International Journal On Engineering Technology and Sciences – IJETS™
ISSN (P): 2349-3968, ISSN (O): 2349-3976
Volume 1 Issue 5, September 2014

DELTA ENCODING FOR DEDUPLICATION OF HYBRID CLOUDS.

Dr.C.Kumar Charliepaul [1]
Principal
A.S.L Pauls College of Engg & Tech,
Coimbatore .
charliepaul1970@gmail.com

Felcia Jerlin [2]
Assistant professor / CSE
Dr.G.U POPE College of Engineering,
Tuticorin
jerlinfelcia@gmail.com

Abstract— *Cloud computing is internet-based computing within which giant teams of remote servers area unit networked to permit the centralized knowledge storage, and on-line access to laptop services or resources. Cloud computing depends on sharing of resources to realize coherence and economies of scale, like a utility over a network. At the muse of cloud computing is that the broader thought of converged infrastructure and shared services. In each enterprise moving databases to the cloud, there is a lot of knowledge within the cloud and this may increase exponentially. However because the knowledge will increase the management of knowledge becomes a troublesome task at hand. to create knowledge management scalable in cloud computing, deduplication has been a widely known technique. Knowledge deduplication could be a specialised knowledge compression technique for eliminating duplicate copies of redundant knowledge in storage and by keeping just one physical copy and referring alternative redundant knowledge to it copy. Though deduplication will increase the potency {of knowledge of information} storage there\'s a lot of privacy considerations raised as user's sensitive data area unit liable to each corporate executive and outsider attacks. The matter of deduplication got to be solve with differential privileges in cloud computing employing a semi trustworthy third party World Health Organization will then perform the task of deduplication on the users behalf. Delta secret writing is planned for extremely redundant knowledge. It will additional improve the storage potency of enormous knowledge chunks.*

I.INTRODUCTION

Cloud computing is an emerging concept of information technology service. It consists of both infrastructure as well as the application services and focuses on types of users requirements. The users can identify their needs through the software in Cloud. In cloud computing, the word cloud is used as a metaphor for *the "Internet,"* so the phrase *cloud computing* means a type of "Internet-based computing," where different services such as servers, storage and applications are delivered to an organization's computers and devices through the Internet.

The current Virtual Machine (VM) resources scheduling in cloud computing environment mainly considers the current state of the system but seldom considers system variation and historical data, which always leads to load imbalance of the system. A scheduling strategy on load balancing of VM resources based on genetic algorithm. According to historical data and current state of the system and through genetic algorithm, the approach computes ahead the influence it will have on the system after the deployment of the needed VM resources and then chooses the least-affective solution, through which it achieves the best load balancing and reduces or avoids dynamic migration. This strategy solves the problem of load imbalance and high migration cost by traditional algorithms after scheduling

Currently in cloud computing, it mainly considers the current system condition in VM resources scheduling but rarely considers the pervious condition before scheduling and the influence on system load after scheduling which usually leads to load imbalance. Most of the load balancing exists in VM migration. Yet, when the entire VM resources are migrated, due to the large granularity of VM resources and the great amount of data transferred in migration and the suspension of VM service, the migration cost becomes a problem.

A cloud service has three distinct characteristics that differentiate it from traditional hosting. It is sold on demand, typically by the minute or the hour; it is elastic which means that a user can have as much or as little of a service as they want at any given time and the service is fully managed by the provider. A cloud can be private or public. A public cloud sells services to anyone on the Internet. A private cloud is a proprietary network or a data center that supplies hosted services to a limited number of people. When a service provider uses public cloud resources to create their private cloud, the result is called a virtual private cloud. Private or public, the goal of cloud computing is to provide easy, scalable access to computing resources and IT services.

Cloud computing is a general term for anything that involves delivering hosted services over the Internet. These services are broadly divided into three categories

- Infrastructure-as-a-Service (IaaS)
- Platform-as-a-Service (PaaS)
- Software-as-a-Service (SaaS).

The "Infrastructure-as-a-Service" as the road, it's the basis for communication. It's the bottom layer that you build your platform on.

