International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 9 (2017) pp. 1815-1818

A Modern Approach in Cloud Computing Storage by Using Compression and Crypto Mechanism

Sheik Saidhbi[#1] and Dr. Komati Thirupathi Rao[#2]

[#1]*Research Scholar, Bharathiar University, Coimbatore, Tamil Nadu, India.*

[#2]*Professor and Associate Dean-Academics, K L University, AP., India.*

[#1]***ORCID: 0000-0001-5334-8981***

Abstract
In order to communicate the compressed secure multimedia data over the internet by using the mechanism of compression and crypto algorithm is a modern and innovative idea in cloud computing. The cloud computing is an inevitable field of data communication and resource sharing in the modern era. The essence of its functioning, its boundaries, the development of new applications, becoming increasingly agile and collaborative, inspiring subjects for research. In the new century, it appears that the ability of data centers is limited and runs out. The economy in sequences the trend of technological development and the solution is the adoption of grid services and/or utility computing as well as the use of virtualization to maximize the available resources. Existing services and applications become more distributed, paradigms like Service-Oriented Architecture emerge in response to integration and service orchestration, and the organization and technologies used in data centers evolve. In this research paper focus on the role of compression and crypto algorithm in cloud data storage mechanism.

Keywords: Cloud, Compression, Storage, Secure and Data.

INTRODUCTION

In general the cloud applications are categorized based On-demand self-service - A consumer can unilaterally provide computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provider. Broad network access and the capabilities are available over the network and accessed through standard mechanisms that promote the use by heterogeneous thin or thick client platforms (e.g., mobile phones, tablets, laptops, and workstations). Resource is pooling and the providers computing resources are combined to serve multiple consumers using a multi-client model, with different physical and virtual resources, dynamically assigned and reassigned according to consumer demand.

There is a sense of location independence since, the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify the location at a higher level of abstraction (e.g., country, state, or data center). Examples of resources include storage, processing, memory, and network bandwidth. Rapid elasticity - Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time.

Measured service and the cloud systems automatically control and optimize resource use, by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g. storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the used service.

Software as a Service (SaaS) - The capability provided to the consumer, ability to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through either a thin client interface, such as a web browser (e.g., web-based email), or a program interface. The consumer does not manage or control the underlying cloud infrastructure, including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user specific application configuration settings.

Platform as a Service (PaaS) - The capability provided to the consumer is to deploy onto the cloud infrastructure, consumer-created or acquired applications created using programming languages, libraries, services, and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly configuration settings for the application-hosting environment.

Infrastructure as a Service (IaaS) - The capability provided to the consumer is the provision of processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over

operating systems, storage, and deployed applications, and possibly limited control of select networking components (e.g., host firewalls). From the figure 1.1 illustrates the principles behind the proposed approach,

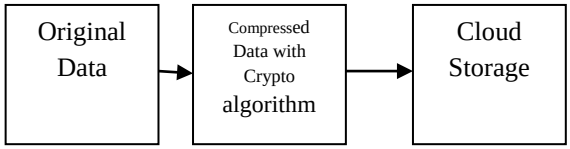

Figure 1.1 Layout for Cloud data storage modern approach

RELATED WORK

Data compression technique is used for saving disk space, reducing the time needed for communication or the time needed for data transfer and more. Data to be handled as well as software has been growing, and the amount of information communicated between systems has also been constantly increasing. In these circumstances, data compression technology is regarded as an important factor to support the information infrastructure. Data compression technology is not only used in PCs, but also in many fields like modems, routers, digital cameras, facsimiles, CDs, MDs(Mini disk), video on demand(VOD), TV conference system, DVDs, Digital telephones and other fields yet the data compression technique generally refer to data stored in 'files' and file transfer over phone lines.

In MS-DOS, used programs like ARC and PKZIP and ubiquitous. UNIX used has the COMPRESS and COMPACT utilities and WINDOWS used have WINZIP utilities. Types Compression: There are two types of data compression, "lossless" data compression and the "lossy" compression. Lossless data compression is used when the data has to be uncompressed exactly as it was before compression. This type of compression used when storing database records, spreadsheets and word processing files. Text files are stored using lossless techniques, since losing a single character can in the worst case, make the text dangerously misleading. Archival storage of master sources for images, videos data and audio data generally needs to be lossless as well. However there are strict limits to the amount of compression that can be obtained with lossless compression. Lossless compression ratio is generally in the range of 2:1 to 8:1.

