Core End To End Traffic Retandary Elimination for Reducing Cloud Bandwidth Cost

[1]Vamsi Priya Challa, [2]M Vijetha

[1]Research Scholar, [2]Assistant Professor, Department of Computer Science & Engineering

KODADA Institute of Technology & Science For Women, Kodad

Abstract: In today's era, End to End TRE (Traffic redundancy) over the cloud computing is a recent technology and recent research topic under cloud computing environment. The Cloud computing and its mechanism to transfer the data, managing its component is an important process for complete process. Where the component such as third party auditor, data centre and virtual machine take participate. There are process involve in the middle such as Packet forwarding, data storage, transmission over the different channel takes process in user activity. In this area of research recent technique PACK takes an advantage of chunking scheme with execution of SHA-1 as security signature approach. According to understanding algorithm PACK face little problem formulation such as in finding high level prediction with efficient signature scheme. The algorithm exhibit reliable prediction over the data. In this dissertation work, our work presented as EPACK, which is the extension version of PACK to identify redundancy prediction over the transmitted data. Algorithm used of similarity measure algorithm, which take use of similarity score measure value from the available data and storage. The algorithm makes use of Meta information, topical relation in between the file parameters, SHA-2 for signing information and outperforms the high level prediction as compare to existing technique. The proposed algorithms also take use of security concern in between any transmission of data in different components. A highly security enhancement of ECC algorithm named HECC (Hyper Elliptic curve cryptography) is used to perform security level in this approach. This can performs over the PACK & EPACK algorithm with comparison parameter by performing data transmission using JDK framework. The proposed scheme and result comparison analysis shows the efficiency of our approach while comparing with PACK algorithm. The further approach over the simulation parameter and environment outperform its best while comparing with the existing PACK technique.

Keywords: Cloud Computing; TRE Algorithm; CORE; CSP, LAN.

1. INTRODUCTION

Cloud computing is a technology that uses data transmission technique assumes a vital part for any of different administrations. There are some administrations which are given by the Cloud computing innovation [1]. Cloud computing is intended to give benefits as pay according to utilize premise or can state on request benefits.

With cloud, there is no longer a need to build an expensive and high accessible repetitive framework, as one would need to do with a customary IT operational framework, in light of the fact that a bomb first attitude has been naturally incorporated with the structure of Cloud computing. The cloud technology was organized on the understanding that a specific part in the framework will give output.
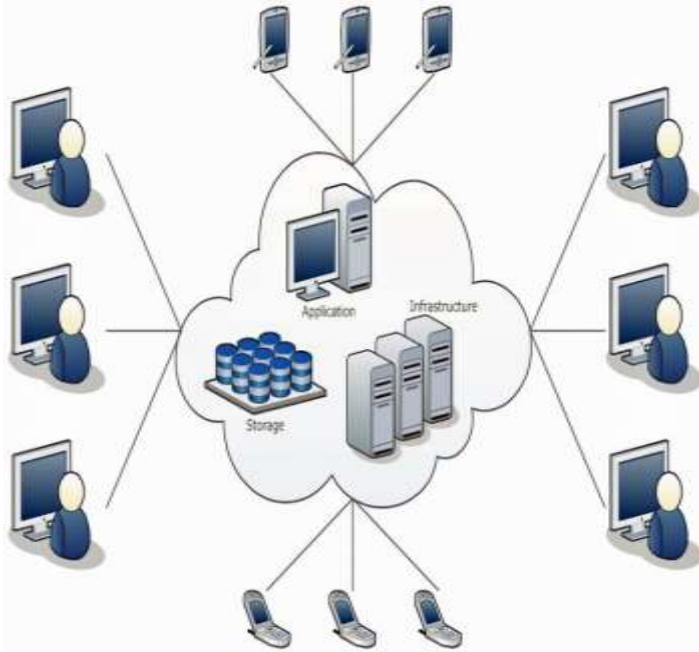
Figure 1.1: Cloud Computing Environment.

Cloud computing is a current mechanical advancement in the processing field in which principally centered around outlining of administrations which can be given to the clients in same path as the essential utilities like sustenance, water, gas, power and communication. In this innovation administrations are created and facilitated on the cloud (a system intended for putting away information called datacenter) and afterward these administrations are offered to clients dependably at whatever point they need to utilize. The cloud facilitated administrations are conveyed to clients in pay-per-utilize, multi-tenure, versatility, self-operability, on-request and practical manner[3]. Cloud computing is turned out to be well known in light of above say administrations offered to clients. Every one of the administrations offered by servers to users are given by cloud specialist organization (CSP) which is working same as the ISP (Internet specialist organization) in the web figuring. In the web make some creative improvement in virtualization and dispersed figuring and getting to of fast system with minimal effort draw in center of clients toward this innovation. This innovation is composed with the new idea of administrations provisioning to clients without buying of these administrations and put away on their nearby memory [4].