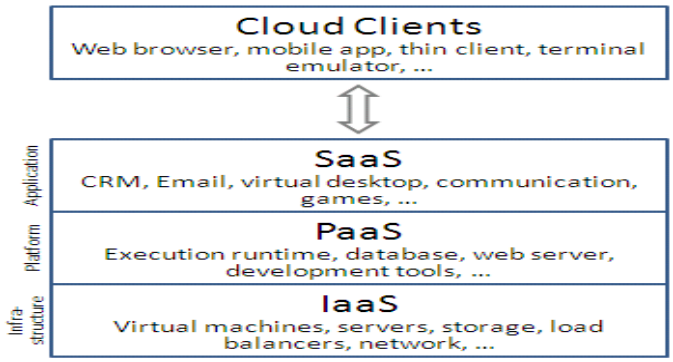

Fig. 1.1 Cloud Computing Services

The platform are the cars traveling on the infrastructure. PaaS rides on IaaS. But on the top of that, the goods and passengers inside the cars are the SaaS. It's the end user experience. It's the end result.

I.1COMMUNICATIONS IN CLOUD COMPUTING

For service developers, making services available in the cloud depends on the type of service and the device(s) being used to access it. The process may be as simple as a user clicking on the required web page, or could involve an application using an Application Program Interface (API) accessing the services in the cloud. Telcos are starting to use clouds to release their own services and those developed by others, but using Telco infrastructure and data. The list of communications in cloud computing are as follows

1.Using the Communications Services: When in the cloud, communications services can extend their capabilities, or stand alone as service offerings, or provide new interactivity capabilities to current services. Cloud-based communications services enable businesses to embed communications capabilities into business applications, such as Enterprise Resource Planning (ERP) and Customer Relationship Management (CRM) systems. They can be accessed from any location and linked into current services to extend their capabilities, as well as stand alone as service offerings.

2.Accessing through Web APIs: Accessing communications capabilities in a cloud-based environment is achieved through APIs. These APIs open up a range of communications possibilities, only limited by the media and signaling capabilities within the cloud. By using the Web APIs, these complexities of the code can be simplified and the media can be delivered to the remote device more easily.

3.Media Server Control Interfaces: When building communications capabilities into the "Core of the cloud," where they will be accessed by another service, the Web 2.0 APIs can be used, as well as a combination of SIP or Voice XML and the standard media controlling APIs such as MSML, MSCML and JSR309

4.Communications Scalability: To deliver on the scalability requirements for cloud-based deployments, the communications software should be capable of running in virtual environments. This allows for easily increasing and decreasing session densities based on the needs at the time, while keeping the physical resource requirement on servers to a minimum.

II.LITERATURE SURVEY

Jin Li, Yan Kit Li, Xiaofeng Chen and Patrick P, A Hybrid Cloud Approach for Secure Authorized Deduplication[1] Data deduplication is one of important data compression techniques for eliminating duplicate copies of repeating data and has been widely used in cloud storage to reduce the amount of storage space and save bandwidth. To protect the confidentiality of sensitive data while supporting deduplication, the convergent encryption technique has been proposed to encrypt the data before outsourcing. To better protect data security, it makes the first attempt to formally address the problem of authorized data deduplication. Different from traditional deduplication systems, the differential privileges of users are further considered in duplicate check besides the data itself. And to present several new deduplication constructions supporting authorized duplicate check in a hybrid cloud architecture. Security analysis demonstrates that it is secure in terms of the definitions specified in the proposed security model. As a proof of concept, it is to implement a prototype of our proposed authorized duplicate check scheme and conduct testbed experiments using our prototype. It is to show that the proposed authorized duplicate check scheme incurs minimal overhead compared to normal operations.

The problem of privacy preserving deduplication in cloud computing and propose a new deduplication system supporting for

Differential Authorization: Each authorized user is able to get his/her individual token of his file to perform duplicate check based on this privileges. Under this assumption, any user cannot generate a token for duplicate check out of his privileges or without the aid from the private cloud server.

Authorized Duplicate Check: Authorized user is able to use his/her individual private keys to generate query for certain

file and the privileges he/she owned with the help of private cloud, while the public cloud performs duplicate check directly and tells the user if there is any duplicate. The security requirements considered is to lie in two folds, including the security of file token and security of data files. For the security of file token, two aspects are defined as unforgeability and indistinguishability of file token. The details are given below.

