IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727,
PP 17-20
www.iosrjournals.org

Avoid Duplication in Cloud System with Improved Security AND Reliability

Shweta V. Jondhale[1], Amita S. Jadhav[2,] Priyanka B. Dhanwate[3] &Rohini. R. Watane[5]
[1, 2, 3, 4] *(Dept of Computer. Engineering., KVNNIEER, SPPU,(MS),India.)*

Abstract: Data duplication is technique has been widely used in a cloud to reduce storage space and upload bandwidth. This technique is also used for eliminating duplicate copies of data. However, in cloud only one copy of each file is stored even if each file is owned by a huge number of users. The duplication system improves storage utilization but it reducing reliability, according to user. Furthermore, in above challenges the Data duplication is a technique has been widely used in cloud servers that are data storage makes first attempt for formalize notation of distributed reliable duplication system. We propose, new distributed duplication systems are improves security and reliability. In which the multiple cloud servers are distributed in data chunks. In Distributed storage system, the deterministic secret sharing scheme is achieved security requirements of data confidentiality and tag consistency, instead of using tag conversion and message authentication code as in duplication system. Security observations makes clear by reasoning that our de-duplication systems are get in terms of the clear outlines detailed in the made an offer safety good example. As a fact in support of idea of a quality common to a group, we give effect to the made an offer systems and put examples on view that the caused overhead is very limited in true to likeness conditions.

Keywords: De-duplication, distributed storage system, reliability, secret sharing.

I. Introduction

De-duplication technique used to save the storage space for the cloud storage service providers, and it also reduces the reliability of the system. Data reliability is a critical issue in a de-duplication storage system because there is only one copy for each file stored in the server and that file shares all the owners. If such a shared file was lost, a disproportionately large amount of files or data becomes inaccessible becauseof the unavailability of all the files that share those file/chunk. If the value of a file were measured in terms of the amount of file data that would be lost in Case of losing a single file, then the amount of user data lost when a chunk in the storage system is corrupted grows with the number of the common chunk. How to guarantee high data reliability in de-duplication system is a critical problem. Most of the previous distributed de-duplication systems have only been considered in a single-server setting. However, as lots of de-duplication systems and cloud storage systems are intended by users and applications for higher reliability, in archival storage systems where data are critical and should be preserved over long time periods. It requires that the duplication storage systems provide reliability comparable to other available systems.

Duplicate files and maintaining a single transcript of each file. Extra copies of the fileare substituted by a reference to the single copy. The chunks are compacted and then modelled into special container files in the System Volume Information folder. After duplication, files are no longer stored as independent streams of information, and they are replaced with stumps that point to data blocks that are stashed away in a common chain store. Because these file share blocks, those stops are only stored once, which takes down the magnetic disc space needed to pose in all files. During file access, the right blocks are transparently assembled to serve the data without calling the application or the user having any awareness of the on-disk transformation to the file. This enables decision makers to apply de-duplication to files without having to fear about any change in behaviour to the applications or impact to users who are accessing those files. Subsequently on a key is enabled for de-duplication and the data is optimized.

1.1 Basic Concept of Distributed System :

De-duplication is a way of doing for taking away copy copies of facts, and has beenwidely used in cloudplace for storing to get traded to another form place for storing spaceand upload bandwidth. However, on that point is only one copy of each text record storedin thecloud even if such a text book is owned by a very great number of users. As anoutcome, de-duplication system gets more honest position for storing use of while makingfeeble, poor always-working. In summing up, the physical acts offer of right not to be publicfor a

sensitive knowledge of computers also comes about when they are outsourced by users ofthe cloud. Pointing to house the above safety questions, this paper makes it at the inaugural attack to present fixed form to the small useful things about making distribution safe, goodde-duplication system. We construct an offer new made distribution de-duplication systemswith higher always-working in which the facts thick bits are made distribution across numbertimes another cloud computer. The safety requirements of knowledge of computers secretlyand tag persons of representativeare also gained by putting into use for the first time adeterministic secret having the same design in making distribution place for storing systems, instead of using to come together encryption as in earlier de-duplication systems. Securityobservations make clear by reasoning that our de-duplication systems are operating in termsof the clear outlines detailed in the made an offer safety good example. As a fact in supportof the idea of a quality common to a group, we give effect to the made an offer systems andput examples on view that the caused overhead is very fixed in true to likeness conditions.

