

Towards a New Data Replication Management in Cloud Systems

Abdenour Lazeb, Université Oran1, Ahmed Ben Bella, Oran, Algeria

Riad Mokadem, Institut de Recherche en Informatique de Toulouse (IRIT), Paul Sabatier University, Toulouse, France

Ghalem Belalem, Université Oran1, Ahmed Ben Bella, Oran, Algeria

ABSTRACT

Applications produce huge volumes of data that are distributed on remote and heterogeneous sites. This generates problems related to access and sharing data. As a result, managing data in large-scale environments is a real challenge. In this context, large-scale data management systems often use data replication, a well-known technique that treats generated problems by storing multiple copies of data, called replicas, across multiple nodes. Most of the replication strategies in these environments are difficult to adapt to cloud environments. They aim to achieve the best performance of the system without meeting the important objectives of the cloud provider. This article proposes a new dynamic replication strategy. The proposed algorithm significantly improves provider gain without neglecting customer satisfaction.

KEYWORDS

Business Model, Cloud Systems, Cost Model, Data Management, Data Replication, Performance, SLA

1. INTRODUCTION

Cloud Computing is utilized to portray a modern lesson of organizing based computing that takes put over the Web.

Services are offered on-demand, that are continuously on, anytime and anyplace. On the other hand, the tenants are billed for utilizing resources through the 'pay as you go' model. The equipment and software services are accessible to common public, undertakings, organizations and businesses markets. But, what commitments does the cloud provider that you have chosen? How long will it take to restart your solution in case of a problem? Can he lose your data? These are classic questions that we regularly asked when we talk about cloud computing the answer is in the SLA established between a cloud provider and its tenants, i.e., consumers (Zhao et al., 2015). SLA includes the service level objectives (SLO) of the tenant, for example, availability and performance, which must be met by the provider (Limam et al., 2019).

In order to satisfy the SLA, data replication is a well known technique that consists in storing multiple copies of data, called replicas, at multiple nodes. In this context, data replication strategies in cloud systems address classic problems such as: (i) which data to replicate? (ii) when to replicate these data? (iii) Where to replicate these data but also to specific issues of the Cloud environment such as (iv) Determine the number of necessary replicas such as the objectives of the tenant will be satisfied while ensuring a profit for the cloud provider (Mokadem and Hameurlain, 2020)..

Some solutions can be brought to this problem:

DOI: 10.4018/IJSITA.2019040101

- The proposal of a cost model allowing replication only if it is necessary.
- Effective placement of data replicas (Djebbar & Belalem, 2012).
- An elastic management of the number of replicas.
- The proposition of an economic model for the cloud provider, such as information replication is advantageous. It is conditioned by a minimization of the punishments paid by the provider which makes it possible to extend its economic profit (Belalem et al., 2011).

To guarantee failure tolerance, a capacity advertising copies data among different copies. These copies store the same set of information, so in case any of copies is lost, information may still be gotten and recouped from the other replicas.

In this paper, we propose an algorithm that mixes all these solutions for good replication management. As result, the main contribution is to improve provider gain over a wide range of cloud and SLA-conditions without neglecting customer satisfaction.

This paper is organized as follows: Section 2 cites Related work. Section 3 explains our strategy aspects. Positioning our strategy is presented in Section 4. The last section contains the conclusions and future work.

2. RELATED WORK

(Xie et al., 2017) set three threshold parameters for dataset conditions among datasets, get to frequencies of datasets, and the storage capacity of information centers. Dataset reliance among datasets and get to recurrence for each dataset are calculated as limitations of the dataset. They utilize the limit esteem of capacity space to restrain information replication to maintain a strategic distance from flood issues and guarantee full errand completion in the corresponding area. They moreover classify information sorts into three categories, settled dataset, free-flexible dataset and constrained-flexible dataset, to develop a mapping between datasets and each information center. By receiving their methodology, they endeavor to assist diminish information development and information exchange cost. Their work I find it a little expensive compared to ours and does not treat current state each time.

Pankowski (2016) proposes a strategy for reproducing NoSQL information. The calculation is called Lorq. The most highlights of Lorq are (a) information replication is realized by implies of reproducing logs putting away upgrade operations, and so-called pulse operation sent by the pioneer, (b) the preparing and replication methodologies ensure that inevitably all operations in each replica are executed within the same arrange and no operation is misplaced. An uncommon consideration is paid to distinctive sorts of consistency, which can be ensured by the framework. They propose a strategy based on data put away by client administrations to ensure diverse consistency levels, in this manner actualizing SLA usefulness, but replicated data types specification, verification and optimality (Burckhardt et al., 2014) are neglecting.

Cheng et al. (2012) presents an Active/Standby capacity show for information storage in HDFS cluster, bringing in a modern strategy, complex occasion preparing, to recognize information sorts in real-time, displaying the copy situation procedure which is more valuable than the first one, based on these strategy, planning, executing, and assessing an versatile copy administration framework to oversee the copy in HDFS. In addition, this strategy is prepared with the capacity to decide and delete disagreeable copies to spare the capacity.

Xiuguo (2016) went towards accomplishing the least taken a toll copies dispersion benchmark in a down to earth way, they propose a reproductions arrangements procedure show, counting the way to distinguish the need of making reproduction, and plan an calculation for copies situations that can effectively decrease the whole taking a toll within the cloud and proposing information sets administration fetched models, including capacity fetched and exchange taken a toll, showing a novel worldwide information set copies arrangements methodology from cost-effective see named MCRP, which is an inexact minimum-cost arrangement. They proposed a cost-effective information

18 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/towards-a-new-data-replication-management-in-cloud-systems/241865

Related Content

Impact of the Interoperability of ERPs on Information Systems Disintegration

Tarek Samara (2016). *International Journal of Strategic Information Technology and Applications* (pp. 94-109).

www.irma-international.org/article/impact-of-the-interoperability-of-erps-on-information-systems-disintegration/186764

Measuring the Financial Benefits of IT Investments on Coordination

Chiara Francalanci and P. Maggiolini (2002). *Information Systems Evaluation Management* (pp. 54-74).

www.irma-international.org/chapter/measuring-financial-benefits-investments-coordination/23427

Evaluation of Decision-Making Support Systems

Gloria E. Phillips-Wren, Manuel Mora and Guiseppe Forgione (2010). *Strategic Information Systems: Concepts, Methodologies, Tools, and Applications* (pp. 131-140).

www.irma-international.org/chapter/evaluation-decision-making-support-systems/36685

Redefining the Information Technology in the 21st Century

Ruben Xing, Zhongxian Wang and Richard L. Peterson (2011). *International Journal of Strategic Information Technology and Applications* (pp. 1-10).

www.irma-international.org/article/redefining-information-technology-21st-century/52068

Comparison of Baccalaureate Nursing Students' Experience of Video-Assisted Debriefing versus Oral Debriefing following High-Fidelity Human Simulation

Colleen Royle and Kathleen Hargiss (2015). *International Journal of Strategic Information Technology and Applications* (pp. 40-49).

www.irma-international.org/article/comparison-of-baccalaureate-nursing-students-experience-of-video-assisted-debriefing-versus-oral-debriefing-following-high-fidelity-human-simulation/136267