Data Replication in Cloud Systems: A Survey

Khaoula Tabet, University Larbi Tebessi, Department of Mathematics and Computer Science, Tebessa, Algeria
Riad Mokadem, Paul Sabatier University, Toulouse, France
Mohamed Ridda Laouar, University Larbi Tebessi, Department of Mathematics and Computer Science, Tebessa, Algeria
Sean Eom, Southeast Missouri State University, Management Information Systems, Cape Girardeau, MO, USA

ABSTRACT

This paper presents a survey of data replication strategies in cloud systems. Based on the survey and reviews of existing classifications, we propose another classification of replication strategies based on the following five dimensions: (i) static vs. dynamic, (ii) reactive vs. proactive workload balancing, (iii) provider vs. customer centric, (iv) optimal number vs. dynamic adjustment of the replica factor and (v) objective function. Ideally, a good replication strategy must simultaneously consider multiple criteria: (i) the reduction of access time, (ii) the reduction of the bandwidth consumption, (iii) the storage resource availability, (iv) a balanced workload between replicas and (v) a strategic placement algorithm including an adjusted number of replicas. Therefore, selecting a data replication strategy is a classic example of multiple criteria decision making problems. The taxonomy we present can be a useful guideline for IT managers to select the data replication strategy for their organization.

KEYWORDS

Classification, Cloud Computing, Data Replication Strategies, Multiple Criteria Decision Making

INTRODUCTION

With the increasing globalization of contemporary business organizations, distributed databases and their management have become one of the key areas in database research. A distributed database is a single logical database scattered across multiple computers. Basically, there are two options for distributing a database: data partitioning or data replication. Data replication is one of the important decisions in organizations (Km & Eom, 2016). It refers to the creation of identical copies of data (replicas). Data partitioning is another strategy for distributing a database that breaks a table into multiple records (horizontal partitioning) or multiple columns (vertical partitioning).

This paper presents a survey of data replication strategies in cloud systems. Data replication improves data availability, response time, fault tolerance, and reduces network traffic. It is frequently used in: (i) DBMS (Pérez, García-Carballeira, Carretero, Calderón, & Fernández, 2010), (ii) parallel and distributed systems (Loukopoulos, Lampsas, & Ahmad, 2005; Benoit, Rehn-Sonigo, & Robert, 2008), (iii) mobile systems (Tos, Mokadem, Hameurlain, Ayav, & Bora, 2016) and (vi) large scale systems, including P2P(Xhafa, Kolici, Potlog, Spaho, Barolli, & Takizawa, 2012) and data Grid

DOI: 10.4018/IJISSC.2017070102

Copyright © 2017, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
systems (Mansouri, Azad, & Chamkori, 2014). Many replication strategies proposed aim to answer the following questions:

- What data should be replicated?
- When should the data be replicated
- Where the new replicas should be placed?

Data replication is a necessary tool for effectively managing a database in distributed database environments. Most of the works in the literature have classified replication strategies based on the following criteria: (i) static vs. dynamic classification (Chervenak, Deelman, Foster et al., 2002; Čibej, Slivnik, & Robič, 2005), (ii) centralized vs. decentralized replication (Sashi & Thanamani, 2011; Amjad, Sher, & Daud, 2012; Grace & Manimegalai, 2014), (iii) server vs. client replication (Doğan, 2009; Steen & Pierre, 2010), (iv) objective function based classification (Mokadem & Hameurlain, 2015), and (v) system architecture based classification (Tos, Mokadem, Hameurlain et al., 2015). However, the existing replication strategies are not adapted to the cloud system. They aim to obtain the best performance without taking the profit of cloud providers or the satisfaction of tenant requirements into account. Creating as many replicas in clouds may not be economically feasible. Hence, replication strategies in such environments should also ensure both a tenant Quality of Service (QoS) and the economic profitability of the provider.

This paper presents a survey of data replication strategies in cloud systems. We propose another classification of replication strategies based on the following five dimensions:

- Static vs. dynamic (Ghemawat, Gobioff, & Leung, 2003; Bai, Jin, Liao et al., 2013);
- Reactive vs. proactive workload balancing (Silvestre, Monnet, Krishnaswamy et al., 2012; Hussein & Mousa, 2014);
- Provider-centric vs. customer-centric (Sakr & Liu, 2012; Sousa & Machado, 2012);
- Minimal blocking probability (Xue, Shen, & Guo, 2015) and energy efficiency and bandwidth consumption (Boru, Kliazovich, Granelli et al., 2015);
- Objective function (Bonvin, Papaioannou, & Aberer, 2011; Kirubakaran, Valarmathy, & Kamalanathan, 2013; Tos et al., 2016).

The rest of this paper is organized as follows. Section 2 reviews existing classifications and how dynamic replication can be classified with respect to multiple criteria. Section 3 proposes a new taxonomy of data replication strategy in Cloud systems with respect to the several multiple dimensions. Section 4 discusses some important factors that impacts performance of these strategies. The final section outlines future research directions.

**REVIEW OF EXISTING REPLICATION STRATEGY CLASSIFICATION**

Most of the proposed replication strategy in cloud systems aim to increase data availability, decrease data access time, reduce access latency, increase reliability of the cloud system, and minimize energy consumption. Finding an optimal number of replicas in the cloud is an important part of the strategy formulation in order to avoid unnecessary replications, which can generate an overhead in the system (Boru et al., 2015). Although storage cost is becoming less important, some replication strategies assume unlimited storage that is not realistic, while others assume a fixed and limited amount of storage. Reducing access latency is also an important objective of replication strategies since it reduces the job execution time (Xue et al., 2015). Most of the replication strategies require a tradeoff between different criteria that impact the system performance. A significant number of replication strategies that have been proposed in the literature are classified according to the following attributes.
Related Content

SatNav or SatNag?: A Case Study Analysis of Evolving HCI Issues for In-Car Computing
www.igi-global.com/chapter/satnav-satnag-case-study-analysis/52425?camid=4v1a

Twitter in Higher Education: New Pedagogy in the Knowledge Era of Globalization
www.igi-global.com/chapter/twitter-in-higher-education/96061?camid=4v1a

National Game Management Database of Hungary
www.igi-global.com/chapter/national-game-management-database-hungary/65013?camid=4v1a