ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

Usability Matrix On Dynamic Datasets For Cloud Storage Solution Framework

Ms.S.Dharani, Ms.S.Shanthi

PG Student, M.E (CSE), Valliammai Engineering College, Chennai, India[1]

Assistant Professor, Deapartment of CSE, Valliammai Engineering College, Chennai, India[2]

ABSTRACT- Data security and privacy of data is one of the major concern in the cloud computing. To overcome this situation of information disclosure, a particular standard of encryption is done to the sensitive data before uploading onto the cloud servers. It becomes clear that, the plain text keyword search is not viable. In the existing system, third parties which have privilege over intermediate datasets are created in order to reduce the frequent access of data from cloud directly that increases the cost. The severity data's in the intermediate sets are encrypted using homomorphism algorithm and least accessible data's are hided. In turn, does not root the inference channel analysis. Identify the most frequent access data and less frequent access data and finding the possible solutions of encryption is the core concept discussed in the existing system. There is a serious flaw which deals with identifying the less access table Vs more frequent access table. The reason is the most frequent access table may have relation with some other table in the database and using those options; the most frequent access table can deduce with some other table and manipulate the data. In the proposed system, using the privacy leakage constrain column wise encryption has been done for unencrypted data's in the intermediate dataset. And a concept encrypting the data thereby finding out reference attribute between data tables are achieved. In addition to the exceeding system, there is an automatic scheduling algorithm to maintain a log based tracking for frequent and un frequent usage of data under the time criteria.

KEYWORDS—Cloud computing, data storage and privacy, privacy preservation, column wise encryption, automatic scheduling.

I. INTRODUCTION

Cloud computing is used to illustrate a variety computing concepts that involve a large number of computers connected through a real-time communication network such as the Internet. Cloud computing is similar to distributed computing over a network, and it has the ability to run a program or application on many connected computers at the same time. It defines the network-based services, which emerge to be providing by real server hardware, and are indeed served up by virtual hardware implicit by software running on one or more real machines. Such virtual servers do not actually exist and can therefore be moved around and scaled up or down without affecting the end user, similar to cloud. cloud computing provides services in software, platform and at infrastructure These cloud services may receive in a Public, Private or Hybrid network[1].The use of such virtualized resources allows the user to completely customize the Virtual Machine (VM) images. The organizations provide private clouds to improve the resource utilization of the available computation facilities. Privacy is an increasingly important aspect of data publishing. Cloud computing is one of the most pre-dominant paradigm in recent trends for computing and storing purposes. The end of this decade is marked with a predominant change in the industrial information technology towards a pay-per-use service business model which provides an optimistic approach for the end users on the software's and storage. Going ahead with many advantages in the computing world and storage resources offered by cloud service providers, the data owners must lay their important information into the public cloud servers which are not within their trusted domains. Data security and privacy of data is one of the major concern in the

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

cloud computing. To conquer this position of information disclosure, the service providers are enforce a exacting standard of encrypting the sensitive data before uploading onto the cloud servers. It become noticeable that, the plain text keyword search is not viable. As the whole amount of data stored in public clouds increases in an exponential manner, it becomes too firm to support well-organized keyword based queries and ranking the matching results on encrypted data.

However, if personal private information is leaked from the database, the service will be regarded as unacceptable by the original owners of the data. There are two approaches to avoiding leaks of private information from public databases: generalization methods and perturbation methods. Generalization methods modify the original data to avoid identification of the records. These methods generate a common value for some records and replace identifying information in the records with the common value. However, detailed information is lost during this process. On the other hand, perturbation methods add noise to data. While perturbed data usually retains detailed information, it also normally includes fake data. In particular, when performing such an evaluation, it is difficult to model the background knowledge of an adversary trying to obtain private information from a database. Even if some fields of records in a database have been anonymized in some manner, an adversary may still be able to identify a record through background knowledge[1][4].

For example, even if ZIP codes are generalized to include just the highest level of regional information in a medical database, this may still be enough to identify a record if there is only one case of a particular disease in that region and an adversary knows that a particular target has had that disease and lives in that region. Since the generalization and perturbation methods take such different approaches, it is very difficult to compare them[4].

