International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 3 Issue: 3 979 - 980

Survey on Network Coding for Multiple Cloud Storage

Hemanth kumar A
MTECH IV Semester, CSE
SVIT, Bangalore, India
hemanth.ankegowda@gmail.com

Sreelatha P K
Assistant Professor, Department of CSE
SVIT, Bangalore, India
sreelatha.pk@.saividya.ac.in

Abstract: Reliability of data can be provided by maintaining redundancy in multiple cloud storage nodes. The simplest form of providing redundancy is replication, where it involves directly replicating data to multiple clouds. However, replication technique involves high computational cost and monetary cost is high when retrieving data in the multiple cloud environments. Therefore, striping data across multiple cloud storage is one of the best solution for providing redundancy and repair cost is very less when compared to replication. There exists many codes, which performs striping of data, erasure coding is an example where provides good redundancy. However, the erasure codes perform recovery of whole file rather a single block. Therefore new design problem arises when erasure codes are used. However, regenerating codes shows better solution for this problem.

Keywords- *Cloud storage, Erasure coding, Network coding, Repair, Regenerating code*

*****

I. INTRODUCTION

Nowadays, the demand for increasing reliability has been rising with the increase in large data storage, applications such as social networks. However, many cloud service providers are trying to provide reliability to their user's data by redundancy scheme such as replication. Replication [4] involves directly copying of data to 'n' number of storage nodes such that, whenever a cloud failures, data can be retrieved from any other cloud but, cloud service providers are charging users for amount of data they retrieve during a repair operation. Therefore, replication technique seems to be the costlier.

The best method for redundancy is using erasure coding schemes. There are many classes of erasure codes. However, all erasure codes provide double fault tolerance mechanism. Though erasure coding uses striping mechanism, it recovers whole file whenever a block of file is lost. Therefore, new coding scheme called regenerating codes provides significant efficiency compared to erasure coding schemes and replication

The following sections are described as follows. In Section 2,detailed explanation of erasure coding is done, In section 3,network coding explained, In section 4, brief description of various repair operations is done. In section 5, explanation of regenerating codes has done. Section 6, concludes the paper.

II. ERASURE CODING

Erasure coding [1][5] scheme approach maintains redundancy of data across multiple cloud storage nodes. Erasure codes based on reed Solomon codes [6] maintains MDS property. MDS stands for maximum distance separable codes. MDS (n, k) is defined as any 'k' out of 'n 'nodes can be used to reconstruct the lost data. An erasure code provides double fault tolerance, i.e., it is assumed that whenever two clouds fail at a time, data can be retrieved from the surviving clouds. It maintains double fault tolerance i.e., (n-k) for n=4 and k=2. The amount of space occupied by data is very less when erasure coding scheme is applied for multiple cloud storage set up. Therefore tradeoff between redundancy and reliability is very good in erasure codes maintaining MDS property. There exists many classes of erasure codes, some of them are based on XOR operation, and linear combination based codes.

Erasure coding involves storage nodes performing encoding operation whenever repair operation is performed. It is necessary to relax the encoding operation at storage nodes in order to preserve network bandwidth and also erasure codes recovers the complete file which corresponds to the size of the original file rather recovering a block of file. Figure below shows erasure codes based on MDS property.

Figure 1.Each storage nodes stores two blocks of data and each data blocks are linearly combined to form four different linearly combined data blocks. Storage size is M=4 blocks A1, A2, B1, B2. It can be observed that k=2 out of the n=4 can be used recover the file.

III. NETWORK CODING

Network coding [7] can be defined as an generalization of conventional routing where it involves operations such as store and forwarding of data. In contrast to conventional routing, network coding performs encoding operation at intermediate nodes and generates the output data for previously received input data. Therefore network coding mixes information at intermediate nodes. The objective of network coding is significantly reducing network bandwidth and power consumption .Therefore; network coding plays an very role in multiple cloud storage performing repair operation where data from the surviving clouds are transferred to regenerate the lost data.

For multiple cloud storage, network coding was introduced in for wireless sensor networks. Many aspects of coding were explored for networked storage applications. The consideration of repair network bandwidth and the notion of code maintenance over time make the problem very unique. The main objective is to find the fundamental trade off between storage and repair traffic.

IV. REPAIR METHODS

Repair is defined as the amount of outbound data being retrieved from the surviving clouds to regenerate the lost data at a new node.There exists number some of the repair methods or operations in order to reconstruct the lost data.

The repair operations includes exact repair, where the lost data is exactly reconstructed at a new node.The functional repair, where the newly generated data is totally different from the old lost data. However, the newly generated data maintains an MDS property upon iterative operations. Finally, exact repair of systematic parts is also called hybrid model which lies between exact repair and functional repair.

V. REGENERATING CODES

Regenerating codes [3] are new classes of codes which divide the file into blocks and store it in different cloud storage nodes. Regenerating codes are very good at reducing repair traffic; many results show that the regenerating codes provide the double fault tolerance like RAID 6 codes based on reed Solomon codes.

There exists different classes of regenerating codes, some of them are EMSR (exact minimum storage regenerating codes), random linear regenerating codes, FMSR (functional minimum storage regenerating codes). A significant Communication efficiency is provided by regenerating codes same as replication and it also offers storage efficiency of erasure codes.

Regenerating codes shows the same properties as erasure correcting codes in terms of storage and availability. Data availability is high when regenerating codes are applied to storage nodes. As a result, a significant reduction in communication overhead can be observed. Therefore, regenerating codes seems to be more beneficial than the previous coding techniques. Regenerating codes uses network coding concepts and applies it to storage system and defines optimal tradeoffs between the amount of data transferred and stored.

The main objective of regenerating codes is to reduce repair traffic and it is done by striping data redundantly in ‘n’ number of storage nodes.

Regenerating codes uses the concept of network coding, which means that nodes perform encoding operations and send encoded data. During repair, each surviving node encodes its stored data chunks and sends the encoded chunks to a new node, which then regenerates the lost data. However, many classes of regenerating codes relax the need of encoding at storage nodes and it is performed by proxy/gateway server.

VI. CONCLUSION

Regenerating codes have found to be the best method in reducing repair traffic and also it provides very good storage efficiency, communication efficiency in a multiple cloud storage environment. Therefore, both the benefits of erasure codes and replication can be seen in regenerating codes.

REFERENCES

[1] Tutorial: Erasure Coding for Storage Applications web.eecs.utk.edu
[2] http://searchstorage.techtarget.com/definition/erasure-coding
[3] Regenerating Codes for Distributed Storage Networks ...
[4] en.wikipedia.org/wiki/Replication_(computing)
[5] Erasure coding and cloud storage storage Eternity, wikibon 2011
[6] www.cs.cmu.edu/~guyb/.../reedsolomon/reed_solomon_codes.html
[7] www.ifp.illinois.edu/~koetter/NWC