Improving the Availability of Replication in Non-dedicated Resources

Yuanyuan Zhang [1], Wei Sun [1], Yasushi Inoguchi [2]
[1]Graduate School of Information Science,
[2]Center for Information Science,
Japan Advanced Institute of Science and Technology,
1-1, Asahidai, Nomi, Ishikawa, 923-1292 Japan
{ yuanyuan, sun-wei, inoguchi}@jaist.ac.jp

Abstract

Grid computing often involves large quantity of heterogeneous resources from autonomous administrative domain, which could be dedicated resources, but most of which could be non-dedicated resources especially in commodity computing or business grid. The availability of data replication should be guaranteed in the non-dedicated resources. We proposed a technique to improve the availability of data replication, which reduce the negative influence of non-dedicated resources as far as possible.

1. Introduction

Data replication is an optimization technique well known in the distributed systems, which can achieve better access and fault tolerance. In grid computing [1] data replication is a main technique for some research field, for example data management, replica management, and grid file system [2, 3, 4].

Different from other distributed systems, grid authorizes the resource owners enough autonomy. On the other hand there are many temporary resources in grid system. Thus all of these kinds of resources can be looked as non-dedicated resources. In the dedicated resource environment an unavailable resource is a fault for data replication, so there are many fault tolerance techniques in data replication. A common method is to produce more copies of the original data. In the non-dedicated resources environment it is not a fault when a resource is unavailable and the unavailable resource might recover in some time.

In this poster we propose a technique of replication for non-dedicated resources.

2. Replication technique

2.1. Definition

We list some terms used in our poster as follows.

- Well-known resource: the regulation of the time when the resource is available or unavailable is known well.
- Less-known resource: there is no clear regulation of the time when the resource is available or unavailable.
- Complementary resource set: resources in the set can complement the unavailable time for each other to form continuous available time, which are shown in figure 1.

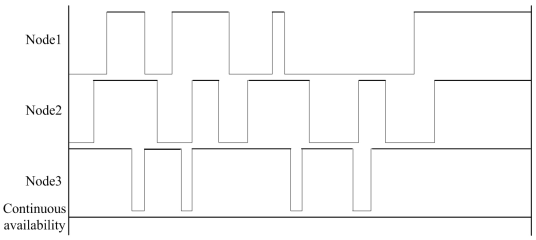

Figure 1: Complementary resource set

2.2. Static replica

According to the requirement of dependability in a system, we can get a degree of Redundancy ***R***. For well-known resources in a grid, the number of replica should be decided by the size of complementary resource set ***S*** and ***R***. Therefore, 1) we should find ***R*** independent complementary resource sets, each of whose size is the possible minimum size in the current resources; 2) the replicas are sent to each node in each complementary resources set. 3) The copy in the node with longest available time is selected as the

master replica. Replica will be updated in terms of the master replica. 4) If a replica is unavailable, a node other than current selected nodes, which can make up for the unavailable one, will be chosen to join the corresponding complementary resource set. If the disable replica is a master replica, a new master replica will be voted after 4), that is, redo 4) and 3). 5) If the disable replica recovers and the data is intact and available, system will hold this replica and this replica will have a high priority to be used to make up for next disable replica. The static replica is shown in Figure 2.

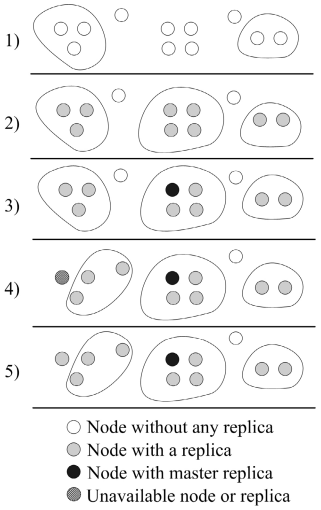

Figure 2: Static replica

The static replica is at the cost of storage space. The performance and complexity of static replica is decided by the quality of resources.

2.3. Mobile replica

For less known resources or if there are not enough resources for static replica, mobile replica has to be used to improve the availability of data replication. According to the degree of Redundancy ***R***, 1) ***R*** nodes are selected to receive the data copies. All the replicas are updated simultaneously. If one replica or more is unavailable, 2) one new replica or more will be created in another nodes. The replica set looks like a "mobile" set. 3) If one replica or more is recovered and the data is intact and available, system will hold this replica and this replica will have a high priority to be used to make up for next disable replica. The mobile replica is shown in Figure 3.

The mobile replica is at the cost of communication. The performance of mobile replica is also decided by the quality of resources.

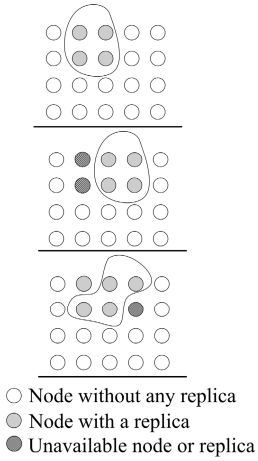

Figure 3: Mobile replica

2.4. Avoid local optimum

Both the static replica and mobile replica are prone to local optimum, because the recovered replica has the high priority just as what are mentioned before. For the static replica the local optimum is not so serious, since the size of each complementary resource set is the minimum, that is, almost all good quality resources have been chosen into the complementary resource sets.

For the mobile replica the times of invalidation for each recovered node will be recorded. If the recorder is beyond the defined threshold, a new node without recorder for this replica will be selected to receive the replica. Moreover, in case that there is no new node in the resources, the node with the minimum recorder will be used again.

3. Conclusion

In this poster we propose a method to improve availability of replication, which is designed to cope with the non-dedicated environment. We hope it is a helpful technique for the system with non-dedicated resources.

4. References

[1] I. Forster and C. Kesselman, *The Grid 2: Blueprint for a New Computing Infrastructure*, Morgan Kaufmann, San Francisco, 2004.

[2] H. Stockinger, A. Samar, K. Holtman et al., "File and Object Replication in Data Grid", Cluster Comuting, 2002.

[3] L. Guy, P. Kunszt, E. Laure et al., "Replica Management in Data Grids", GGF, 2002. Computing", PhD dissertation, 2002.

[4] http://www.ggf.org