Cloud Computing, the long-held dream of computing as a utility, has the potential to make over a large part of the IT engineering, making software even more striking as a service and shaping the way IT hardware is planned and purchased. Developers with modern ideas for new Internet services no longer require the large investment outlays in hardware to organize their service or the human expenditure to operate it. They need not be fretful about over-provisioning for a service whose status does not meet their prediction, thus slaying costly resources, or under-provisioning for one that becomes disgracefully accepted, thus missing potential customers and profits. Moreover, companies with large batch-oriented tasks can get results as rapidly as their programs can scale, since using 1100 servers for one hour costs no more than using one server for 1100 hours. This flexibility of resources, without paying a payment for huge scale, is unmatched in the history of IT[14].

1. To reduce that confusion by instructive terms, providing simple statistics to enumerate comparisons between of cloud and conventional Computing, and identifying the top technical and non-technical obstacles and opportunities of Cloud Computing.

2. Usage-based pricing is not renting.

3. Renting a resource involves paying a negotiated cost to have the resource over some time period, whether or not you utilize the source.

II. RELATED WORK

A. Future Generation Computer Systems

k-anonymization is an important privacy protection mechanism in data publishing. Such mechanisms only protect the data up to the rst release or rst recipient. In practical applications, data is published continuously as new data arrive; the same data may be anonymized differently for a different purpose or a different receiver. In such scenarios, even when all releases are properly k-anonymized, the anonymity of an individual may be accidentally compromised if recipient cross-examines all the releases received with other recipients. Preventing such correspondence attacks faces major challenges. In this paper, it characterize the correspondence attacks and propose an efficient anonymization algorithm to the attacks in the model of continuous data publishing.This paper provides a systematic way to characterize the correspondence attacks and propose an efficient anonymization algorithm to prevent the attacks in the model of continuous data publishing. All the possible attacks like F-attack, C-attack and B-attack is discussed in this paper[2]. This paper provides the following disadvantage: Data with frequent changes like frequent updates / inserts were not considered in this paper and it adds a major drawback in this paper. Due to dynamic environment, the frequent data releases makes this project void.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

B. Cloud Computing and Emerging IT Platforms

Recently, privacy preserving data publishing has received a lot of attention in both research and applications. It refers to static data sets. In this paper, an emerging problem of continuous privacy preserving publishing of data streams which cannot be solved by any straightforward extensions of the existing privacy preserving publishing methods on static data. To undertake the problem, we build up a novel approach which considers both the distribution of the data entries to be published and the statistical distribution of the data stream. An wide-ranging performance using both real data sets and synthetic data sets verifies the effectiveness and the efficiency of our methods. Distribution of the data entries to be published and the statistical distribution of the data stream is the core idea of this project and it's not handled ever before in this perception. A concrete model and an anonymization quality measure, and developed a group of randomized methods[3]. The disadvantages are: This paper didn't clearly explain once the data is mined for getting multidimensional data. Impact due to work load on the data streams is not discussed in this paper.

C. In Cloud, Can Scientific Communities Benefit from the Economies of Scale

The basic idea behind Cloud computing is that resource providers offer elastic resources to end users. We intend to answer one key question to the success of Cloud computing. In Cloud, can undersized or medium-scale logical computing communities benefit from the economies of scale. An enhanced scientific public cloud model (ESP) that encourages small or medium scale research organizations rent elastic resources from a public cloud provider. On a basis of the ESP model we design and implement the Dawning Cloud system that can consolidate heterogeneous scientific workloads on a Cloud site. An innovative emulation methodology and perform a comprehensive evaluation for two typical workloads Two typical workloads: high throughput computing (HTC) and many task computing (MTC), Dawning Cloud saves the resource consumption maximally by 44.5% (HTC) and 72.6% (MTC) for service providers and saves the total resource consumption maximally by 47.3% for a resource provider with respect to the previous two public Cloud solutions. To this end, we wind up that for typical workloads: HTC and MTC, Dawning Cloud can enable scientific communities to benefit from the economies of scale of public Clouds[6][5]. The disadvantages are: A prominent shortcoming of the dedicated system model: for peak loads, a dedicated cluster system cannot provide enough resources, while lots of resources are idle for light loads.

