Chapter 46
Byzantine Fault-Tolerant Architecture in Cloud Data Management

Mohammed A. AlZain
Taif University, Saudi Arabia

Alice S. Li
La Trobe University, Australia

Ben Soh
La Trobe University, Australia

Mehedi Masud
Taif University, Saudi Arabia

ABSTRACT

One of the main challenges in cloud computing is to build a healthy and efficient storage for securely managing and preserving data. This means a cloud service provider needs to make sure that its clients' outsourced data are stored securely and, data queries and retrievals are executed correctly and privately. On the other hand, it may also mean businesses are willing to outsource their data to a third party only if they trust their data are not accessible and visible to the service provider and other non-authorized parties. However, one of the major obstacles faced here for ensuring data reliability and security is Byzantine faults. While Byzantine fault tolerance (BFT) has received growing attention from the academic research community, the research done is generally from the distributed computing point of view, and hence finds little practical use in cloud computing. To that end, the focus of this paper is to discuss how these faults can be tolerated with the authors' proposed conceptualization of Byzantine data faults and fault-tolerant architecture in cloud data management.

DOI: 10.4018/978-1-5225-8176-5.ch046

1. INTRODUCTION

One of the main challenges in cloud computing is to build a healthy and efficient storage for securely managing and preserving data. It means a cloud service provider needs to make sure that its clients' outsourced data are stored securely and, data queries and retrievals are executed correctly (in terms of reliability) and privately (in terms of three security attributes: Confidentiality, Integrity and Availability – CIA) (AlZain, Soh et al., 2013; AlZain, Soh et al., 2011; AlZain, Pardede et al. 2012; AlZain, Soh et al., 2013; AlZain, Li et al., 2015). On the other hand, it also means businesses are willing to outsource their data to a third party (cloud service provider) only if they trust their data are not accessible and visible to the service provider and other non-authorized parties (Hore, Mehrotra et al. 2004). However, one of the main obstacles here for ensuring data reliability and security is Byzantine faults. The focus of this paper is to discuss how these faults can be tolerated with our proposed conceptualization of Byzantine data faults and fault-tolerant architecture in cloud computing.

The remainder of this paper is organized as follows. Section 2 presents the background and related work. Section 3 overviews our conceptualization of Byzantine data faults as well as fault tolerant state machines which are the base of our proposed fault-tolerant architecture in cloud data management (in Section 5). Section 4 details three crucial operations in the proposed fault-tolerant architecture. Section 6 gives a qualitative evaluation of the proposed architecture, while Section 7 concludes the paper.

2. BACKGROUND AND RELATED WORK

If something happens to the data or if the data is corrupted by the service provider, the service provider is responsible for data restoration. The service provider should have a mechanism to recover or back-up the data (Agrawal, El Abbadi et al., 2009). There are three issues to be addressed for wide adaptation of the data storage framework in terms of data security: Cryptographic techniques, Private Information Retrieval, and Data Replication Techniques (Agrawal, El Abbadi et al., 2009). These techniques are commonly used for secure data outsourcing.

HAIL (High Availability and Integrity Layer) (Bowers, Juels et al., 2009) is another example of a protocol that controls multiple clouds. HAIL is a distributed cryptographic system that permits a set of servers to ensure that the client's stored data is retrievable and integral. HAIL provides a software layer to address the availability and integrity of the stored data in multi-clouds (Bowers, Juels et al., 2009).

Agrawal et al. (Agrawal, El Abbadi et al., 2009) discuss the issue of information distribution (in terms of data query and retrieval) with the aim of showing that there is an orthogonal approach which is based on information distribution instead of encryption in the area of data and computer security. The need to communicate important or private information from one party to another instigated most of the work on data security. Agrawal et al. (Agrawal, El Abbadi et al., 2009) introduced Shamir's Secret Sharing algorithm (Shamir, 1979) as a solution for the privacy issue.

Data Replication is one of the important approaches for outsourced data security. The data owner divides data and replicates them into different data storage or multi-clouds. RACS (Redundant Array of Cloud Storage) (Abu-Libdeh, Princehouse et al., 2010) for instance, utilizes RAID-like techniques that are normally used by disks and file systems, but for multiple cloud storage. Abu-Libdeh et al. (Abu-Libdeh, Princehouse et al., 2010) assume that to avoid "vender lock-in", distributing a user's data among multiple clouds is a helpful solution. This replication also decreases the cost of switching providers and offers