2. LITERATURE REVIEW

S. IHM, K. PARK AND V. PAI, "Wide-Area Network Acceleration For The Developing World (WANAX)," In Proc. USENIX ATC, 2010.

Traffic occurs between the sender and receiver by transmitting the same information repeatedly, that information may contain files, documents, video's etc. In recent years various TRE techniques have been developed for eliminating the redundant data. Wanax present, Wide-area network (WAN) accelerators operate by compressing redundant network traffic from point-t communications, enabling higher effective bandwidth. It uses multi resolution scheme [9].

EYAL ZOHAR, ISRAEL CIDON AND OSNAT MOKRYN, "PACK: Prediction-Based Cloud Bandwidth And Cost Reduction System", 2014

In this paper, the author depicts the current approach which is PACK Algorithm for traffic repetition elimination framework is displayed. PACK uses sender based approach and the receiver based approach for redundancy elimination. Here data is divided into small chunks

The receiver computes the respective signature for each chunk and looks for a match in its local chunk store. If the chunk's signature is found, the receiver determines whether it is a part of a formerly received chain, using the chunks' metadata. If affirmative, the receiver sends a prediction to the sender for several next expected chain chunks. Upon a successful prediction, the sender responds with a PRED-ACK confirmation message [10].

S. IHM, K. PARK, AND V. PAI, "Survey On Traffic Redundancy And Elimination Approach For Reducing Cloud Bandwidth And Costs", 2014

In order to removing redundancy we are using TRE approach. Large amount of popular content is transferred repeatedly across network links in the internet. To transfer the information between the sender and receiver data is divided into chunks. Chunking mechanism helps to improve the efficiency by parallel upload/download of different chunks and each chunk generates a signature to transfer data in secure manner [11].

LEI YU, HAIYAING SHEN, KARAN SAPRA, LIN YE AND ZHIPENG CAI, "CORE", 2007

The pay-as-you-go benefit show affects cloud clients to decrease the use cost of data transmission. Activity Redundancy Elimination (TRE) has been appeared to be a

2

viable answer for lessening transmission capacity expenses and hence has as of late caught huge consideration in the cloud condition. By concentrate the TRE methods in a follow driven approach, we found that both here and now (time traverse of seconds) and long haul (time traverse of hours or days) information access can simultaneously show up in the activity, and exclusively utilizing either sender-based TRE or recipient based TRE can't all the while catch both sorts of movement repetition. Additionally, the effectiveness of existing collector based TRE arrangement is powerless to the information changes contrasted with the chronicled information in the reserve. In this paper, we propose a Cooperative end-to-end TRE arrangement (CORE) that can distinguish and expel both here and now and long haul repetition through a two-layer TRE outline with agreeable operations between layers. A versatile forecast calculation is additionally proposed to enhance TRE productivity through powerfully altering the expectation window estimate in view of the hit proportion of recorded forecasts. In addition, we upgrade CORE to adjust to various activity repetition attributes of cloud applications to enhance its operation cost. Broad assessment with a few genuine follows demonstrate that CORE is able to do viably recognizing both here and now and long haul repetition with low extra cost while guaranteeing TRE effectiveness from information changes [12].

MISS. MANE VIDYA MARUTI, PROF. MININATH K.NIGHOT, "Authorized Data De-Duplication Using Hybrid Cloud Technique", 2015

Cloud computing is vital in the information sharing application. Day by day utilization of cloud is expanding. Be that as it may, the issue in distributed computing is each day information transferred on the cloud, so expanding comparative information in cloud. Along these lines it can be lessen the span of comparable information in cloud utilizing the information DE duplication technique. This technique primary point is that expel copy information from cloud. It can likewise spare storage room and data transfer capacity. This proposed technique is to evacuate the copy information yet in which client have appointed some benefit as indicated by that duplication check and every client have their special token. Cloud DE duplication is accomplished utilizing the cross breed cloud design. This proposed technique is more secure and devours less assets of cloud. Additionally it demonstrated that proposed conspire has insignificant overhead in copy evacuation when contrasted with the ordinary DE duplication method. In this paper Content Level DE duplication and additionally File Level DE duplication of record information is looked at over the cloud.

In this framework it likewise proposed new duplication check technique which produce the token for the private record and check content level DE duplication. Client need to present the benefit alongside the focalized key as a proof of possession. It fathomed more basic piece of the cloud information stockpiling which is just endured by various techniques. Proposed strategies guarantee the information duplication safely. Execution of this framework is 98 % more than existing framework [13].