D. Security and Privacy Challenges in Cloud Computing Environments

Cloud computing has generated significant interest in both academia and industry, but it's still an evolving standard. Basically, it aims to secure the economic utility model with the evolutionary development of many existing approach and computer technology. Confusion exists in IT industry about how a cloud differs from existing models and how these differences affect its adoption. Nevertheless, cloud computing is an important standard, with the potential to significantly reduce costs through optimization and increased operating and economic efficiency. Without suitable security and privacy solutions designed for clouds, this potentially revolutionize computing standard could become a huge failure. Several surveys of latent cloud adopters designate that security and privacy is the primary concern hindering its adoption Cloud computing is a model for enabling expedient, on-demand network access to a shared group of configurable computing resources (e.g., networks, servers, storage, applications and services) that can be swiftly provisioned and released with minimal management effort or service provider interaction. This cloud representation promotes availability and is composed of five critical characteristics, three deliverance models, and four exploitation models[7]. The disadvantages are: In clouds, service providers usually don't know their users in advance, so it is complicated to allocate users directly to roles in access control policies.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

III. PRELIMINARIES

A. Preservation Of Cost Using Heuristic Algorithm

The heuristic algorithm is used to identify the data sets that need to be encrypted. . In this the entire data sets are encrypted together and hence there occur a high overhead and high cost .Hence propose a privacy-preserving cost with intermediate data sets [8].

B. Frequent Pattern Mining Algorithm

In this, A tree structure has been developed from the intermediate data sets in order to analyze privacy propagation among data sets. The algorithm defines

The frequent-pattern tree (FP-tree) is a structure that stores quantitative information about frequent patterns in a database. Han defines the FP-tree as the tree structure defined below : One root labelled as "null" with a set of item-prefix sub trees as children, and a frequent-item-header table;

Each sub tree consists of three fields:

Item-name: register which entry is represented by the node;

Count: the number of connections represented by the portion of the path reaching the node;

Node-link: associates to the next node in the FP-tree carrying the same item-name, or null if there is none[9].

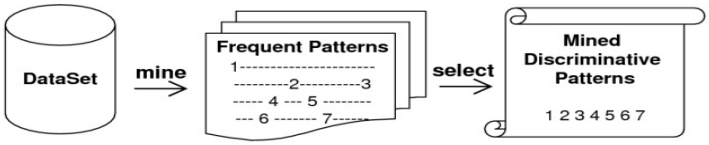

Fig 1: Frequent Pattern Mining

Input: A database DB, represented by FP-tree constructed according to Algorithm 1

Output: The entire set of frequent patterns.

Method: call FP-growth(FP-tree, null).

Procedure FP-growth(Tree, a)

{

(01) if Tree contain a distinct prefix path then // Mining distinct prefix-path FP-tree

{

(02) let P be the distinct prefix-path part of Tree;

(03) let Q be the multipath part with the apex branching node replaced by a null root;

(04) for each combination (denoted as ß) of the nodes in the path P do

(05) generate pattern ß ∪ a with support = minimum support of nodes in ß;

(06) let frequent pattern set(P) be the set of patterns so generated;

}

(07) else let Q be Tree;

(08) for each item ai in Q do { // Mining multipath FP-tree

(09) generate pattern ß = ai ∪ a with support = ai .support;

(10) create ß's conditional pattern-base and then ß's conditional FP-tree Tree ß;

(11) if Tree ß ≠ Ø then

(12) call FP-growth(Tree ß , ß);

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

(13) let frequent pattern set(Q) be the set of patterns so generated;
}
(14) return(freq pattern set(P) ∪ freq pattern set(Q) ∪ (freq pattern set(P) × freq pattern set(Q)))
}

C. Message Digest Algorithm

Message Digest is a small piece of data that results encryption. Message digest algorithm is used for verify the data integrity. MD5 is currently very vulnerable to collision attacks. Consider SHA1 broken since collision attacks are feasible. This means that MD5 executes faster.MD5 Message-Digest Algorithm is a widely used cryptographic hash function that produces a 128-bit (16-byte) hash value.MD5 processes a variable-length message into a fixed-length output of 128 bit. The input message is broken down up into chunk of 512-bit blocks (sixteen 32-bit words); the message is padded so that its length is dividable by 512. The padding works as follows: first a single bit 1, is append to the end of the message. This is followed by as many zero as are required to fetch the length of the message up to 64 bits less than a multiple of 512. The residual bits are filled up with 64 bits instead of the length of the unique message, modulo 264.

