Security in Cloud IaaS using Encryption and Deduplication

Mr. Sanket P. Farkade1 , Prof. Pragati Patil2
M.Tech, Computer Science and Engineering
AbhaGaikwad-Patil College of Engineering & Technology, Nagpur

Abstract:- With the insecure improvement of mechanized data, de-duplication techniques are by and large used to fortification data and limit framework and limit overhead by perceiving and taking out overabundance among data. Instead of keeping different data copies with a similar substance, de-duplication takes out dull data by keeping emerge physical copy and suggesting different abundance data to that copy. De-duplication has become much thought from both the insightful world and industry in light of the way that it can altogether upgrades stockpiling use and extra storage space, especially for the applications with high de-duplication extent, for instance, recorded limit systems. Different de-duplication structures have been proposed considering distinctive de-duplication strategies, for instance, client side or server-side de-duplications, record level or square level de-duplications. Especially, with the approach of conveyed stockpiling, data de-duplication frameworks end up being all the more appealing and segregating for the organization of continually growing volumes of data in dispersed stockpiling organizations which motivates attempts and relationship to outsource data stockpiling.

Keywords*: Deduplication, Cloud storage system, reliability, secret sharing*

*****

Introduction

Cloud computing is Web based improvement and utilization of PC innovation. It is a model for empowering helpful, on-request organize access to a mutual pool of configurable figuring assets. In idea, it is a model move whereby points of interest are disconnected from the clients who no longer responsible for the innovation foundation "in the cloud" that backings them. The term cloud is utilized as an image for the Web. It is a style of processing in which as opposed to keeping information all alone hard drive or refreshing applications for your necessities, you utilize an administration over the web at other area which is overseen by the outsider. Average Cloud computing administrations give normal business applications online that are gotten to from a web program, while the product and information are put away on the servers over the Web on a compensation for-utilize premise. Every one of the expenses related with setting up a server farm, for example, getting a building, equipment, repetitive power supply ,cooling frameworks, redesigning electrical supply, and keeping up a different Fiasco Recuperation site can be passed on to an outsider merchant. Since the client is charged just for PC administrations utilized, Cloud computing expenses are a small amount of customary innovation uses.

Cloud give diverse sorts of organization model, for example, open cloud, group cloud, private cloud, half and half cloud. Every one of them have diverse properties and the client can utilize any of them as indicated by their prerequisite. Cloud additionally gives distinctive sorts of administrations to clients. These administrations are comprehensively isolated into three classifications: Framework as a Service(IAAS), Stage as a Service(PAAS), and Programming as an Administration (SAAS).

Building improvement and its determination are two separating successful factors for any business/affiliation. Cloud computing is a late development perfect model that enables affiliations or individuals to give diverse organizations in a steady and viable way. Cloud computing shows an open door for inescapable systems to control computational and stockpiling advantages for accomplish assignments that would not ordinarily be possible on such resource obliged devices.

1.1 Deployment Models

Deploying cloud computing can differ depending on requirements. There are four different deployment models, each with specific characteristics that support the needs of the services and users of the clouds in particular ways.

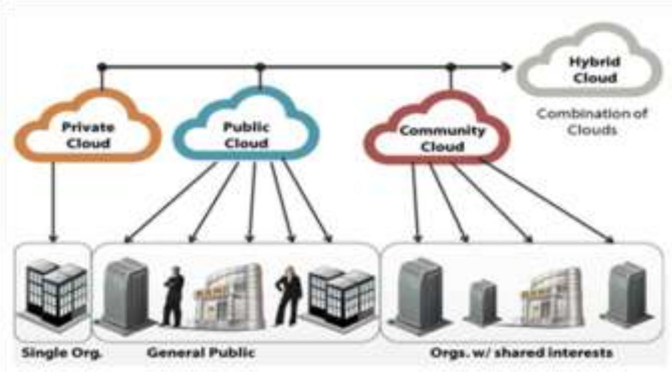


Figure 1.1: Development Models of Cloud

- Private Cloud : The cloud infrastructure has been deployed and is maintained and operated for a specific organization. The operation may be in-house or with a third party on the premises.