Unforgeability of file token/duplicate-check token: Unauthorized users without appropriate privileges or file should be prevented from getting or generating the file tokens for duplicate check of any file stored at the Security-Critical Section Problem (S-CSP). The users are not allowed to collude with the public cloud server to break the unforgeability of file tokens. The S-CSP is honest but curious and will honestly perform the duplicate check upon receiving the duplicate request from users. The duplicate check token of users should be issued from the private cloud server in our scheme.

Indistinguishability of file token/duplicate-check token: It requires that any user without querying the private cloud server for some file token, he cannot get any useful information from the token, which includes the file information or the privilege information.

Data Confidentiality: Unauthorized users without appropriate privileges or files, including the S-CSP and the private cloud server, should be prevented from access to the underlying plaintext stored at S-CSP. In another word, the goal of the adversary is to retrieve and recover the files that do not belong to them. In the system, compared to the previous definition of data confidentiality based on convergent encryption, a higher level confidentiality is defined and achieved.

Waraporn Leesakul, Paul Townend and Jie Xu ,Dynamic Data Deduplication in Cloud Storage [11] Cloud computing plays a major role in the business domain today as computing resources are delivered as a utility on demand to customers over the Internet. Cloud storage is one of the services provided in cloud computing which has been increasing in popularity. The main advantage of using cloud storage from the customer's point of view is that customers can reduce their expenditure in purchasing and maintaining storage infrastructure while only paying for the amount of storage requested, which can be scaled-up and down upon demand. With the growing data size of cloud computing, a reduction in data volumes could help providers reducing the costs of running large storage system and saving energy consumption. So data deduplication techniques have been brought to improve storage efficiency in cloud storages. With the dynamic nature of data in cloud storage, data usage in cloud changes overtime, some data chunks may be read frequently in period of time, but may not be used in another time period. Some datasets may be frequently accessed or updated by multiple users at the same time, while others may need the high level of redundancy for reliability requirement. Therefore, it is crucial to support this dynamic feature in cloud storage. However current approaches are mostly focused on static scheme, which limits their full applicability in dynamic characteristic of data in cloud storage. A dynamic deduplication scheme for cloud storage is aiming to improve storage efficiency and maintaining redundancy for fault tolerance.

Cloud computing has recently emerged as a popular business model for utility computing system. The concept of cloud is to

Provide computing resources as a utility or a service on demand to customers over the Internet. The concept of cloud Computing is quite similar to grid computing, which aims to achieve resource virtualisation. In grid computing, the organisations sharing their computing resources, such as processors, in order to achieve the maximum computing capacity, whereas cloud computing aims to provide computing resources as a utility on demand, which can scale up or down at any time, to multiple customers. This makes cloud computing play a major role in the business domain, whereas grid is popular in academic, scientific and engineering research. In general, it is to define cloud computing as a business model that provide computing resources as a service on demand to customers over the Internet.
The essential characteristics of cloud computing have been defined in. Cloud providers pool computing resources together to serve customers via a multi-tenant model. Computing resources are delivered over the Internet where customers can access them through various client platforms. Customers can access the resources on-demand at any time without human interaction with the cloud provider. From a customers' point of view, computing resources are infinite, and customer demands can rapidly change to meet business objectives. This is facilitated by the ability for cloud services to scale resources up and down on demand leveraging the power of virtualization. Moreover, cloud providers are able to monitor and control the usage of resources for each customer for billing purposes, optimization resources, capacity planning and other tasks.

Cloud storage is one of the services in cloud computing which provides virtualized storage on demand to customers. For example, customers can use cloud storage as a backup service, as opposed to maintaining their own storage disks. Organizations can move their archival storage to the cloud which they can achieve more capacity at the low-cost, rather than buying additional physical storage. Applications running in the cloud also require temporary or permanent data storage in order to support the applications. As the amount of data in the cloud is rapidly increasing, customers expect to reach the on-demand cloud services at any time, while providers are required to maintain system availability and process a large amount of data.