1) The advantages

MD5 twisted to be very broken with regards to collisions (we can produce a collision in a few seconds of work on a PC) and SHA-0 is also broken in that respect; SHA-1 is a bit blistering; How a hash function achieves collision resistance is a bit of a miracle since the complete function is totally known, with no secret value; it just mixes the data too much for the best cryptographers to loosen the process[12][10].

D. Flowtimescheduling Algorithm

Hard periodic real-time scheduling problems are motivated by special requirements of real-time systems arising in safety critical environments, e. g. the avionics or automotive industry. Here each task $t_i=(c(t_i),p(t_i))$ releases a job of running time $c(t_i)$ periodically exactly every $p(t_i)$ time units. The problem is to find stating offsets that ensure that no collision of different jobs occurs.

1) Binary Search

We represent a binary tree by a linked data structure in which each node is an object. Each node contains the fields key and possibly other satellite data.
left: points to left child.
right: points to right child.
p: points to parent. p[root[T]] = NIL.
keys must satisfy the binary-search-tree property[11].
If y is in left sub tree of x, then key[y] □ key[x].
If y is in right sub tree of x, then key[y] □ key[x].

A)Tree-Minimum(X)

Less frequent access data is scheduled by using this pseudo code;

while left[x] ≠ NIL
x := right[x]
end while
return x
end TREE-MINIMUM

B)Tree-Maximum(X)

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

Most frequent access data is scheduled by using this pseudo code;

```
while right[x] ≠ NIL
x := right[x]
end while
return x
end TREE-MAXIMUM
```

C)Advantage
1.Binary search can interact poorly with the memory hierarchy.
2.It is faster than linear search[11][13].

IV. MOTIVATING EXAMPLE AND PROBLEM ANALYSIS

Existing techniques are efficiently used for achieving data staging and data storage for privacy concern on a set of vantage to reduce the computational cost of encryption or decryption of data sets in a cloud system with a minimum outlay. Surplus data used for improvising the efficient optimal solutions is based on the dynamic upper bound privacy which is polynomial bounded by the number of service requests and the number of distinct data items in cloud. This is partial as most of the existing staging or privacy upper bound targets towards a class of services that access and process the decrypted data and thereby inherit the severity of data when access time sequence is more. Alternatively, a constraint optimization problem can be defined as a regular constraint to find a solution to the problem whose expenditure, evaluated as the sum of the cost functions, is minimized. Third parties who have privilege over intermediate datasets are created in order to reduce the frequent access of data from cloud directly that increases the cost. Hence the procedure of anonymization and homomorphic type of encryption are done in the system. In turn, avoids the possibility of inference channel analysis.

1) Disadvantage

The major disadvantage of the system is the relation between a particular sensitive data with the other data should be identified properly and it should be anonymized. Frequent access pattern on the data may get changed in timely manner.

V. PROPOSED WORK

Our proposed system is designed to identify only the important and critical intermediate datasets that needs to be encrypted for security purposes, hence reducing encryption/decryption cost and thus maintaining data privacy. One way for evaluating this upper bound for a partial solution in our existing paradigm is to consider each constraint separately and mining the data in order to restrict access when the user claims to find the original information. For each constraint, the maximal possible value for any of these values is an upper bound may recover privacy-sensitive partial column level encryption. Hence an column wise encryption in the unencrypted data's of intermediate datasets are proposed. Additional a feature of encrypting on the basis of reference attribute between the data tables are achieved to reduce the cost complexity when accessing the data. An automatic scheduling strategy is involved to maintain an log report of the frequent and infrequent usage of intermediate dataset under time conditions. As a result, the algorithm requires an upper bound on the cost that can be obtained from extending a partial solution, and this upper bound should be significantly reduced with our approach over existing ones. Pattern is a form or model (or, more abstractly, a set of rules) that can be used to make or to generate things or parts of a thing.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