- Community Cloud : The cloud infrastructure is shared among a number of organizations with similar interests and requirements. This may help limit the capital expenditure costs for its establishment as the costs are shared among the organizations. The operation may be in-house or with a third party on the premises.
- Public Cloud : The cloud infrastructure is available to the public on a commercial basis by a cloud service provider. This enables a consumer to develop and deploy a service in the cloud with very little financial outlay compared to the capital expenditure requirements normally associated with other deployment options.
- Hybrid Cloud : The cloud infrastructure consists of a number of clouds of any type, but the clouds have the ability through their interfaces to allow data and applications to be moved from one cloud to another. This can be a combination of private and public clouds that support the requirement to retain some data in an organization, and also the need to offer services in the cloud.

1.2 Service Models

Once a cloud is established, use of cloud computing services in terms of business models can differ depending on requirements. The primary service models being deployed are of three types. Each of service provides different properties and are used according to user requirements.

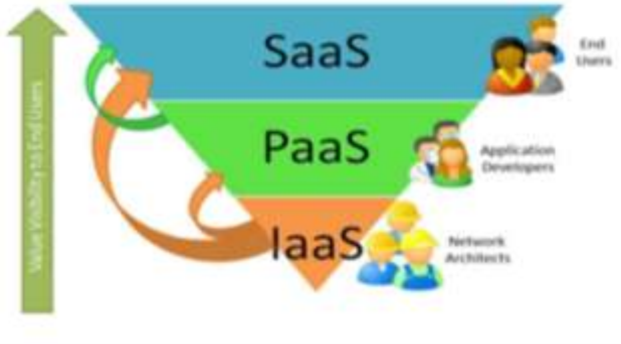


Figure 1.2: Service Models of Cloud

Software as an Administration (SAAS) : In this model, purchasers can get to and utilize an application or administration that is facilitated in the cloud. Cloud suppliers introduce and work application programming in the cloud and cloud clients get to the product from cloud customer. This take out the need to introduce and run the applications on the clients possess PC which improves upkeep and support of the product. Microsoft is extending its association around there, and as a component of the Cloud computing alternative for Microsoft Office 2010, its Office Web Applications are accessible to Office volume permitting clients and Office Web Application memberships through its cloud-based Online Administrations.

Platform as an Administration (PAAS) : In this model, customers has admittance to the stages, permitting them to introduce their own particular programming and applications in the cloud. The working frameworks and system get to are not overseen by the shopper. The cloud supplier conveys a processing stage i.e. OS, database, web server and so on. Application engineers can create and run their product arrangement on a cloud stage without the cost and unpredictability of purchasing and dealing with the fundamental equipment and programming layers.

Infrastructure as an Administration (IAAS) : It is a type of Cloud computing that gives virtualized registering assets over the Web. It offers very versatile assets that can be balanced on-request. A third part supplier has equipment, programming, servers, stockpiling and other framework segments in the interest of its clients. IaaS clients pay on a for each utilization premise, normally by the hour, week, or month.

Proposed System

In proposed system single file is never stored at a single place nor is duplicated at various database locations.