3. PROBLEM IDENTIFICATION

There are different techniques with the cloud computing data storage, its applicability over the data centre, server and accessing is performed by different user. Multiple file upload and its usage make duplication of data over the server.

As the study is taken and performed with different techniques and different result from the algorithms were monitored such as Content based, Chunking based, Hash Based and other different technique for data processing, security approach over data store. The techniques for security over the cloud data is also performed by different services to make it more secure and accessible [23].

Upon verifying different scenario and the available technique different short comes with the Existing algorithm SHA-1 with file based de-duplication which is taken as base for our research work.

The following are the monitored points which identified as problem and further analysed and performed further with enhancements.

1. Previous technique such as file based scheduling doesn't over count all its data parts, or internal division which can further duplicate over the large amount of data. Thus an efficient monitoring is required which can further be monitor file duplicity with data division.

2. AES algorithm takes an advantage of asymmetric encryption technique which is used by base paper, but still when we talk about the multiple tenants, multiple ownership and multiple user over the data. Thus a security of key sharing is still a challenging issue which is faced by authors.

3. The Key length taken for the purpose of security in previous research is not considerable today . Today's scenario required an efficient and long length key for security purpose.

4. The existing approach for security uses MD5 for the hashing for content matching, but the MD5 algorithm faces collision issue with value generation. Hashing algorithm is the best practice to have long hash value.

5. A combination of MD5 and AES is taken for the consideration which is fighters more secure while talking about key exchange; again an extra procedure is required to do the key exchange. Thus it exhibit extra computational time as well as computation cost for cloud server.

6. A file level de-duplication algorithm by file hash MD5 is used, which can come under collision scenario and product false result when it terms to large number of server data files.

7. The existing algorithm take advantage over previous traditional techniques but still more refinements are required as per today's standard. Thus a better security, hashing mechanism can make it more reliable and executable to tackle with current security and cloud scenario in the world.

4. PROPOSED METHODOLOGY

As per our observation about the previous technique and their disadvantage in different terms and scenario's. Our work present a new approach which is highly secure and consumes low computational time and thus computational cost over the large number of structured available dataset.

Our work propose a new algorithm EPACK algorithm with more added secure algorithm HECC (Hyper Elliptic Curve Cryptography) is performed along with SHA-2 and Similarity measure score as more stable hashing approach. Our algorithm also checks for proper redundancy using more secure and reliable parameters.

HECC Algorithm For Security Intermediate Level:

The security of hyper elliptic Curve Cryptosystem depends on the discrete logarithm problem. This problem helps to avoid the eavesdropper from breaking of keys even both Q and P values are known publicly. Different types of curve have to study to understand about public key (Q), group point (P) and Hyper elliptic Curve Discrete Logarithmic problem (HECDLP).

Hyper elliptic curve E of genus $g >= 1$ over finite filed F is the set of solution $(x, y) \in F*F$ to the equation

$$E: y^2 + h(x) y = f(x) \qquad (1)$$

Where $h(x)$ is a polynomial of degree g and $h(x) \in F(x)$, $f(x)$ is a monic polynomial of degree $2g+1$ and $h(x) \in F(x)$. The curve E is said to be non-singular curve, if there are no pairs $(x, y) \in F*F$. The polynomial $f(x)$ and $h(x)$ are chosen such that it has to satisfy the following equations

$$2y + h(x) = 0 \qquad (2)$$

$$h'(x) y - f'(x) = 0 \qquad (3)$$

Algorithm Pseudo Code

Input: Data Packet, Receiver Address.

Output: Redundancy Evaluation, Result Evaluation.

Algorithm Begins:

Receiving user input packets();

If(data packet received)

{

Hecc Enc(data);

EnPacket transmission ();

Predict Attempting();

Meta property evaluation ();

SigSha2();

EndIf;

}

If(PRED-ACK)

{

observePred();

processPredACK();

4

Parameter computation & evaluation();

EndIf;

}

The figure below represent the complete flow of the proposed scenario which represent our work and computes parameters efficiently.

5. RESULT ANALYSIS

This section explains about the results which are obtained by applying existing and proposed algorithm.

Table 1.1: Statically Analysis Of Obtained Result.

