A NOVEL DYNAMIC RESILIENT SCHEME ADAPTING FOR DISTRIBUTED ENVIRONMENTS

D. Porchelvi*

R. Lakshmi Priya*

Abstract-
For the past few years, the immerse popularity of Grid resource sharing services has produced a significant stimulus to content-delivery overlay network. In existing system, DHT is the data structure used for that mapping key to the nodes of a network based on a consistent hashing function. Most of the DHT overlays require O (n) hops per lookup request with O (n) neighbors per node, where n is the network size. A computing resource is constantly portrayed by a resource type such as CPU and memory, and resource attributes representing the quantity and specialty. Grid resource management scheme that is data locality-preserving is proposed. Dynamism-resilient resource management algorithm, cluster-token forwarding algorithm, and deadline-driven resource management algorithms are proposed. On adopting these algorithms consumes less overhead, speedy and dynamism-resilient multi resource discovery. Experimental results shows well-organized resource clustering, reduces communications cost, and enhances resource discovery success rate flexibly surviving in large scale applications.

Keywords- Clustering technique, DHT, Grid Computing, Range Queries.

* PSN College of Engineering and Technology

I. INTRODUCTION

A Computational Grid [1], or Grid, is a collection of Resources (computational devices, networks, online Instruments, storage archives, etc) that can be used as an ensemble. Locality preserving hashing ensures that similar keys are assigned to similar objects. This can enable a more efficient execution of range queries.

In grid, this functionality is provided by an information service. In an information service computing resources are characterized by sets of attributes, as for example the type of the operating system, network address, and CPU speed or storage capacity. A fundamental function of such a system is the search for resources with specific combinations of attribute values [2]. Due to a large number of resources, indexing of the attributes becomes necessary.

A distributed hash table (DHT) is a class of a decentralized distributed system that provides a lookup service similar to a hash table; (key, value) pairs are stored in a DHT, and any participating node can efficiently retrieve the value associated with a given key. Responsibility for maintaining the mapping from keys to values is distributed among the nodes, in such a way that a change in the set of participants causes a minimal amount of disruption. This allows a DHT to scale extremely large numbers of nodes and to handle continual node arrivals, departures, and failures.

To use a DHT overlay for resource discovery in a Grid system, all Grid nodes are organized into a DHT overlay. The descriptors of available resources are regarded as files and are distributed among the nodes. Resource queries are regarded as file lookups and are routed to the nodes having the descriptors of the required resources. Therefore, DHT overlays map the resource providers and consumers in Grids in a distributed manner.

Grids apply centralized or hierarchical control .Heterogeneous Grid resources are governed by different organizations spread throughout the World Wide Web. But they have limitations in supporting large-scale dynamic Grid applications with a varying demand of resources. The available resources called as files and are distributed among the nodes. Resource queries called as file lookups and are routed to the nodes consisting of the descriptors of the required resources. In existing, there is no proper specification for locating the resources in DHT. In Proposed system, Hierarchical Cycloid Overlay Architecture is used to collect the close resources and forms the cluster. Physical Proximity is used for the cluster Formation. The resources are grouped based on functionalities of each resources.

The available resources and the request resources of all nodes are collected in one cluster by using the cluster token forwarding technique. Discovery of resources can be done by using this PRP. Different Priorities of Resources are accumulated. The available requests are immediately processed without waiting for the next periodical resource allocation. The resource requests are processed in the ascending order of their deadlines.

The rest of this paper is structured as follows: Section II presents a concise review of DHT overlay and HCO. Section III briefly introduces the properties of DHT related to clustering. Section IV describes the System architecture. Section V presents the locality preserving clustering. Section VI presents the cluster token forwarding algorithm and deadline driven resource management algorithm. Section VII concludes this paper.

II. RELATED WORKS

The Locality-Preserving resource clustering and discovery algorithms based on the HCO architecture. To support multi resource queries, Mercury [7] uses multiple DHT overlays. It uses one DHT for each resource, and processes multi-resource queries in parallel in corresponding DHT overlays. However, depending on multiple DHT overlays leads to high overhead for DHT maintenance. Some of the approaches [8], [9] organizes all Grid resources into one DHT overlay, and assigns all descriptors of one resource to one node. MAAN maps attribute values to the Chord identifier space via uniform locality preserving hashing.

To facilitate efficient range queries, andrzejak and Xu proposed a CAN-based approach for grid information services [9]. Most of the single DHT-based approaches assign one node to be responsible for all descriptors of one resource, leading to an imbalanced distribution of workload. Also, if one of the nodes fails in dynamism, many descriptors will be lost at a time.