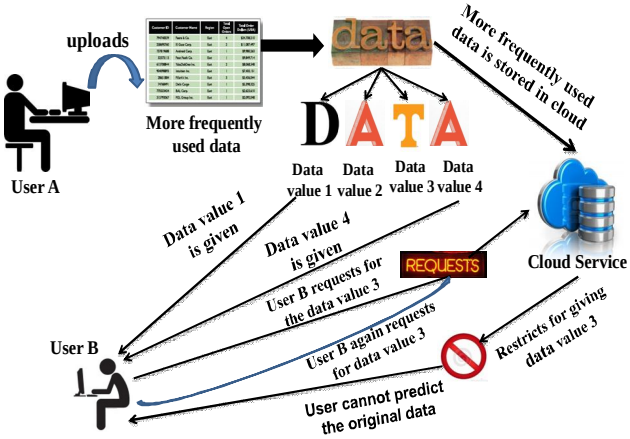

Fig 3: Architecture diagram

The proposed architecture describes the most frequent access data and less frequent access data and finding the possible solutions of encryption is the core concept .The most frequent access table can infer with some other table and manipulate the data. In the proposed system, using the privacy leakage constrain column wise encryption has been done for unencrypted data's. The important and critical intermediate datasets are encrypted for security purposes. It reduces encryption/decryption cost and thus maintaining data privacy. upper bound for a partial solution is used. Hence an column wise encryption in the unencrypted data's of intermediate datasets are proposed. An automatic scheduling strategy is involved to maintain an log report.

In fig 3 user A uploads a datasets of an organization . These frequently accessible datasets are stored in the intermediate datasets of cloud storage. It avoid the cost of accessing the cloud frequently. The severity estimation of the datasets are identified in order to perform inference channel analysis to avoid privacy leakage. The process of anonymization is done for maintaining privacy.As the size of the data utilization reduces cost of the utilized data get reduced. the data's are encrypted and user A requests cloud service and hence the original data cannot be predicted by unauthorized user.

The parameter that we have to decide upon is called support of an item set. In order to identify the item sets that are accessed most commonly in the cloud storage environment. These frequently accessible datasets are stored in the intermediate datasets of cloud storage so as to avoid the cost of accessing the cloud frequently. The entity is a individual object, position or occurrence for which data is collected. The relationship is the interface between the entities.

The table with common references with another tables are identified so as to prevent the privacy leakage through inference channel analysis. Inference channel analysis is a control used in the output of databases to stop a person who has access to only summary information from being able to determine (infer) a particular value for a particular record. The severity estimation of the datasets are identified in order to perform inference channel analysis to avoid privacy leakage.

1) Anonymity typically refer to the state of an individual's individual identity, or individually specialized information, being publicly unknown. This process of anonymization is done for maintaining privacy as well as to reduce the utilization size of the data. As the size of the data utilization reduces cost of the utilized data get reduced. The use of encryption/decryption is the skill of communication. In time of war, a cipher, often incorrectly called a code.

2) Encryption is the translation of information into a structure, called a cipher text, that cannot be easily understood by unauthorized people. The anonymized datasets in the previous module are checked for its relativity with other datasets

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

belongs to another table and doubly encrypted. The algorithmic concepts of Message digest Double Encryption Algorithm has been utilized in this module to perform the double encryption of relational datasets.

3) Scheduling is the method by which threads, processes or data flows are given right to use system resources . This is usually done to load balance a system effectively or achieve a target quality of service. New updated datasets are traced effectively for every timestamps and the process of Anonymizing and privacy preserving has been done. This modules fulfills the proposed strategy thereby monitoring the system dynamically by utilizing the algorithmic technique of Flow time scheduling Algorithm.

