simulation parameter setting while taking results.

Table 5.1 Simulation Parameters

Parameter Name	Value
Simulation time limit	1000s
Number Of Organisation	20
Number Of Nodes Per Organisation	5
Default CPU	100
Default Disk	500
Default Disk	1000
node[10].CPU	500
node[15].CPU	500
node[20].Memory	10000
node[24].Memory	10000
node[21].CPU	200
node[5].CPU	600
node[28].Disk	9000
node[35].Disk	5000

B. Available (free) Resources Graphs

Figure 5.1 shows availability of resources along with time. Hierarchy structure of resource request results into availability of resources to perform task at each node. If a particular node is running out of resource then it forwards request to gateway node. Gateway node maintains information about resources available at each node, so forwards resource request to appropriate node. This results in uniform utilization of resources at each node.

Proposed algorithm fails in some cases as shown in Figure 5.2. Every node advertises its resource information to Gateway node after allocation of resource as well as after releasing resources. If Gateway node forwards some request to node while advertise packet is still in network, it may so happen that enough resources are not available with node and the node is not able to process request. In Figure 5.2 this happened for one request.

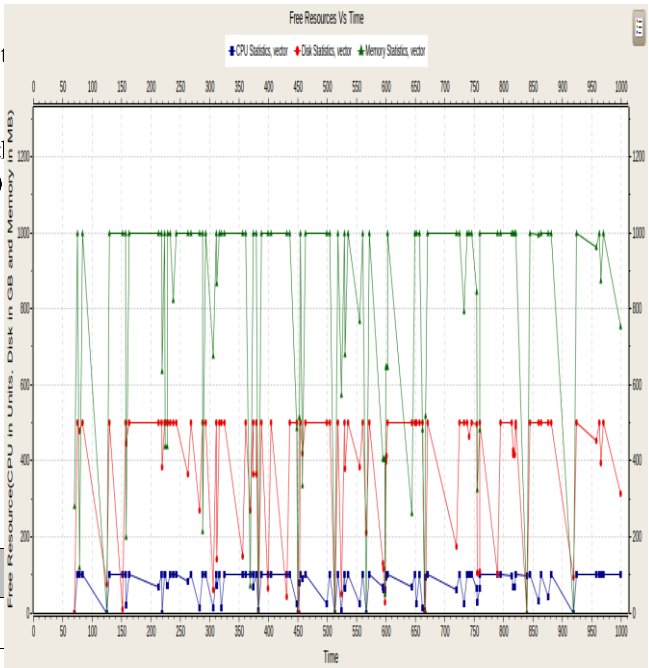

Figure 5.1 Available Resources Along with Time

ISSN: 2278 – 1323

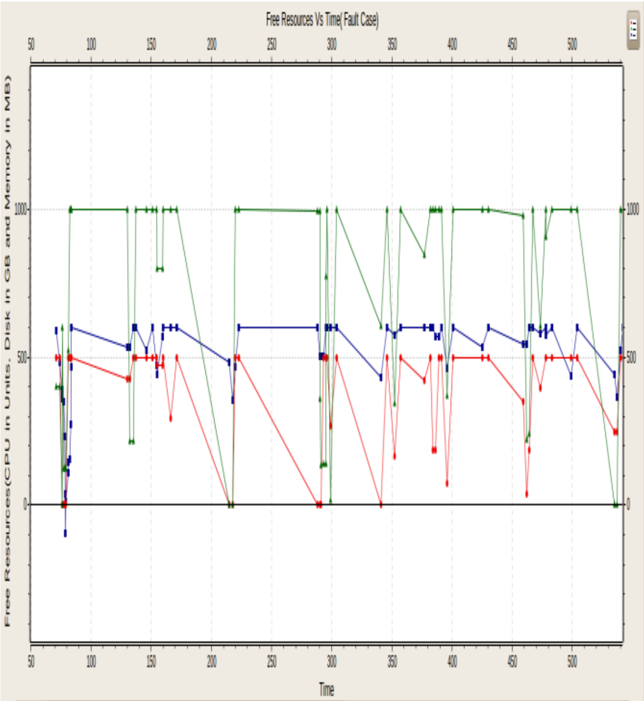

Figure 5.2 Available Resources along with Time (Fault Case)

C. Path Length to Different Resources Discovered by Ant

Gateway nodes forwards ***Forward-Ant*** messages according to pheromone value of outgoing link. If pheromone value is not present then it is forwarded thorough random outgoing link. ***Backward-Ant*** follows the same path as ***Forward-Ant*** in reverse direction and updates pheromone value at each node it visits. Pheromone value *ph* is calculated as follows:

ph= (resource value)/ hopcount;