By using a single DHT, HCO generates much lower maintenance overhead than the methods based on multiple DHTs. HCO is more dynamism-resilient than the methods based on a single DHT by distributing resource descriptors of one resource among a number of nodes. More importantly, few of the current approaches [5] can achieve the program/data locality to facilitate low overhead and quick resource discovery.

HCO establishes program/data locality [12], which enables users to discover their required physically close resources from their nearby nodes on the DHT overlay. These features contribute to the high scalability and an efficiency characteristic of HCO in Grid resource management.

SOMO [10] performs resource management in DHT networks by embedding a tree structure into a DHT structure. Some approaches [15], focus on weaving all attributes of a resource into one or a certain number of IDs for resource clustering and discovery in a DHT overlay.

All resource discovery operations are conducted in the overlay layer in a distributed manner. One major challenge in resource clustering is to keep the logical proximity and physical proximity of resource nodes consistent. A landmark clustering is adopted to generate proximity information [9].

Cycloid [6] assigns an object to the node whose ID is closest to its ID. It provides two main functions: Insert (ID, object) stores an object to a node responsible for the ID and Lookup (ID) retrieves the object through DHT-based searching. Each node maintains a routing table recording its neighbors in the overlay network for object lookups. Like all other DHTs, when a node joins in cycloid, it generates its cycloid ID and relies on a bootstrap node to find its neighbors and connect to them.

III. DISTRIBUTED HASH TABLE

DHTs characteristically emphasize the following properties:

1. Autonomy and Decentralization: the nodes collectively form the system without any central coordination.
2. Fault tolerance: the system should be reliable (in some sense) even with nodes continuously joining, leaving, and failing.
3. Scalability: the system should function efficiently even with thousands or millions of nodes.

A. DHT BASED HIERARCHY

A DHT-based hierarchy is prescribed for locality preserving grid resource clustering and discovery by the architecture and processing layers of HCO. Different clusters are formed by various group of nodes based on their physical proximity. All the operations in resource discovery are conducted in the overlay layer in a distributed manner. The major challenge in resource clustering is to keep the logical proximity and physical proximity of resource nodes consistent. A landmark clustering is adopted to generate proximity information [13], [14].

HCO is also applicable to Chord [3] and Pastry [4], i.e., a hierarchical DHT based on Chord or Pastry. The lower-layer of each cluster is responsible for storing resource directories of physically close nodes in the hierarchical structure. Nodes in a cluster are responsible for directories of different resources. In HCO, node ID and the Hilbert number Hi are generated by the node itself.

B. CLUSTERING

- A cluster is an ordered list of objects, which have some common characteristics.
- The distance between two clusters involves some or all elements of the two clusters.
- The clustering method determines how the distance should be computed.
- Fuzzy clustering is a method of cluster analysis used for the formation of clusters.

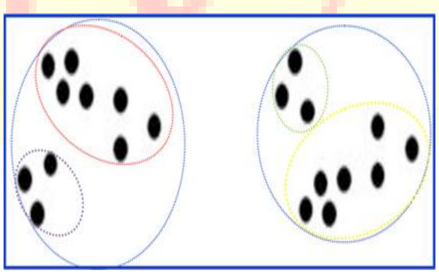

Figure 1-Clustering

A cluster consists of a set of loosely connected computers that work together so that in many respects they can be viewed as a single system. Clusters emerged as a result of convergence of a number of computing trends including the availability of low cost microprocessors, high speed networks, and software for high performance distributed computing.

Clusters are usually deployed to improve performance and availability over that of a single computer, while typically being much more cost-effective than single computers of comparable speed or availability.