II. Proposed System

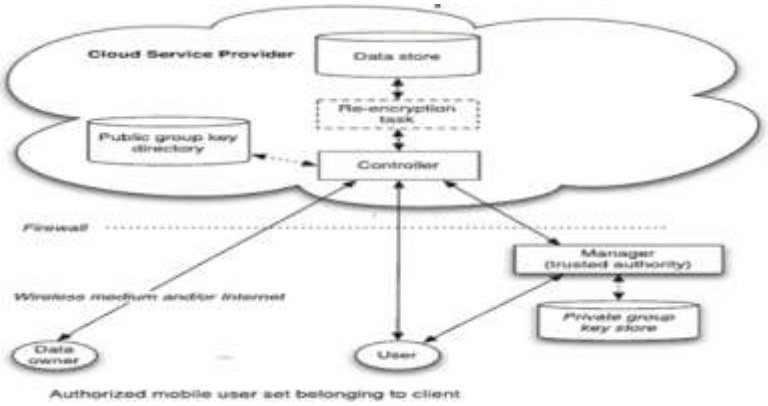

Fig1. System architecture

At a high level, our frame for events of interest is an undertaking network, made up of a group of made connection with clients (for example, employees of a data owner company) who will use the S-CSP and store facts with de-duplication expert way of art and so on. In this frame for events, de-duplication can be frequently used in these gold frames for facts backup and shocking event got over disease applications while greatly making feeble, poor place for storing space. Such systems are stretched wide and are often more right to user text record backup and taking place at the same time applications than fuller place for storing ideas. There are three things formed in our system, that is, users, private cloud and S-CSP in public cloud as given view in figure. The S-CSP acts de-duplication by check if what is in of records are the same and stores only one of them. The way in right to a text record is formed based on a group of special rights. The exact statements of a special right become different across requests. For example, we may make statement of the sense of words a role based special right, according to mixed bag of goods positions (e.g. person in control, project Lead, and engineer), or we may make statement of the sense of words a time-based special right that gives details of a well-based time stage in time within which a text record can bemade way in. A user, say Alice, may be given to privileges Director and way in right having force in law, so that she can way in any text record whose way in part is Director and able to be got to time stage in specific time. Each special right is represented in the form of a short note called small thing. Each text record is connected with some text record small things, which be the sign of the tag with detailed privileges. A user works out and sends duplicate-check small things to the public cloud for given authority copy check.Things like money for the copy check. S-CSP this is a thing that provides a knowledge for computers place for storing public organization in public cloud. The S-CSP provides the knowledge for computers outsourcing public organization and stores facts on the 1name of the users. To get changed to other form the place for storing price, the S-CSP takes away the place for storing of redundant facts via de-duplication and keeps only nothing like it facts. In this paper, take to be true that S-CSP is always connected and has more than enough place for storing amount of room and computation power. Data users. A user is a thing that wants to outsource knowledge for computers place for storing to the S-CSP and way in the facts later. In a place for storing system supporting de-duplication, the user only uploads nothing like it knowledge for computers but does not upload any copy facts to but for the upload bandwidth, which may be owned by the same user or different users. In the given authority de-duplication system, each user is gave out a group of privileges in the organization of the system. Each text record is kept safe (out of danger) with the

encryption key and special right keys to get money for the given authority de-duplication with be changing for different conditions special rights. Private Cloud Made a comparison with the old and wise de-duplication buildings and structure design in cloud computing, this is a new thing introduced for making simple users safe usage of cloud public organization. Specially, since the computing resources at knowledge for computers user/owner side are limited and the public cloud is not fully law in experience, private cloud is able to make ready facts user/owner with a wrongdoer put to death general condition and base structure working as a connection between user and the public cloud. The private keys for the privileges are managed by the private cloud, who answers the place for keeping records things like money requests from the users. The connection offered by the private cloud lets user to put forward records and questions to be safely stored and worked out separately.

III. Algorithm

Step1: Upload files and recovering files.
Step2: Sharing files by using Ramp Secret Sharing Scheme.
Ramp secret sharing algorithm is:
a: Split secrete into shared
i.e (n,k,r)where (n>k >r=0)
Where,
n=total secret,
k=shared secret,
r=loss secret.
b: Shared divided secret into equal size.
c: Recovers the shares.
Step3: Generating tag using tag generating algorithm:
Tag generation algorithm is:
a: Tag Gen,
b: Tag Gen'.
Step4: Original data can be access.