This results in discovery of shortest path. Figure 5.3 shows path length to CPU

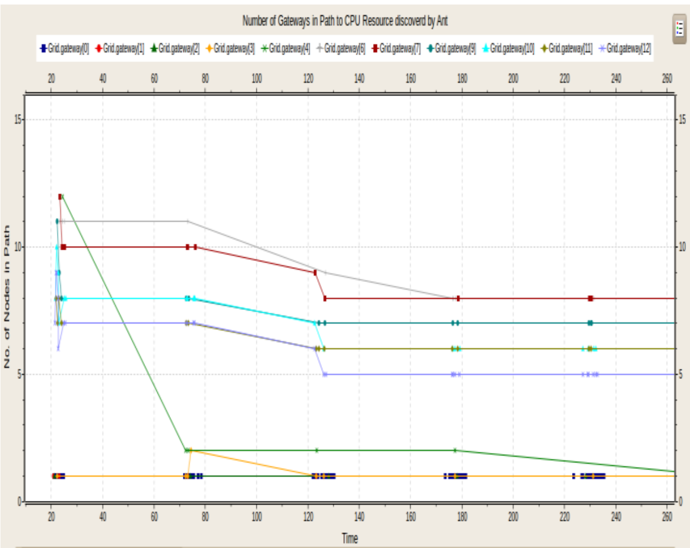

Figure 5.3 Path Length to Global CPU Resource

VI. Conclusion

Proposed algorithm performs fast discovery of global resources using ant colony optimization technique. From the results it is clear that proposed algorithm not only considers shortest path but also considers resource values to find best resource from other organisation. Dijkstra's shortest path algorithm requires every gateway node to know distance between every other node, which in turn increases memory requirement for Gateway nodes.

Future work includes finding out approach to decide after how much amount of time we should retransmit Ant Message, for how long Gateway node holds information given by Ant Packet etc. Future work will also include analysis of traffic generated by Ant messages and how to minimize it.

VII. REFERENCES

[1] Hesham A. Ali, Mofreh M. Salem and Ahmed Hamza, "Framework for scalable autonomous P2P resource discovery for the grid implementation", International Journal of Computer Science and Engineering, Vol. 1.3, summer, 2007.

[2] M. Dorigo and Gianni Di Caro, "Ant Colony Optimization: A New Meta-Heuristic" 0-7803-5536-9/99 1999 IEEE.

[3] S. C. Guo, H.Wan, G. B.Wang, and M. Xie, "Analysis of grid resource compensation in market-oriented environment," Maintenance and Reliability, vol. 45, no. 2, pp. 36–42, 2010.

[4] K. Krauter, R. Buyya, and M. Maheswaran, "A taxonomy and survey of grid resource management systems for distributed computing," Software—Practice and Experience, vol. 32, no. 2, pp. 135–164, 2002.

[5] M. Dorigo and C. Blum, "Ant colony optimization Meta-Heuristic" book.

[6] Grid Service Reliability Modeling and Optimal TaskSchedulingConsidering Fault Recovery IEEE TRANSACTIONS ON RELIABILITY, VOL. 60, NO. 1, MARCH 2011.

[7] A Framework for Scalable Autonomous P2P Resource Discovery for the Grid Implementation International Journal of Computer Science and Engineering Vol. 1 No. 3, WASET, 2007.

[8] I. Foster, "The Grid: A new infrastructure for 21st century science,"Physics Today, vol. 55, no. 2, pp. 42–47, 2002.

[9] Y. S. Dai, M. Xie, and K. L. Poh, "Reliability of grid service systems,"Computers and Industrial Engineering, vol. 50, no. 1–2, pp. 130–147,2006.

[10] S. C. Guo, H.Wan, G. B.Wang, and M. Xie, "Analysis of grid resource compensation in market-oriented environment," Eksploatacja I Niezawodnooeæ—Maintenance and Reliability, vol. 45, no. 2, pp. 36–422010.

[11] C. L. Li and L. Y. Li, "Multiple QoS modeling and algorithm in computational grid," Journal of Systems Engineering and Electronics, vol.18, no. 2, pp. 412–417, 2007

- Mr. LaxmanJ. Deokate is currently pursuing Master degree in Information Technology,Pune University, Maharashtra, India.

- *Mr.V.V.Puri has completed Master degree in Computer Science and Engineering, Pune University, Maharashtra, India.*