IV. SYSTEM ARCHITECTURE

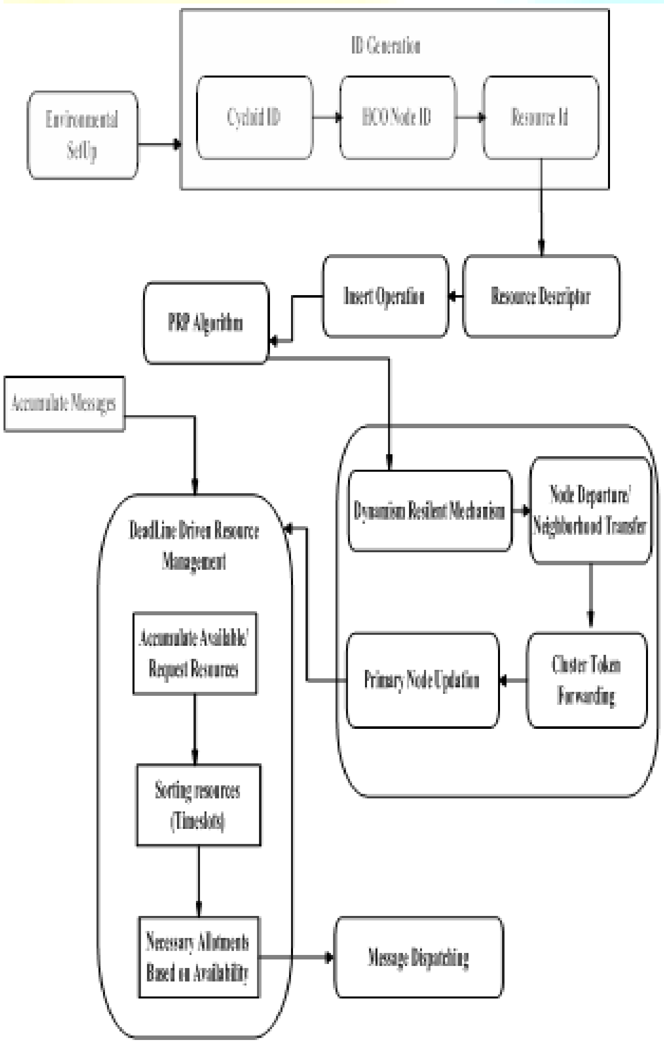

Figure 2 -Architecture

V. LOCALITY PRESERVING GRID RESOURCE MANAGEMENT

A. ENVIRONMENTAL SETUP

Creation of Nodes is the initial process. In grid computing the systems will be interconnected, the task will be shared by the systems. In this module the grid environment is constructed. In this Module the cycloids are formulated. Cycloid is a look up efficient overlay network generalized from the cube connected cycles (CCC). A d dimensional cycloid is constructed having n= d* 2d nodes.

B. IMPLEMENTING HCO NETWORK

1. Id Generations are done here. Initially location of cluster (id l) or node is computed, followed by Position within a cluster (id l). Idl are grouped.
2. Distances between two node id s are evaluated by distance equation.
3. HCO Node constraints are evaluated. Distance between nodes in one cluster denoted as Hilbert Number (H) of a node is calculated; Nodes with closer H are closer together.
4. Each Node has a consistent Number; hence this number is computed.

C. LOCALITY PRESERVING RESOURCE CLUSTERING

Resource Descriptor is formulated. A node reports its available resources or requests for resources using a resource descriptor Dr, Resource Descriptor Dr <RF, ID,RA, IP>; where RF, ID, and RA are the resource functionality, identifier, and resource attribute ,IP refers to the IP address of the resource owner or requester.

The objects with the same ID are stored in the same node in this DHT Overlay. HCO computes the reliable hash value Hr of a resource r's resource functionality, and uses IDr = (Hr, Hi) to represent the ID of resource r.

The directory node's resource are used to represent the resource r whose Dr is stored in the directory node. The logical distances between the nodes i and a number of directory nodes represent the physical distances between nodes i and the directory nodes' resources.

Therefore, if a node has resource options in a number of directory nodes, it should choose the resource in the logically closest directory node in the overlay.

D.RESOURCE DISCOVERY

HCO applies the cycloid self-organization mechanism to cope with the problems of to make use of the physical proximity of the network nodes and to diminish the operation cost. Hence Dynamism Resilient Resource Management is applied. If a node leaves its descriptors are forwarder to the next close system.

On Applying Cluster Token forwarding Algorithm, it accumulates messages of available resources & request of all nodes in one cluster. Primary node in each cluster periodically generates the token and transfers to other cluster.

E. RESOURCE ALLOCATION

The deadline driven resources management technique is applied. As per this approach the problem of required resources allocation in time line fashion. Each node reports its available resources in future time slots. Instant demand of resources may be got from the requestor's .Availability Resource List is got from the Nodes, requested Resource list is got from the directory node. Both the list are sorted and placed. Depending on the availability and deadline the resources list and request list are forwarded.

VI. ALGORITHM DESCRIPTION

Algorithm1. Pseudo code for node i's operations in HCO.

Generate the consistent hash value for its resources available list and resources request list .Request resources are processed using Proximity aware Randomized Probing algorithm. As Per this approach, the probing is made on nearest nodes.

Here the proximity-aware randomized probing algorithm (PRP) is developed based on the algorithm in [11] to resolve the problem. In the PRP algorithm, sequential probing is first applied by a node. If no response is received during a predefined time period, the node randomly chooses two nodes in an increasing range of proximity and repeats the probing process.

Since resource descriptors are allocated to different nodes in a cluster based on resource functionality, the probed nodes should be the directory nodes of the requested resource.

