Research of Indirect Replication Algorithm in Distributed Storage System*

Wang Yijie, Li Sikun
Institute of Computer, National University of Defense Technology, Changsha, Hunan, 410073, China
Email: wwyyjj1971@vip.sina.com **Tel.:** (0731)4573663

ABSTRACT

Replication is the key technology of distributed storage system. In this paper, according to the intrinsic characteristic of distributed storage system, based on the peer-to-peer model, the indirect replication algorithm is proposed. In the indirect replication algorithm, the data object is partitioned into several data blocks, and then these data blocks are encoded in order that there is data redundancy between data blocks. Compared with the traditional replication algorithm, the indirect replication algorithm has less granularity of replication、less bandwidth cost and storage cost, and can provide higher availability 、 durability and security. The results of performance evaluation show that the encode time and decode time is proportional to the square of data size, and that if the number of encoded data blocks used to recover data object increases, the decode time is decreased greatly.

Keywords: Distributed storage system, replication, data availability, performance evaluation.

* This work is supported by the National Grand Fundamental Research 973 Program of China (No.2002CB312105), A Foundation for the Author of National Excellent Doctoral Dissertation of PR China (No.200141), and the National Natural Science Foundation of China (No.69903011, No.69933030).