Providers need a way to dramatically reduce data volumes, so they can reduce costs while saving energy consumption for running large storage systems.With the aim to reduce storage space, in traditional deduplication systems, duplicated data chunks identify and store only one replica of the data in storage. Logical pointers are created for other copies instead of storing redundant data. Deduplication can reduce both storage space and network bandwidth. However such

techniques can result with a negative impact on system fault tolerance. Because there are many files that refer to the same data chunk.

Amrita Upadhyay, Pratibha R Balihalli, Shashibhushan Ivaturi and Shrisha Rao, Deduplication and Compression Techniques in Cloud Design [2] Deduplication and compression in cloud computing aims at reduction in storage space and bandwidth usage during file transfers. The design depends on multiple metadata structures for deduplication. Only a copy of the duplicate files is retained while others are deleted. The existence of duplicate files is determined from the metadata. The files are clustered into bins depending on their size. They are then segmented, deduplicated, compressed and stored. Binning restricts the number of segments and their sizes so that it is optimum for each file size. When the user requests a file, compressed segments of the file are sent over the network along with the file-to-segment mapping. These are the uncompressed and combined to create a complete file, hence minimizing bandwidth requirements. Cloud technology has emerged and become a very important aspect of many business fields.

Computer resources are used and accessed through networks. Storage in a public cloud in terms of space and time is vital for backup and recovery, email systems, etc. It focuses on the infrastructure services dealing with storage and network usage. Deduplication and compression, are some of the important data optimization services that the cloud offers.
A new architecture for deduplication on the cloud using functionalities like segmentation, compression and binning. It talks about local block-level deduplication which is verified using the Eucalyptus environment. A global deduplication across various users can have security issues and have designed multiple metadata structures which enable faster lookups and enhance user experience. A cloud based on the proposed architecture has better storage efficiency and lesser bandwidth consumption. The architecture benefits both cloud service providers and users. Considerable amounts of space can be saved in the cloud by means of deduplication and compression. Segmentation of a file reduces it to smaller chunks which are easier to transfer over the Internet. Sending unique compressed segments minimizes bandwidth consumption significantly. This of course results in reduction of cost and increase in storage efficiency, while also improving the user experience.

In case of duplicate files, almost 80% of redundant data is removed. The transfer is not only less costly but also faster at the user end. Deduplication is done in two situations: for existing files and for incoming (new) files. Files are initially segregated into different bins depending on their size. When segmentation is done with constant size for all files irrespective of their size, it either leads to a large number of segments or very few segments. Binning helps in deciding the size of each segment to be formed, based on the size of the parent file. This decreases the time to process and save the segments, also response time over the network. Based on test results, a reduction of 47.5% in the processing time is seen due to this process.

For a user, bandwidth is a limiting factor when accessing a public storage cloud over the Internet. With the advent of mobile devices which connect to cloud, the end user uses cloud services that may require frequent uploads and downloads. A user is also charged by the cloud service provider based on the storage space and/or the bandwidth consumed. It provides a solution which aims at minimizing both, without loss of data. When a user requests for a file, it is reconstructed and given to him. This file at the user end is no different from the one uploaded, making the whole process transparent to the user. Experiments show that the bandwidth utilization as measured using iperf is reduced by 31%. For cloud storage, Atmos offers scalability, automated data placement and information services all over the globe.

sNikolai Samteladze and Ken Christensen,DELTA: Delta Encoding for Less Traffic for Apps[8] The number of applications (or apps) in the Android Market exceeded 450,000 in 2012 with more than 11 billion total downloads. The necessity to fix bugs and add new features leads to frequent app updates. For each update, a full new version of the app is downloaded to the user's smart phone; this generates significant traffic in the network. The proposed system is to use delta encoding algorithms and to download only the difference between two versions of an app and the implementation delta encoding for Android using the bsdiff and bspatch tools and evaluate its performance. The app update traffic can be reduced by about 50%, this can lead to significant cost and energy savings.

In 2012 the Android Market had more than 450,000 applications available and over 11 billion total app downloads with more than 1 billion new downloads happening every month. The introduction of new features and bug fixes make it usual for an app to have updates released every few weeks. Each update results in a download of the entire new app to a smart phone that has the previous version of the app installed. Such app updates increase the user's download traffic, and thus also increases traffic on the cellular infrastructure. It address how to reduce the volume of app update traffic on cellular networks and also reduce the load on app servers in data centers that serve apps to smart phone users.