Algorithm2. Cluster Token Forwarding Algorithm

1. Token is generated by the primary node in each cluster. After that, primary node receives a token, and then it will forward its available resource descriptor.
2. After receiving a token, the primary node forwards the token to a node in another cluster using PRP.
3. As per the availability of resources in the nodes the request will be executed. This algorithm is for Resource Discovery.

Algorithm3. Pseudo code for Resource Allocation performed by a directory node

1. Generate Available Resource List and Request Resource List. Sort the ARL and RRL in ascending order of the time T.
2. Availability Resource List is got from the Nodes, requested Resource list is got from the directory node. Both the list are sorted and placed.

3. To handle the resource requests immediately, we define a parameterized threshold E for T to determine if a request is urgent or not. Depending on the availability and deadline the resources list and request list are forwarded.

VII. CONCLUSION

There is a demand for scalable and efficient resource management scheme to maintain distributed performance in a dynamic wide-area environment. Locality-preserving algorithms are proposed to enable users dynamically and to discover physically close resources. The HCO scheme employs a single large DHT overlay with low overhead. A balanced workload distribution is achieved that is resilience to resource failure. Cluster-token forwarding and deadline-driven resource management algorithms employed increases efficiency.

References

[1] I. Foster and C. Kesselman, "Globus: A Metacomputing Infrastructure Toolkit," The Int'l J. Supercomputer Applications and High Performance Computing, vol. 11, no. 2, pp. 115-128, 1997.

[2] F. Berman et al., "Adaptive Computing on the Grid Using Apples," IEEE Trans. Parallel Distributed Systems, vol. 14, no. 4, pp. 369-382, Apr. 2003.

[3] I. Stoica et al., "Chord: A Scalable Peer-to-Peer Lookup Protocol for Internet Applications," IEEE/ACM Trans. Networking, vol. 11, no. 1, pp. 17-32, Feb. 2003.

[4] A. Rowstron and P. Druschel, "Pastry: Scalable, Decentralized Object Location and Routing for Large-Scale Peer-to-Peer Systems,"Proc. IFIP/ACM Int'l Conf. Distributed Systems Platforms Heidelberg (Middleware '01), 2001.

[5] B.Y. Zhao, L. Huang, J. Stribling, S.C. Rhea, A.D. Joseph, and J. Kubiatowicz , "Tapestry: An Infrastructure for Fault-Tolerant Wide-Area Location and Routing," IEEE J. Selected Areas in Comm.,vol. 22, no. 1, pp. 41-53, Jan. 2004.

[6] H. Shen, C. Xu, and G. Chen, "Cycloid: A Scalable Constant-Degree P2P Overlay Network," Performance Evaluation—P2PComputing Systems, vol. 63, no. 3, pp. 195-216, 2006.

[7] A.R. Bharambe, M. Agrawal, and S. Seshan, "Mercury: Supporting Scalable Multi-Attribute Range Queries," Proc. ACM SIGCOMM, pp. 353-366, 2004.

[8] M.Cai et al., "MAAN: A Multi-Attribute Addressable Network for Grid Information Services," Proc. J. Grid Computing, 2004.

[9] D. Oppenheimer et al., "Scalable Wide-Area Resource Discovery," Technical Report TR CSD04-1334, EECS Dept., Univ. of California, 2004.

[10] A. Andrzejak and Z. Xu, "Scalable, Efficient Range Queries for Grid Information Services," Proc. Second Int'l Conf. Peer-to-Peer Computing (P2P '02), 2002.

[11] Z. Zhang, S.-M. Shi, and J. Zhu, "Somo: Self-Organized Metadata Overlay for Resource Management in P2P DHT," Proc. Int'l Workshop Peer-to-Peer System (IPTPS), 2003.

[12] H. Shen and C. Xu, "Locality-Aware and Churn-Resilient Load Balancing Algorithms in Structured Peer-to-Peer Networks," IEEE Trans. Parallel and Distributed Systems, vol. 18, no. 6, pp. 849-862, June 2007.

[13] F.P. Preparata and J. Vuillemin, "The Cube-Connected Cycles: A Versatile Network for Parallel Computation," Comm. ACM, vol. 24, no. 5, pp. 300-309, 1981.

[14] S. Ratnasamy, M. Handley, R. Karp, and S. Shenker, "Topologically- Aware Overlay Construction and Server Selection," Proc. INFOCOM, 2002.

[15] Z. Xu, M. Mahalingam, and M. Karlsson, "Turning Heterogeneity into an Advantage in Overlay Routing," Proc. INFOCOM, 2003.