

## **A clustering-based hybrid replica control protocol for high availability in grid environment**

### **ABSTRACT**

In recent years, with the emergence of grid computing system, the number of distributed sites has become very large. When thousands of sites involved in a grid computing system, data replication can improve data availability, communication cost and provide fault-tolerance in the system. In the literature, many replica control protocols have been proposed for managing replicated data. However, in large scale distributed system, most of these protocols still require a bigger number of replicas for maintaining consistency, thus degrade the performance of the protocols. Therefore, in this study, we proposed a new replica control protocol named Clustering-Based Hybrid (CBH) protocol. CBH protocol employs a hybrid replication strategy by combining the advantages of two common replica control protocols into one to improve the performance of the existing protocols. We analyzed the communication cost and availability of the operations and compare CBH protocol with recently proposed replica control protocol named Dynamic Hybrid (DH) protocol. A simulation model was developed using Java to evaluate CBH protocol. Our results show that the proposed protocol provides lower communication cost and higher data availability than DH protocol.

**Keyword:** Replica control protocol; Data replication; Data availability; Communication cost