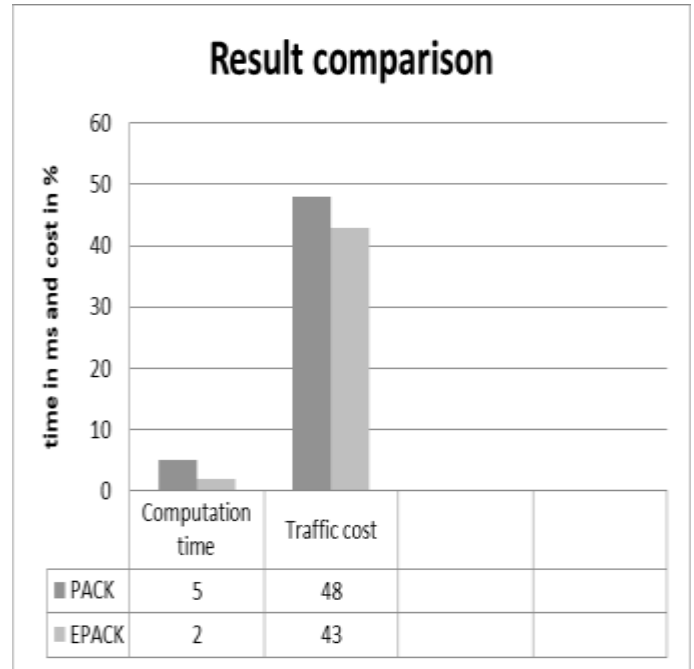| TECHNIQUE	EXISTING	PROPOSED EPACK
Encryption Technique	No Security	HECC Encryption over Packet transmission
Key Size	128	256
Hashing Approach	MD5	SHA-2
De-duplication	Content level	Unlimited block level chunking deduplication
Overall	AES+MD5+Content level de-duplcation approach	AES+ SHA2 + Categorial Similarity measure score approach

Figure 1.2: Comparison Line Graph For Technique Analysis.

Working on text based graphical algorithm such as our proposed technique authentication will be efficient and easy to visualize and in order to make it easier for user to use, such outlier technique will be efficient to use.

5. CONCLUSION & FUTURE WORK

Our proposed algorithm makes use of comparison parameter as computation time as well as computation cost to compute the comparison analysis. The algorithm is developer in Java language with Java net-beans tool setup using Intel i3 processor, 750 GB RAM. The comparison analysis and execution result shows that our proposed approach EPACK outperform best while comparing with existing algorithm.

A consistent proposed EPACK algorithm provides a high level security along side de-duplication approach with data store. There are still further work can be done to prove our work in knowledge and for industry use. The following works are left for future work.

1. The real time implementation can be done, which can apply over the industry level cloud infrastructure and

to find it more secure, reliable than the other alternate available over the web.

2. More study over the Hashing can be done, thus that an removal of that part can be done, which can make it more fast accessible.

3. A study of system can be derived and performed with different operating system with other file format values.

REFERENCES

1.   Jens, F. (September 2008). *Defining cloud services and cloud computing*.

2.   L. Wang, G. Laszewski, M. Kunze and J. Tao, ―Cloud computing: a perspective study‖, J New Generation Computing, 2010, pp 1-11.

3.   P. Mell and T. Grance, ―Cloud Computing Definition‖, National Institute of Standards and Technology, Version 15, 10-7-09.

4.   Mythry Vuyyuru, Pulipati Annapurna, K.Ganapathi Babu, A.S.K Ratnam, "An Overview of Cloud Computing Technology", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-3, July 2012.

5.   Satya Sudheer Varma Surimalla, Krishna. B,"Eliminating Redundancy in Cloud – Databases Using Authorized Hybrid Cloud Approach", International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online) Vol. 3, Issue 3, pp: (355-358), Month: July - September 2015.

6.   A. Anand, A. Gupta, A. Akella, S. Seshan, and S. Shenker, "Packet caches on routers: The implications of universal redundant traffic elimination," in *Proc. SIGCOMM*, 2008, pp. 219–230.

7.   PACK source code, 2011 [Online]. Available: http://www.eyalzo. com/projects/pack.

8.   A. Gupta, A. Akella, S. Seshan, S. Shenker, and J. Wang, "Understanding and exploiting network traffic redundancy," UW-Madison, Madison, WI, USA, Tech. Rep. 1592, Apr. 2007.

9.   S. Ihm, K. Park, and V. Pai, "Wide-area network acceleration for the developing world (WANAX)," in *Proc. USENIX ATC*, 2010, pp. 12–18.

10.   Eyal Zohar, Israel Cidon, and Osnat Mokryn, "PACK: Prediction-Based Cloud Bandwidth and Cost Reduction System", IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 22, NO. 1, FEBRUARY 2014.

11.   Ms. Pratibha A Kuber, Ms. Garima Makhija "HyPACK: Hybrid Prediction-Based Cloud Bandwidth and Cost Reduction System", International Journal Of Engineering And Computer Science ISSN:2319-7242 Volume 4 Issue 3 March 2015, Page No. 10707-10709.

12.   Lei Yu, Haiyaing Shen, Karan Sapra, Lin Ye and Zhipeng cai, "CoRE: Cooperative end to end traffic redundancy elimination for reducing cloud bandwidth cost",IEEE Transaction on parallel and distributed system, Vol 6, no. 1, January 2007.