Chapter 7

Data Access Management System in Azure Blob Storage and AWS S3 Multi-Cloud Storage Environments

Yaser Mansouri
Adelaide University, Australia

Rajkumar Buyya
The University of Melbourne, Australia

ABSTRACT

Multi-cloud storage offers better Quality of Service (QoS) such as availability, durability, and users’ perceived latency. The exploitation of price differences across cloud-based storage services is a motivate example of storing data in different Geo-graphically data stores, where data migration is also a choice to achieve more cost optimization. However, this requires migrating data in tolerable time from the perspective of users. This chapter first proposes a comprehensive review on different classes of data stores inspiring data migration within and across data stores. Then, it presents the design of a system prototype spanned across storage services of Amazon Web Services (AWS) and Microsoft Azure employing their RESTful APIs to store, retrieve, delete, and migrate data. Finally, the experimental results show that the data migration can be conducted in a few seconds for data with a magnitude of Megabytes.

INTRODUCTION

Cloud computing has gained significant attention from the academic and industry communities in recent years. It provides the vision that encompasses the movement of computing elements, storage and software delivery away from personal computer and local servers into the next generation computing infrastructure hosted by large companies such as Amazon Web Service (AWS), Microsoft Azure, and Google. Cloud computing has three distinct characteristics that differentiate it from its traditional counterparts: pay-as-
you-go model, on-demand provisioning of infinite resources, and elasticity (Buyya, Shin, Venugopal, Broberg & Brandic, 2009).

Cloud computing offers three types of resources delivery models to users: (i) Infrastructure as a Service (IaaS) which offers computing, network, and storage resources, (ii) Platform as a Service (PaaS) which provides users tools that facilitate the deployment of cloud applications, and (iii) Software as a Service (SaaS) which enables users to run the provider’s software on the cloud infrastructure.

One of the main components of IaaS offering by cloud computing is Storage as Services (StaaS) that provides an elastic, scalable, available, and pay-as-you-go model, which renders it attractive for data outsourcing, both for the users to manipulate data independent of the location and time and for firms to avoid expensive upfront investments of infrastructures. The well-known Cloud Storage Providers (CSPs)—AWS, Microsoft Azure, and Google—offer StaaS for several storage classes which differ in price and performance metrics such as availability, durability, the latency required to retrieve the first byte of data, the minimum time needed to store data in the storage.

The data generated by online social networks, e-commerce, and other data sources is doubling every two years and is expected to augment to a 10-fold increase between 2013 and 2020—from 4.4 ZB to 44 ZB1. The network traffic, generated from these data, from datacenters (DCs) to users and between DCs was 0.7 ZB in 2013 and is predicated to reach 3.48 ZB by 20202. The management of such data in the size of several exabytes or zettabytes requires capital-intensive investment; the deployment of cloud-based data stores (data stores for short) is a promising solution.

Moving the data generated by data-intensive applications into the data stores guarantees users the required performance Service Level Agreement (SLA) to some extent, but it causes concern for monetary cost spent in the storage services. Several factors substantially contribute to the monetary cost. First, monetary cost depends on the size of the data volume that is stored, retrieved, updated, and potentially migrated from one storage class to another one in the same/different data stores. Second, it is subject to the required performance SLA (e.g., availability3, durability, the latency needed to retrieve the first byte of data) as the main distinguishing feature of storage classes. As the performance guarantee is higher, the price of storage classes is more. Third, monetary cost can be affected by the need of data stores to be in a specific geographical location in order to deliver data to users within their specified latency. To alleviate this concern (i.e., monetary cost spending on storage services) from the perspective of application providers/users, it is required to replicate data in an appropriate selection of storage classes offered by different CSPs during the lifetime of the object regarding to the satisfaction of latency for Put (write), Get(read), and data migration from the perspective of users.

The use of multiple CSPs offering several storage classes with different prices and performance metrics respecting their Quality of Service (QoS) in terms of availability and network latency. In this direction, we designed algorithms that take advantage of price differences across CSPs with several storage classes to reduce monetary cost on storage services for time varying workloads (Mnasouri, Y.; Buyya, R., 2016). This mandates data migration across data stores. As a concern for users, it is important to migrate data in tolerable time. This chapter sheds a light on this gap through the following contribution:

- We provide a taxonomy of cloud-based data stores offered by the well-known cloud providers to bring the attention of researchers for future research directions.
- We provide modules using RESTful API of AWS and Microsoft Azure to store, retrieve, delete, and migrate data for AWS’ and Microsoft Azure’ data stores.
Related Content

Security and Privacy Issues in Secure E-Mail Standards and Services
[www.igi-global.com/article/security-privacy-issues-secure-mail/37580?camid=4v1a](www.igi-global.com/article/security-privacy-issues-secure-mail/37580?camid=4v1a)

Security Vulnerabilities, Threats, and Attacks in IoT and Big Data: Challenges and Solutions
[www.igi-global.com/chapter/security-vulnerabilities-threats-and-attacks-in-iot-and-big-data/234809?camid=4v1a](www.igi-global.com/chapter/security-vulnerabilities-threats-and-attacks-in-iot-and-big-data/234809?camid=4v1a)

Firewall
[www.igi-global.com/chapter/firewall/61225?camid=4v1a](www.igi-global.com/chapter/firewall/61225?camid=4v1a)

Ethics in Software Engineering
[www.igi-global.com/chapter/ethics-software-engineering/13483?camid=4v1a](www.igi-global.com/chapter/ethics-software-engineering/13483?camid=4v1a)