When an app is updated, the entire app file, an APK file for an Android app is downloaded to a mobile device such as a smart phone. This is the case even if only a few bytes have changed from the previous version of the app. The proposal is to update an app by transferring only the difference between the old and new versions and then applying such a delta patch locally in the smart phone. Delta encoding is widely used to reduce the size of update files. It was first developed for Unix. Modern software, such as the Google Chrome browser, uses delta encoding algorithms to reduce the size of updates.

There has been a previous attempt to use delta encoding algorithms to reduce the size of Android APK files. To the best of our knowledge, this previous work has not resulted in a working code or a prototype system. The contribution is the first working prototype of delta encoding for Android app updates. The experimental results show that app updates can be reduced in size by 48% on average. Delta encoding or more

generally, data differencing is an approach to storing or transferring data using calculated differences between files rather than files themselves. A delta encoding algorithm takes old and new versions of a file as an input and computes the difference between them. Such difference data, which will refer to as a patch, can be used to construct the new version of the file from the old one. The measures of interest for delta encoding implementations are the size of the patch and the amount of memory and the time required to construct a patch and to apply it.

Taking advantage of the internal structure of the input files can significantly reduce the size of the generated patch. One of the most well-known optimizations of the bsdiff tool is Courgette, which is used to generate patches for the Google Chrome browser. Courgette decreases the size of a patch by taking into account the specifics of a compiled application file. Such files contain a lot of internal references, which completely change their values from even small changes in the source code. Courgette uses disassembling to find all the internal pointers and recover their values. This optimization allows Courgette to achieve a 10 times smaller patch size for the Google Chrome browser than that of bsdiff.

III. DESCRIPTION OF THE SYSTEM

Each data copy (i.e.,) a file or a block is associated with a token for the duplicate check.

- **S-CSP:** This is an entity that provides a data storage service in public cloud. The S-CSP provides the data outsourcing service and stores data on behalf of the users. To reduce the storage cost, the S-CSP eliminates the storage of redundant data via dedu-plication and keeps only unique data. It is to assume that S-CSP is always online and has abundant storage capacity and computation power.
- **Data Users:** A user is an entity that wants to out-source data storage to the S-CSP and access the data later. In a storage system supporting deduplication, the user only uploads unique data but does not upload any duplicate data to save the upload bandwidth, which may be owned by the same user or different users. In the authorized deduplication system, each user is issued a set of privileges in the setup of the system. Each file is protected with the convergent encryption key and privilege keys to realize the authorized deduplication with differential privileges.
- **Private Cloud:** Compared with the traditional deduplication architecture in cloud computing, this is a new entity introduced for facilitating user's secure usage of cloud service. Specifically, since the computing resources at data user/owner side are restricted and the public cloud is not fully trusted in practice, private cloud is able to provide data user/owner with an execution environment and infrastructure working as an interface between user and the public cloud. The private keys for the privileges are managed by the private cloud, who answers the file token requests from the users. The interface offered by the private cloud allows user to submit files and queries to be securely stored and computed respectively.

Typically, assume that the public cloud and private cloud are both "honest-but-curious". Specifically they will follow our proposed protocol, but try to find out as much secret information as possible based on their possessions. Users would try to access data either within or out of the scopes of their privileges.

All the files are sensitive and needed to be fully protected against both public cloud and private cloud. Under the assumption, two kinds of adversaries are considered, that is

- External adversaries which aim to extract secret information as much as possible from both public cloud and private cloud;

Internal adversaries who aim to obtain more information on the file from the public cloud and duplicate-check token information from the private cloud outside of their scopes.

