Accepted Manuscript

Fault tolerant storage and data access optimization in data center networks

Yang Qin, Weihong Yang, Xiao Ai, Lingjian Chen

PII: S1084-8045(18)30127-9

DOI: 10.1016/j.jnca.2018.04.001

Reference: YJNCA 2115

To appear in: *Journal of Network and Computer Applications*

Received Date: 2 September 2017

Revised Date: 19 February 2018

Accepted Date: 6 April 2018

Please cite this article as: Qin, Y., Yang, W., Ai, X., Chen, L., Fault tolerant storage and data access optimization in data center networks, *Journal of Network and Computer Applications* (2018), doi: 10.1016/j.jnca.2018.04.001.

Fault Tolerant Storage and Data Access Optimization in Data Center Networks

Yang Qin*, Weihong Yang, Xiao Ai and Lingjian Chen

Harbin Institute of Technology (Shenzhen), Shenzhen, China

*corresponding author: csyqin@hit.edu.cn

Abstract—Data center, of which performance affects the cloud computing, is the key part of cloud computing technology. As cloud computing has penetrated into our daily life, the data stored in data center have increased rapidly, which increases the requirements of fault tolerance of distributed storage system. In the meantime, the efficiency of data access should be considered when applying network coding scheme in data center networks. This paper tackles these two issues. The first part of this paper proposes a fault-tolerant storage based placement strategy. We consider three factors: the efficiency of data access, the load of storage node, and the expectations of recovery time of failure nodes. When the loads of nodes are balance and recovery time for failure nodes is shorter, we can achieve a high fault-tolerance storage in data center networks. Simulation results show that the fault-tolerant placement has low expectation of recovery time and can achieve load balance. In the second part of this paper, we design a priority-queue based scheduling policy of data access to improve the efficiency of data access when deploying network coding in data center networks. Simulation results show that the proposed priority queue scheduling policy can achieve low delay by reducing the number of frozen servers.

Keywords—data center; distributed storage systems; data placement strategy; fault tolerance

1 Introduction

The development of distributed computing technology has recently motivated the research of improving the fault tolerance for distributed systems. One of the methods is to design reliable and reasonable data placement strategy. The traditional data fault-tolerant placement strategies are sequential and random placement strategies. The latest strategies take the efficiency of placing and data access into consideration, while the fault tolerance has been ensured.

The sequential placement strategy places the data blocks on the candidate nodes in a certain order. Generally, it used to be applied in the chained declustering (Hsiao and DeWitt, 1990; Lee and Thekkath, 2001) and some distributed hash table structure (Dabek et al., 2001; DeCandia et al., 2007; Rowstron and Druschel, 2001). The sequential placement strategy is simple and easy to implement. However, it's likely to trigger cascading failures. Random placement strategy selects candidate nodes randomly to place data blocks. Random placement strategy is applied in many modern data centers, such as GFS (Ghemawat et al., 2003) and Cassandra (Laksham Avinash and Prashant Malik, 2010). Although placing data blocks randomly can reduce the impacts of cascading failures on the reliability, it has some disadvantages on load balance of nodes. HDFS (Hadoop Distributed File System) (Borthakur, 2008; Shvachko et al., 2010) place the second and third duplicates on the same rack, which aim to save the replication and transmission time. However, it will take a large amount of data transmission time to repair data from the remote nodes when the local duplicates lose. Moreover, it is likely to lose load balance to select candidate nodes randomly.

Therefore, in the first part of this paper, we design a fault-tolerant storage based placement strategy for data center networks to improve the efficiency of data access. In the proposed strategy, the recovery time and the load balance have been taken into consideration jointly.

In a distributed system, the data are stored in different nodes and will be accessed by more than one user. Previous concerns (Dimakis et al., 2010; Dimakis et al., 2007) have focused on that network coding can improve reliability of storage and facilitate data recovery. Today's research shows that coding can also improve the user's access efficiency. The main influence of user access efficiency is the queuing delay of access request. In an uncoded distributed storage system, when a user request is scheduled, the user request is divided into several sub-requests, and each sub-request can only be