1. Advantages of Proposed work

Automatic Scheduling process may enable the system synchronized as it is with the current situation. Finding all possible data and encrypt based on the relationships will provide more value and waitage to the system.

2. Experimental Evaluation Module

Computer performance is characterized by the amount of useful work accomplished by a computer system compared to the time and resources used. Depending on the framework, high-quality system performance this revise the strategies of Short response time for a given piece of work, High throughput, Low utilization of computing resource, High availability of the computing system or application Fast data compression and decompression High bandwidth / short data transmission time The performance evaluation has been done and visualized graphically.

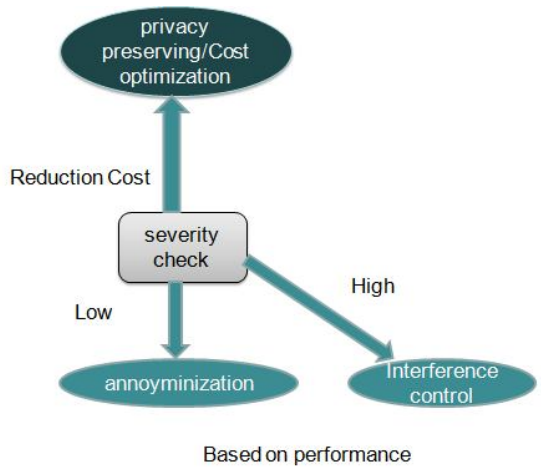

Fig 4:Experimental Evaluation Module

VI. CONCLUSION

In this paper, proposed an approach that identifies which part of intermediate data sets needs to be encrypted while the rest does not, in order to accumulate the privacy preserving. We have modeled the problem of saving privacy-preserving cost as a constrained optimization problem which is addressed by decomposing the privacy leakage constraints. we also investigate privacy aware efficient scheduling of intermediate data sets in cloud by taking privacy preserving as a metric together with other metrics such as storage and computation cost. We are planning to further preserve privacy and cost optimization of datasets that are accessible through cloud by considering many other factors such time span of usage, availability of servers and so on.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

REFERENCES

[1] http://en.wikipedia.org/wiki/Cloud_computing
[2] D. Zissis and D. Lekkas, "Addressing Cloud Computing Security Issues," Future Generation Computer Systems, vol. 28, no. 3, pp. 583- 592, 2011.
[3] R. Buyya, C.S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, "Cloud Computing and Emerging It Platforms: Vision, Hype, and Reality for Delivering Computing as the Fifth Utility," Future Generation Computer Systems, vol. 25, no. 6, pp. 599-616, 2009.
[4] http://www.vmware.com/in/cloud-computing/overview.html
[5] M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R. Katz, A.Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M.Zaharia, "A View of Cloud Computing," Comm. ACM, vol. 53,no. 4, pp. 50-58, 2010.
[6] L. Wang, J. Zhan, W. Shi, and Y. Liang, "In Cloud, Can Scientific Communities Benefit from the Economies of Scale?," IEEE Trans. Parallel and Distributed Systems, vol. 23, no. 2, pp. 296-303, Feb. 2012.
[7] H. Takabi, J.B.D. Joshi, and G. Ahn, "Security and Privacy Challenges in Cloud Computing Environments," IEEE Security & Privacy, vol. 8, no. 6, pp. 24-31, Nov./Dec. 2010
[8] https://www.google.co.in/search?q=heuristi +algorithm+in+cloud+computing
[9] http://www.wikibook.org/wiki/data_mining_algorithms_in_r/frequentpatternmining/the_fp-growth_algorithmhttp://www.research.ibm.com/people/n/nikhil/papers/thesis.pdf.
[10] http://en.wikipedia.org/wiki/MD5.
[11] http://www.ijava2.com/binary-search-tree- sample-algorithm-pseudo code/
[12] http://www.isi.edu/~touch/pubs/sigcomm95.html
[13] http://en.wikibooks.org/wiki/Data_Mining_Algorithms_In_R/Frequent_Pattern_Mining/The_FP-Growth_Algorithm.
[14] M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A View of Cloud Computing," Comm. ACM, vol. 53, no. 4, pp. 50-58, 2010.