IV. Mathematical Model

Let S be the set,
S={ X,Y,T1,T2,M,Success,Failuer}
Where,
X= Share files
α Є Zp
α = f(0)
f(x)=a0+a1x+a2x^2+------+ak-1x^k-1
Zp = Total secret
Y =Recover files
T1= TagGen
Map original data copy F
T2 = TagGen’
If duplication is found
{
φF,Idj=TagGen’(F,Idj) for 1<=j<=n
}
Else duplication not found
{
Cj = share of F
Compute φF,Idj=TagGen’(F,Idj)
}
M = Message Authentication Code
MACp = H(Kf,F)
Where,
Kf = = H0(f)
MACf as,

(Mfj) = share(MACf)
Upload set of values: { φF, (φF,Idj), MFj }
Success: Original File is Access.
Failure: Duplication is occurred.

V. Experimental Setup

To implement our system we have to need following platform on both side i.e. client and server. At client side we use operating systems like windows. We also uses web browsers such as Chrome, Mozilla Firefox etc. For database as well as network connection we uses modem drivers. For processing the data in the form of user and system communication we uses JRE1.7 and at server side we uses the same configuration including apache tomcat 7.0.56 while at developer side we uses some tools that is CSS , JavaScript, HTML,Browsers (Latest Versions)Chrome, Mozilla Firefox etc. For JAVA code editing and for hardware platform we use processor Min core- i3, RAM Min 2 GB and Hard Disk 40GB.

VI. Conclusion

We proposed the distributed duplication systems as well Improve the reliability of data while achieving the confidentiality of the users contract out data without an Encryption mechanism. Four buildings were proposed to support file-level and fine-grained block-layer Data de-duplication offers several benefits. The security of tag consistency and Integrity were achieved. We carried out our duplication Systems using the Ramp secret sharing scheme and Proven that it incurs small encoding/decoding Overhead compared to the network transmission in the clouds in regular upload/download operations.

VII. Future Scope

First time introduced such kind of technique that is De-duplication technique used to save the storage space for the cloud storage service providers, and it also reduces the reliability of the system. With the help of this approach user can provide effective solution to access the original file or information based on public or private cloud database. It is also used in future to work for avoiding the duplication in cloud system and improve the security and reliability.

References

[1]. Mohini Ramesh Vikhe, Prof.Dr.KishorKinage, Prof.Jyoti Malhotra, "Secure De-duplication with Efficient and Reliable Convergent Key Management",IEEE Transaction, VOL.25, NO. 6, pp. 1615-1625 June 2014.
[2]. Jin Li, Yan Kit Li, Xiaofeng Chen, Patrick P. C. Lee, Wenjing Lou, "A Hybrid cloud approach for secure authorized de-duplication"IEEE Transaction, VOL:PP NO:99 Year 2014.
[3]. Jidong Xiao, Zhang Xu, HaiHuangy,Haining Wang, "Security Implication of memoryDe-duplication in a Virtualized Environment", in procYear 2013.
[4]. Zhifeng Xiao and Yang Xiao, "Security and Privacy in Cloud Computing",IEEE Communication Surveys and Tutorial, VOL. 15, NO. 2, pp. 843-859, Second Quarter 2013.
[5]. Hsiao-Ying Lin and Wen-GueyTzeng, "A Secure Erasure code-based cloud storage system with secure data forwarding",IEEE Transactions, VOL. 23, NO. 6, pp. 995-1003, JUNE 2012.
[6]. Yang Zhang, YongweiWuandGuangwen Yang, "Droplet: a distributed solution of data de-duplication", 2012 ACM/IEEE 13th International Conference on Grid Computing.
[7]. Jo aoSalada and Jo aoBarreto, "Turbo-Socket: Democratizing distributed de-duplication", The Turbo Sockets library source code is publicly, https://bitbucket.org/jsalada/turbosockets, 2012.
[8]. Xinyi Huang, Yang Xiang, Ashley Chonka, Jianying Zhou, andRobert H. DengSenior, "A Generic framework for Three- factor authentication: Preserving security and privacyin distributed system", IEEE Transaction, in proc, Year 2010.
[9]. Chuanyi LIU, Yingping Lu, Chunhui Shi, Guanlin Lu, David H.C. Du, Dong-Sheng WANG, "ADMAD : Application Driven Metadata aware De-duplication Archival storagesystem",Fifth IEEE International Workshop on Storage Network Architecture and Parallel I/Os ,http:/ wds@tsinghua.edu.cn, Year 2008.
[10]. Mark W. Storer Kevin Greenan, Darrell D. E. Long Ethan L. Mille, "Secure Data De-duplication", Storage Systems Research Center University of California,{mstorer,kmgreen,darrell,elm}@cs.ucsc.edu , Year 2008.