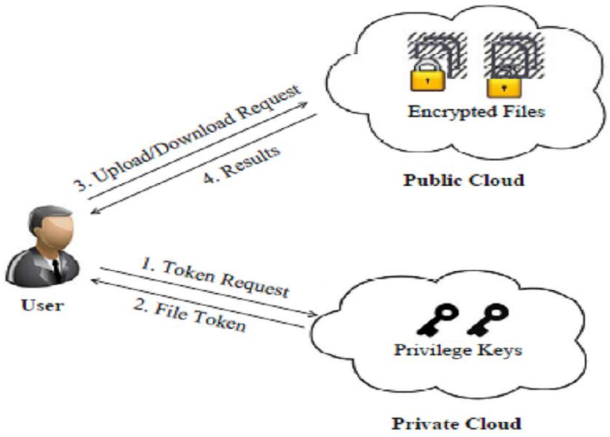

Fig. 3.2.1 Architecture for Authorized Deduplication

Convergent encryption provides data confidentiality in deduplication. A user(or data owner) derives a convergent key from each original data copy and encrypts the data copy with the convergent key. In addition, the user also derives a tag for the data copy, such that the tag will be used to detect duplicates. Here, the tag correctness property holds (i.e.,) if two data copies are the same, then their tags are the same. To detect duplicates, the user first sends the tag to the server side to check if the identical copy has been already stored. Note that both the convergent key and the tag are independently derived, and the tag cannot be used to deduce the convergent key and compromise data confidentiality. Both the encrypted data copy and its corresponding tag will be stored on the server side.

The problem of privacy- preserving deduplication in cloud computing and propose a new deduplication system supporting for

- **Differential Authorization:** Each authorized user is able to get his/her individual token of his file to perform duplicate check based on his privileges. Under this assumption, any user cannot generate a token for duplicate check out of his privileges or without the aid from the private cloud server.

- **Authorized Duplicate Check:** Authorized user is able to use his/her individual private keys to generate query for certain file and the privileges he/she owned with the help of private cloud, while the public cloud performs duplicate check directly and tells the user if there is any duplicate.

IV.DELTA ENCODING

The proposed system is to use Delta Encoding which can handle highly redundant data and can handle large amount of data chunks with ease

The proposed system is to use Delta Encoding which can handle highly redundant data and can handle large amount of data chunks with ease.

- **Load Balancer -** After hashing process with SHA-1, clients send a fingerprint (hash value) to a deduplicator via the load balancer. The load balancer responds to requests from clients sending to any one of deduplicators according to their loads at that time.
- **Deduplicators -**A component designed for identifying the duplication by comparing with the existing hash values stored in metadata server.
- **Cloud Storage -** A Metadata Server to store metadata, and a number of File Servers to store actual files and their copies.
- **Redundancy Manager -** A component to identify the initial number of copies, and monitor the changing level of Quality of Service (QoS).

There are three events which we simulated: upload, update, and delete. The upload event is when the file is first uploaded to the system. If files already exist in the system, and have been uploaded again, the number of copies of the files will be recalculated according to the highest level of QoS, this is for an update event. For a delete file event, users can delete their files, but the files will not permanently deleted from the system if there are any other users refer to the same files.

Deduplicator calls a hash value of the uploaded file from client, and then checks for any duplicates with the same existing hash value in metadata server. If it is a new file, the new metadata of the file will be added to the system and the file will be uploaded to file server. The replicas of the file will be created according to the level of QoS of the upload file.

- **Update:**In the case of existing file, the metadata of the file will be updated and the system may need to create or delete the replicas of the file according to the maximum value of QoS of the file.

- **Delete:**The deduplicator checks the number of files which refer to the same hash value user wants to delete. If there is only one reference to the hash, all replicas of the file will be deleted. On the other hand if there are any other files that refer to the hash, only the metadata will be updated, and the number of replicas of the file may need to decrease according to the maximum value of QoS.

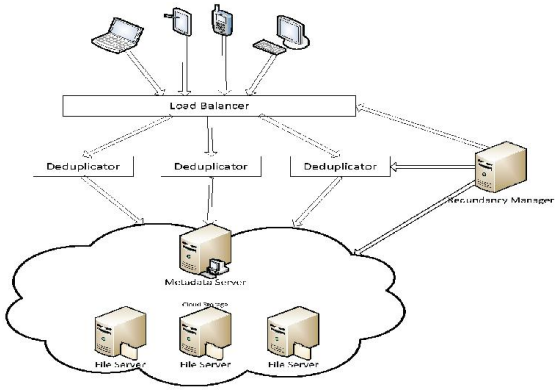

Fig. 5.1 Delta Encoding

V. DELTA ENCODING TECHNIQUE

To find out the duplicates, the prevalent method is used a hashing algorithm to find the integrity of data. The techniques are as follows:

- **Duplicate Detection:** The data stream is first chunked by the CDC approach, fingerprinted by Secure Hash Algorithm-1(SHA-1), duplicate-detected, and then grouped into container of sequential chunks to preserve the backup-stream locality.
- **Resemblance Detection:** The Duplicate Adjacency(DupAdj) Resemblance detection module in DARE first detects duplicate-adjacent chunks in the containers formed in Step. After that improved super-feature module further detects similar chunks in the remaining nonduplicate and non-similar chunks that may have been missed by the DupAdj detection module when the duplicate-adjacency information is lacking or weak.
- **Delta Compression:** For each of the resembling chunks detected, reads its base-chunk, then delta encodes their differences. In order to reduce the number of disk reads, an Least Recently Used(LRU) and locality-preserved cache is implemented here to

prefetch the base-chunks in the form of locality-preserved containers.

- **Storage Management:** The data not reduced, (i.e.,) non-similar or delta chunks, will be stored as containers into the disk. The file mapping information among the duplicate chunks, resembling chunks, and non-similar chunks will also be recorded as the file recipes to facilitate future data restore operations.

VI. Conclusion

In this paper, the notion of licensed knowledge deduplication was proposed to shield the info security by including differential privileges of users within the duplicate check. we tend to conjointly bestowed delta coding which might handle extremely redundant knowledge and might handle great amount of information chunks with ease.There area unit 3 events that we tend to simulated: transfer, update, and delete.To find out the duplicates, the prevailing technique is employed a hashing algorithmic program to seek out the integrity of information. We showed that our licensed duplicate check theme incurs lowest overhead compared to focused encoding and network transfer.

.

References

[1]Jin Li,Yan Kit Li, Xiaofeng Chen and Patrick P, "A Hybrid Cloud Approach for Secure Authorized Deduplication", (2014).
[2]Amrita Upadhyay, Pratibha R Balihalli, Shashibhushan Ivaturi and Shrishab Rao, "Deduplication and Compression Techniques in Cloud Design", (2012).
[3]Anderson. P , Zhang. Z, "Fast and secure laptop backups with encrypted deduplication", In Proc. of USENIX LISA(2010).
[4]Bellare.M,Keelveedhi.S,RistenparDupless.T, "Serveraided encryption for deduplicated storage", In USENIX Security Symposium(2013).
[5]Bugiel. S, Nurnberger. S, Sadeghi. A,Schneider. T, "Twin clouds: An architecture for secure cloud computing", In Workshop on Cryptography and Security in Clouds (2012).
[6]Li. J, Chen. X, Li. X, Lee. P, "Secure deduplication with efficient and reliable convergent key management", In IEEE Transactions on Parallel and Distributed Systems(2013).
[7]Ng. C, Lee. P, "Revdedup: A reverse deduplication storage system optimized for reads to latest backups", In Proc. of APSYS (2013).
[8]Samteladze. N , Christensen. N, "DELTA: Delta Encoding for Less Traffic for Apps",(2012).
[9]Stanekiii. J, Sorniotti. A, Androulaki. E, Kencl. L,"A secure data deduplication scheme for cloud storage", In Technical Report(2013).
[10]Storer.M,Greenan. W, Long. D.D.W , Miller. E.L,"Secure data deduplication", In Proc. of StorageSS(2008).
[11LeesakulW,Townend.P,Xu.J,"DynamicData Deduplication in Cloud Storage", (2014).
[12]Zhang. K, Zhou. X, Chen. X, Ruan. Y, "Privacy-aware data intensive computing on hybrid clouds", In Proceedings of the 18th ACM conference on Computer and communications security (2011).

Kumar Charlie Paul, Principal of A.S.L Pauls College of Engineering & Technology. Had did many National and International Conferences and published many papers in journals. He also guided many students for their Ph.D project works. Having more than 23 years of experience in teaching field.

Felcia Jerlin, Assistant Professor / CSE of Dr.G.U Pope College of Engineering, . Doing Ph.d Had did many National and International Conferences and published many papers in journals. she also guided many students for their UG and PG project works.