Lossy compression, in contrast, works on the assumption, that the data does not have to be stored perfectly. Much information can be simply through away from images, video data and audio data, and when decompressed, the data will still be of acceptable quality. Lossy data compression concedes a certain loss of accuracy in exchange for greatly increased compression. Lossy methods can be used on speech and graphics, and they are capable of acting dramatically higher compression ratios. The question of which is a 'better', lossless or lossy technique is pointless. Each has its own uses, with lossless techniques better in some cases and lossy techniques better in others. In fact, lossless and lossy technique are used together to obtain the higher compression ratio.

PROPOSED WORK

In general, cloud data compression /data compression consists of talking a stream of symbols and transforming them into codes. If the compression is effective, the resulting stream of codes will be smaller than the original symbols. The decision is output a certain code for a symbol or set of symbols is based on a model. So, data compression consists of two components, modeling and coding. According to the recent applications of cloud computing, it is the next big thing.

Cloud computing promises reduced cost, flexibility; improve the time to market, higher availability, and more focus on the core business of an organization. Virtually all players of the IT industry are jumping on the cloud computing. The issue that seems to be able to stop cloud computing are security concerns. One of the most prominent issues for cloud computing is privacy; that is, the protection of sensitive data from unauthorized users. Privacy is defined is the way to secure data protection and the control is outsourced to the cloud provider.Coding is frequently used to refer to the entire data compression process instead of just a single component of that process. For example, the phrase "Huffman coding" describes a data compression technique, which actually mean the coding method used in conjunction with a model to compress data. In data compression metaphor, coding would be the wheels, but modeling would be the engine. Information theory uses the term entropy as a measure of how much information is encoded in a message. The higher the entropy of a message, the more information it contain. The entropy of a symbol is defined as the negative logarithm of its probability. To determine the information content of a message in bits, the base 2 logarithm as follows:

Number of bits = - Log base 2 (probability)

The entropy of an entire message is simply the sum of the entropy of all individual symbols. For example, if the probability of the character 'e' is 1/16, the information content of the character is four bits. So the character string 'eeeee' has a total content of 20 bits. If this string is encoded in the standard 8-bit ASCII, 40 bits are required to encode this message. Encoding characters using EBCDIC or ASCII require every character to be encoded using the same number of bits. To compress data, we need to encode symbols with exactly the number of bits of information the symbol contains. If the character 'e' only gives us four bits of information, then it should be coded with twelve bits. But this will introduce lots of error, with most of the codes in a message being too long and some being too short. This coding problem can be solved by Shannon-Fano coding and Huffman coding two different ways of generating variable length codes when gives a probability table for a given set of symbols.

The Encryption is a possible way to fulfill the confidentiality requirement of data that is being stored in an untrusted cloud database. The request data are used to retrieve data from a cloud database. After the completion of the framework, we proceeded to tests in order to validate experimentally the suitable operation of the framework and to evaluate different parameters associated with the compression, decompression, encryption and decryption of multimedia contents. As may be seen in this table, the obtained results show that compression followed by encryption may be effective for efficient and secure storage of multimedia content. On the other hand, if we apply first encryption followed by compression, the compression rate is very small due to the higher degree of randomness introduced by encryption. These results confirm the unsuitability of encryption followed by compression for efficient multimedia storage. The properties of an encryption function determine which kinds of method can be applied on the encrypted data without decrypting them and, consequently, how it will be stored in to the cloud service provider. The interface for the cloud data access is illustrated by the figure 1.2.

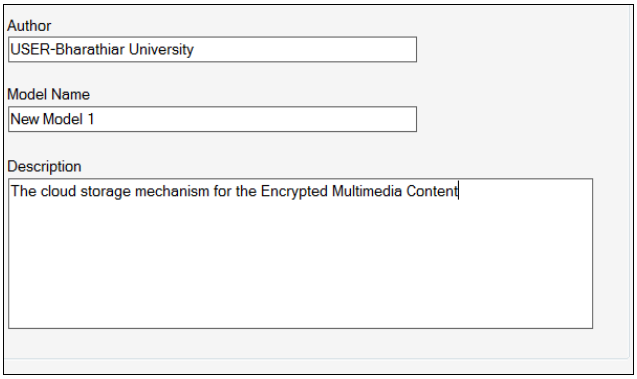

Figure 1.2 Interfaces

The data for the transmission is used for the encryption by using any one of the crypto algorithm. A deterministic encryption scheme (as opposed to a probabilistic encryption scheme) is a cryptosystem which always produces the same ciphertext for a given plaintext and key. The property of indistinguishability under chosen plaintext attack is considered a basic requirement for most provably secure public key crypto systems. The Cloud Encryption Component is shown by the following figure 1.3.