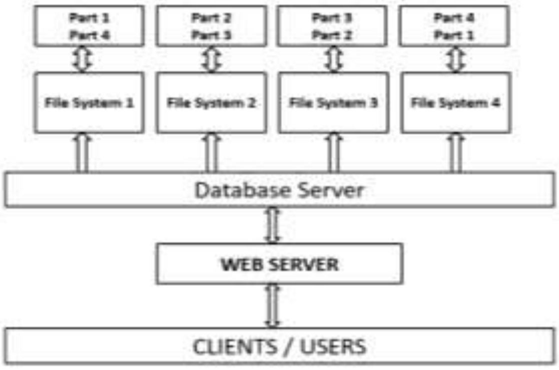
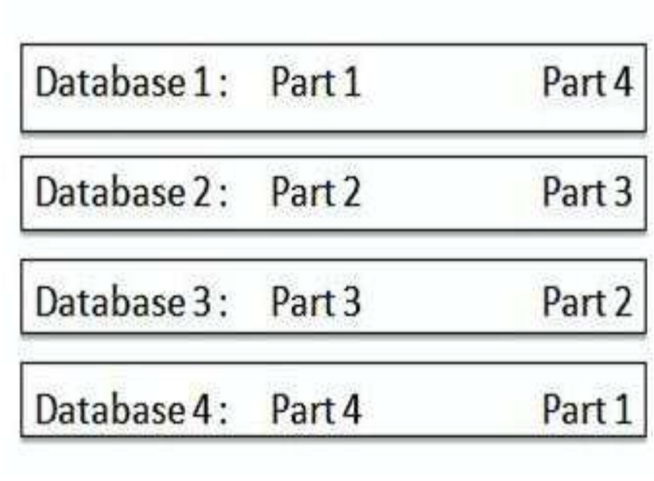


Figure 3.1: Proposed Architecture

Each file is split twice and properly inserted into different file systems. We have considered 4 databases to be used.

The file split sequence is shown as follows:



Database 1:	Part 1	Part 4
Database 2:	Part 2	Part 3
Database 3:	Part 3	Part 2
Database 4:	Part 4	Part 1

Figure 3.2: 4 Database for Deduplication

MODULES

4.1 GUI Designing

A graphical user interface or GUI is a type of interface that allows users to interact with electronic devices through graphical icons and visual indicators such as secondary notation, as opposed to text-based interfaces, typed command labels or text navigation. GUIs were introduced in reaction to the perceived steep learning curve of command-line interfaces (CLIs)which require commands to be typed on the keyboard.

The actions in a GUI are usually performed through direct manipulation of the graphical elements.In addition to computers, GUIs can be found in hand-held devices such as MP3 players, portable media players, gaming devices, smartphones and smaller household, office and industrial equipment. The term GUI tends not to be applied to other low-resolution types of interfaces with display resolutions, such as video games (where HUD is preferred), or not restricted to flat screens, like volumetric displays because the term is restricted to the scope of two-dimensional display screens able to describe generic information, in the tradition of the computer science research at the PARC (Palo Alto Research Center).

4.2 Database Designing

Database design is the process of producing a detailed data model of a database. This data model contains all the needed logical and physical design choices and physical storage parameters needed to generate a design in a data definition language, which can then be used to create a database. A fully attributed data model contains detailed attributes for each entity.

The term database design can be used to describe many different parts of the design of an overall database system. Principally, and most correctly, it can be thought of as the logical design of the base data structures used to store the data. In the relational model these are the tables and views. In an object database the entities and relationships map directly to object classes and named relationships. However, the term database design could also be used to apply to the overall process of designing, not just the base data structures, but also the forms and queries used as part of the overall database application within the database management system (DBMS).

The process of doing database design generally consists of a number of steps which will be carried out by the database designer.

4.3 Connecting Website to Cloud Database

This section provides a sample script that creates a very simple webpage. You can use this webpage to test that your MySQL database is working. You can also use it as a very simple calculator. You copy the script and paste it into a text editor. Then you modify the script with your own hostname, user name, password, and database instance name information and save the changes. Finally, you copy the script to your cloud server and execute the script to display the simple webpage and test your connection to your database instance.

Your web server must be in the same region as your database instance.

4.4 File Encryption and Splitting

If the file contains sensitive information, you can encrypt the file while compressing it. Option -e encrypts the file with the given password, and the receiver should know this password for decrypting it. If the file size exceeds the specified limit after compressing also, then split the files

4.5 Removing Duplications and Testing

Removing duplication means repeated data should be deleted so that this space will be made available for another purpose. so the less space will require and another task can be perform with that space and after that twisting is done.

