Cloud Storage Monitoring System analyzing through File Access Pattern

Guide:
Prof. Sajitha N
Assistant Professor

By:
V R ROHITH
1BG16CS117

Contents

- Introduction

- Problem statement

- Literature survey

- System Architecture

- Result Analysis

- Conclusion

- Reference

Introduction

- The data replication services of cloud storage duplicate the files in real time to increase the availability of the files which in turn increase the hardware cost.The cloud enables to access the same files and applications by multiple users from different locations.
- Replication is used to reach highest availability at high cost. It is degrading the performance of the service when the cost benefits accrued from the replication.

Problem Statement

To propose a Cloud Storage Monitoring System to optimize the Cloud Storage and data availability by analyzing File Access Pattern.

Literature Survey

1. Augustus Devarajan A and SudalaiMuthu T, Cloud Storage Monitoring System analyzing through File Access Pattern, 2019 Second International Conference on Computational Intelligence in Data Science (ICCIDS)

CLOUD STORAGE ARCHITECTURE

- Cloud Client Interface Layer
- Cloud Data Management Layer
- Cloud Storage Layer
- Design Principles of Cloud Storage
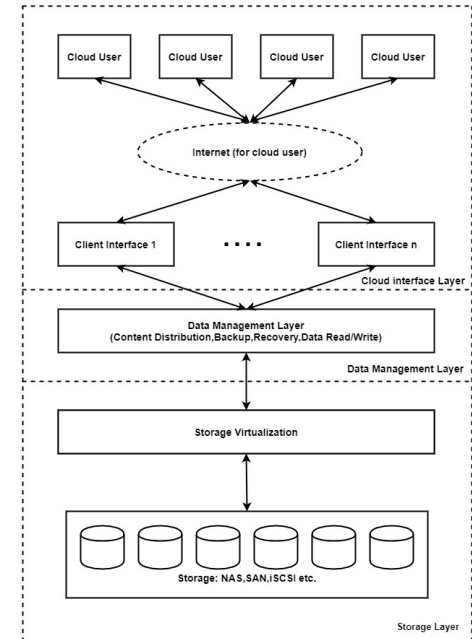
CLOUD STORAGE MONITORING (CSM) SYSTEM :

Cloud Storage Monitoring Architecture :

A prediction and ranking based system is proposed to handle the de-duplication in cloud storage with the following design objectives.
• Identify the frequency on access pattern
• Provide prediction on file access
• Identify the duplication of files on cloud storage
• Build storage efficient system.
• Increase efficiency of the system.
• Block duplication of files in future

Contd.

K-means ranking Algorithm:

Input: k (the number of clusters), D (a set of lift ratios)

Output: a set of k Clusters

Method: Arbitrarily choose k objects from D as the initial cluster centres:

Repeat:

1. Reassign each object to the cluster to which the object is the most similar, based on the mean value of the objects in the cluster;

2. Updated the cluster means, i.e.; calculate the mean value of the objects for each cluster **Until no change**;

Contd.

The de-duplication process are :

A. Comparison based on File attributes.

B. Comparison based on delta version and hashing.

C. Data de-duplication

    1. Data Compression.

    2. Single-Instance Storage.

    3. Data Comparison.

2. S. Ramamoorthy and B. Baranidharan, CloudBC- A Secure Cloud Data access Management system, 2019 IEEE

- The proposed hybrid framework restricts the malicious data modification and deletion among the community cloud environment.
- The random hash values and secure block chain among the cloud user nodes will enhance the security of data access and modification process.
- The security performance also shows that the proposed model effectively restrict the malicious user activity on this platform.

3. Isaac Odun-Ayo and Olasupo Ajayi, An Overview of Data Storage in Cloud Computing , 2017 International Conference on Next Generation and Information Systems(ICNGNS)

- This paper presents the state of the art from some literature available on Cloud storage. The study was executed by means of review of literature available on Cloud storage.
- It examines present trends in the area of Cloud storage and provides a guide for future research.
- The expected result at the end of this review is the identification of trends in Cloud storage, which can beneficial to prospective Cloud researches, users and even providers.

| Service Interface |
| Storage Overlay |
| Metadata Management |
| Storage Management |
| Network and Storage Infrastructure |