The complexity in virtual machine environments can also be more challenging than their traditional counterparts, giving rise to conditions that undermine security. For example, paging, check pointing, and migration of virtual machines can leak sensitive data to persistent storage, subverting protection mechanisms in the hosted operating system intended to prevent such occurrences. Moreover, the hypervisor itself can potentially be compromised. For instance, a vulnerability that allowed specially crafted File Transfer Protocol (FTP) requests to corrupt a heap buffer in the hypervisor, which could allow the execution of arbitrary code at the host, was discovered in a widely used virtualization software product, in a routine for Network Address Translation (NAT).

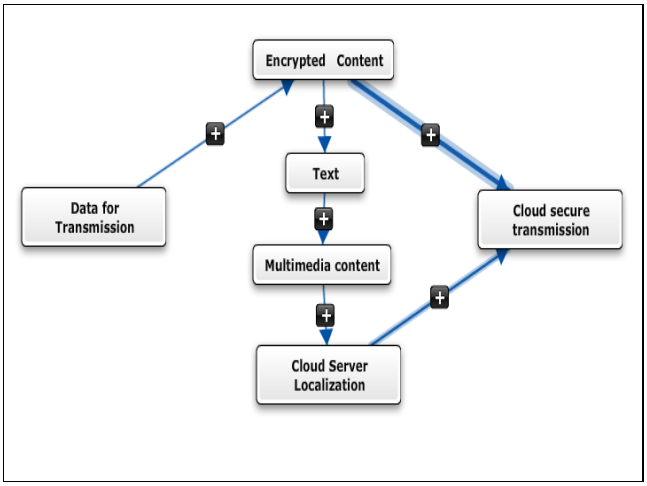

Figure 1.3 Cloud Encryption Components

Most virtualization platforms have the ability to create software-based switches and network configurations as part of the virtual environment to allow virtual machines on the same host to communicate more directly and efficiently. For example, for virtual machines requiring no external network access, the virtual networking architectures of most virtualization software products support same-host networking, in which a private subnet is created for intra-host communications. Traffic over virtual networks may not be visible to security protection devices on the physical network, such as network-based intrusion detection and prevention systems

CONCLUSION

Determining the security of complex computer systems composed together is also a long-standing security issue that plagues large-scale computing in general, and cloud computing in particular. Attaining high-assurance qualities in implementations has been an elusive goal of computer security researchers and practitioners and, as demonstrated in the examples given in this report, is also a work in progress for cloud computing. Nevertheless, public cloud computing is a compelling computing paradigm that agencies need to incorporate as part their information technology solution set.

REFERENCES

[1] Yared Keleta, J.H.P. Eloff, H.S. Venter, Proposing a Secure XACML Architecture Ensuring Privacy and Trust, Research in Progress Paper, University of Pretoria, 2005.
[2] Brian Krebs, Amazon: Hey Spammers, Get Off My Cloud! The Washington Post, July 1, 2008.
[3] Eileen Kowalski et al., Insider Threat Study: Illicit Cyber Activity in the Government Sector, U.S. Secret Service and Carnegie Mellon University, Software Engineering Institute, January 2008.

[4] Michael Krigsma, Amazon S3 Web Services Down. Bad, Bad News for Customers, ZDNET, February 15, 2008.
[5] Sushil Kumar, Oracle Database Backup in the Cloud, White Paper, Oracle Corporation, September 2008. [Lab95] Stephen Labaton, 2 Men Held in Attempt to Bomb I.R.S. Office, New York Times, December 29, 1995.
[6] 20-Year Term in Plot to Bomb IRS Offices, Nation In Brief, Los Angeles Times, August 10, 1996.
[7] Neal Leavitt. Is Cloud Computing Really Ready for Prime Time?, IEEE Computer, January 2009.
[8] James Maguire, How Cloud Computing Security Resembles the Financial Meltdown, Datamation, internet.com, April 27, 2010.