Conclusion

Cloud computing has gone to an advancement that leads it into a valuable stage. This suggests most of the principal issues with conveyed processing have been tended to a degree that fogs have been able to be fascinating for full business abuse. This however does not suggest that each one of the issues recorded above have truly been grasped, recently that the concurring threats can be persisted to a beyond any doubt degree. Cloud computing is in this way still as much an examination subject, as it is a business part promoting. For better mystery and security in circulated processing we have proposed new de-duplication advancements supporting affirmed duplicate check in cross breed cloud auxiliary arranging, in which the duplicate check tokens of records are made by the private cloud server with private keys. Proposed structure consolidates confirmation of data proprietor so it will complete better security issues in appropriated registering.

REFERENCES

[1] 1.J. Li, X. Chen, M. Li, J. Li, P. Lee, and W. Lou, "Secure deduplication with efficient and reliable convergent key management," in *IEEE Transactions on Parallel and Cloud Systems*, 2014, pp. vol. 25(6), pp. 1615–1625.

[2] 2.M. Li, C. Qin, P. P. C. Lee, and J. Li, "Convergent dispersal: Toward storage-efficient security in a cloud-of-clouds," in *The 6th USENIX Workshop on Hot Topics in Storage and File Systems*, 2014.

[3] 3.M. Bellare, S. Keelveedhi, and T. Ristenpart, "Dupless: Serveraided encryption for deduplicated storage," in *USENIX SecuritySymposium*, 2013.

[4] J. Xu, E.-C. Chang, and J. Zhou, "Weak leakage-resilient client-side duplication of encrypted data in cloud storage," in *ASIACCS*, 2013, pp. 195–206.

[5] D. Harnik, B. Pinkas, and A. Shulman-Peleg, "Side channels in cloud services: Deduplication in cloud storage." *IEEE Security & Privacy*, vol. 8, no. 6, pp. 40–47, 2010.

[6] A. Rahumed, H. C. H. Chen, Y. Tang, P. P. C. Lee, and J. C. S. Lui, "A secure cloud backup system with assured deletion and version control," in *3rd International Workshop on Security in Cloud Computing*, 2011.

[7] A. Rahumed, H. C. H. Chen, Y. Tang, P. P. C. Lee, and J. C. S. Lui, "A secure cloud backup system with assured deletion and version control," in 3rd International Workshop on Security in Cloud Computing, 2011.

[8] S. Halevi, D. Harnik, B. Pinkas, and A. Shulman-Peleg, "Proofs of ownership in remote storage systems." in ACM Conference on Computer and Communications Security, &. Cheng. Danezis, and V. Shmatikov, Eds. ACM, 2011, pp. 491–500.

[9] D. Harnik, B. Pinkas, and A. Shulman-Peleg, "Side channels in cloud services: Deduplication in cloud storage." IEEE Security & Privacy, vol. 8, no. 6, pp. 40–47, 2010.

[10] P. Anderson and L. Zhang, "Fast and secure laptop backups with encrypted de-duplication," in Proc. of USENIX LISA, 2010.

[11] H. Shacham and B. Waters, "Compact proofs of retrievability," in ASIACRYPT, 2008, pp. 90–107.

[12] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable data possession at untrusted stores," in Proceedings of the 14th ACM conference on Computer and communications security, ser. CCS '07. New York, NY, USA:

[13] Z. Wilcox-O'Hearn and B. Warner, "Tahoe: the least-authority filesystem," in Proc. of ACM StorageSS, 2008.

[14] J. S. Plank, S. Simmerman, and C. D. Schuman, "Jerasure: A library in C/C++ facilitating erasure coding for storage applications - Version 1.2," University of Tennessee, Tech. Rep. CS-08-627, August 2008

[15] J. R. Douceur, A. Adya, W. J. Bolosky, D. Simon, and M. Theimer, "Reclaiming space from duplicate files in a serverlessCloud file system." in ICDCS, 2002, pp. 617–624.