Cloud Storage Layered Model

SYSTEM ARCHITECTURE
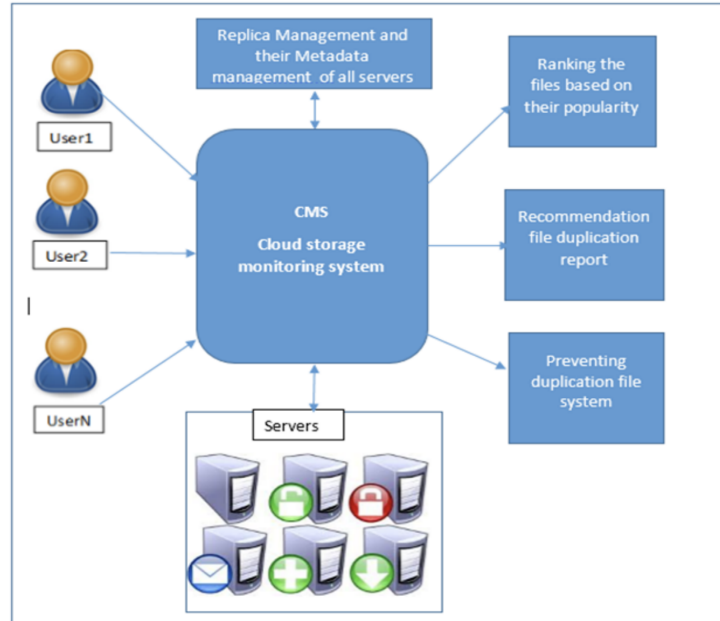
Result Analysis:

- The proposed K-means algorithm allocate ranking for files. It can validate by simulation testing through CloudSim with having samples of five files and 3 storage servers with each 5GB storage. The results are tabulated as shown in Table 3.

| S.No	File Name(.Extn)	File Types	File Size in MB	Servers			No. of Duplication
				Server1 5GB	Server2 5GB	Server3 5GB	
1	Main.java	docx	0.08	0	5	0	5
2	Check_ds	pdf	0.11	0	0	2	2
3	Image_01	jpeg	0.41	0	0	3	3
4	Help_VS	mp3	0.55	3	4	0	7
5	Eng_TST	mp4	1	0	0	2	2
Used Space in GB				0.5	1.4	0.9	
Available Space in GB				4.5	3.6	4.1	

Cont.

- The similar environment is also simulated without using the proposed CSM system as shown in Table 4 and the results are compared as shown in Figure 2.
- The proposed CSM system yield better performance in utilising the storage space using deduplication technique.
- There are five different files with the size of 0.08 MB, 0.11 MB, 0.41 MB, 0.55 MB and 1 MB are used for the experiments.

S.No	File Name(Extn)	File Types	File Size inMB	Servers		
				Server 1 5GB	Server 2 5 GB	Server 3 5 GB
1	Main.java	docx	0.08	0	5	0
2	Check_ds	pdf	0.11	0	0	2
3	Image_01	jpeg	0.41	0	0	3
4	Help_VS	mp3	0.55	3	4	0
5	Eng_TST	mp4	1	0	0	2
Used Space in GB				0.8	1.9	1.4
Available Space in GB				4.2	3.1	3.6

Cont.

- The CSM system has reduced the usage space as 6.66 %, 13.88 %, 12.19 % for sever-1, server-2 and server-3 respectively than the system without using CSM system.

- The deduplication is carried out to reduce the usage of storage space. The average de-duplication is 3.8 GB.

Available Storage Space in GB
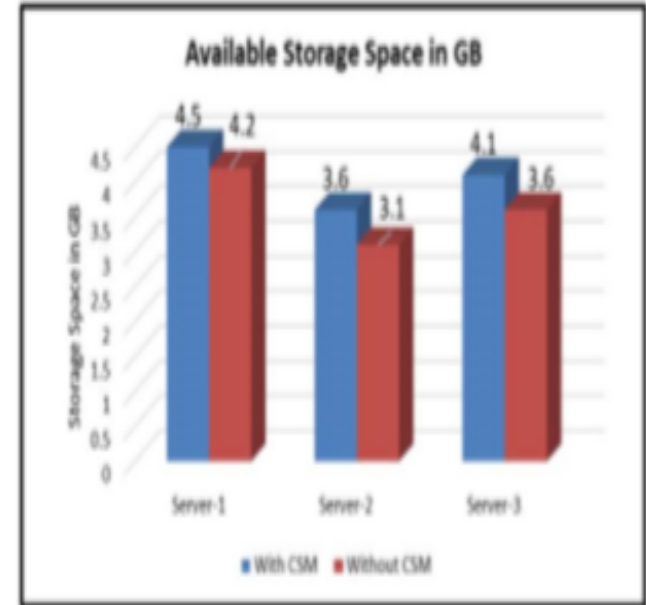
Conclusion

- The Cloud Storage Monitoring (CSM) system is proposed to increase the storage space availability in IaaS-Cloud environment.

- The frequency and popularity of the files are used for ranking. A prediction algorithm evaluates the ranking of files.

- The CSM system has given better performance as average of 10.91% reduction more than "without using CSM" system and yield the average de-duplication as 3.8 GB.

- Thus proposed CSM system provides an efficient data storage mechanism.

References

1. Augustus Devarajan A and SudalaiMuthu T, Cloud Storage Monitoring System analyzing through File Access Pattern, 2019 Second International Conference on Computational Intelligence in Data Science (ICCIDS).

2. S. Ramamoorthy and B.Baranidharan, CloudBC- A Secure Cloud Data access Management system, 2019 IEEE.

3. Isaac Odun-Ayo and Olasupo Ajayi, An Overview of Data Storage in Cloud Computing , 2017 International Conference on Next Generation and Information Systems(ICNGNS).