PACK: Prediction-Based Traffic Redundancy Elimination System & provide secure encryption in Cloud

Amavi A. Vispute
Computer Science & Engineering
JSCOE ,Hadapsar
Pune, India
amuvispute@gmail.com

Prof.H.A.Hingoliwala.
Computer Science & Engineering
JSCOE ,Hadapsar
Pune ,India
Ali_hyderi@yahoo.com

***Abstract*—** In this paper, we use concept of PACK (predictive ACKs), which act like a traffic redundancy elimination (TRE) system, designed for cloud computing customers. TRE is designed on cloud to reduce traffic as well as cost regarding TRE Computation and storage will be optimized. The main advantage of the Pack Cloud-server is its ability to span end clients TRE effort, thus minimizing processing costs prompted by the TRE Algorithm. Unlike previous solutions Pack does not require server to continuously keep track on customer to maintain the status of the server.Pack maintain computing environment that combine server and client movement to maintain cloud elasticity. Pack is based on TRE technology; TRE is used to eliminate the transmission of redundant content as well as allow client to use newly received chunk to identify previously received chunks chains, which in turn can be used as reliable predictors future transmitted chunks.In our proposed work we are using encryption concept. we will send the chunks in encrypted format. For encryption we are using AES algorithm which is based on symmetric block cipher. This is using for security Purpose. We are going to secure our file from other traffics.

Keywords- *Cloud Computing, Traffic Redundancy Elimination, Predictive Acks, Network optimization, Secure Hash Algorithm-1,Advanced Encryption standard*

*****

I. INTRODUCTION

In cloud computing Environment, Cloud customers pay only for the actual use of computing resources, storage, and bandwidth, according to their changing needs, utilizing the clouds scalable and elastic computational capabilities. cloud customers, applying a judicious use of the clouds resources, are motivated to use various traffic reduction techniques, in particular traffic redundancy elimination (TRE)[1], for reducing bandwidth costs.

To the best of our knowledge ,no one previous works have been addressed the requirement for cloud computing-friendly,end-to-end TRE[2] which form PACK. TRE is used to eliminate unnecessary transmission of content and, therefore, Important to reduce network costs.Current End-To-End solution are sender based here cloud load balancing and optimization done on server side which require full synchronization between client and server.but there is lack of synchronization so lose efficiency.Most of its computational efforts on cloud side so less cost-effective.

In this paper We have presented pack, a receiver-based, Cloud-friendly, end-to-End TRE that is based on speculative fiction the theory is that the latency and reduce operating costs to maintain a consistent Server pack required Customer status thus enabling cloud elasticity and mobility, While long-term redundancy protection. Also, Pack Content-based access enables eliminating redundancy a three-way without having to implement multiple server clients the handshake.

1. To maintain Load balance at both server and Client
2. Improve efficiency
3. Reduce Computational Cost
4. Eliminate Redundancy
5. Create Active user friendly Environment at Cloud

The main contribution of this work is that in the proposed system,for provide much more security over network than SHA-1 algotithm[3] we are using AES cryptographic algorithm[4].AES is a symmetric block cipher it uses same key for both encryption and Decryption. This is using for security Purpose. We are going to secure our file,data from other traffics.

When encryption and Decryption perform on chunk size will be reduced so that it may reduce bandwidth cost and also required less buffered storage space.We are using encryption and Decryption technique for security purpose and in existing system we use SHA-1 algorithm which is not much resistance against attacker like Brute-force attack so we are using AES algorithm which having more resistance power to face attack over network.

Following are objective which is provided by AES algorithm:

1. Resistance against all known attack
2. Speed and Code Compactness on a wide range of platform
3. AESign simplicity
4. Block size and Key size can vary making algorithm versatile.
5. Easy to implement

The remaining of this paper is described as follows .We first define system related work in II section .In section III Proposed system model ,frame work ,In section IV we analyze Optimization.Then in section V,We identify more Secure_PACK. Then performance evaluation. Finally conclusion is given in last section.

II. RELATED WORK

A. WANAX

WANAX[5] is a flexible and scalable WAN accelerator targeting developing regions.In WANAX,MRC is used which is a Chunking technique that provides high compression and high throughput by maintaining a small memory footprint. an intelligent load shedding technique that exploits MRC to maximize effective bandwidth by adjusting disk and WAN usage as appropriate a mesh peering protocol that exploits higher speed local peers when possible instead of fetching only over slow WAN links. Disadvantage in WANAX is Hardware installation cost is very expensive.

B. A Low-bandwidth Network File System(LBFS)

Benjie chen and and David Mazieres are proposed[6] LBFS which is a network file system that saves bandwidth by taking advantage of commonality between files. LBFS breaks files into chunks based on contents, using the value of a hash function. It indexes file chunks by their hash values. Advantages of LBFS are it avoids sending redundant data, Require magnitude less bandwidth and indexing help to reduce redundancy. Disadvantage is not suitable for application which require very High bandwidth.
Eg.video, 3D video etc.

C. SmartRE

K. C. Lan and C. M. Chou invent a SmartRE[7] is An Architecture for Coordinated Network-wide Redundancy Elimination. It provides a naive link-by-link view and adopts a network-wide coordinated approach. It is suitable for handling heterogeneous resource constraints and traffic patterns and for incremental deployment. They address several practical issues in the design to ensure correctness of operation in the presence of network dynamics.Advantages are it enable more effective utilization of the available resources at network devices, can apply to Datacenter and MultiHop wireless network .Disadvantage is It having designing problem in Dynamic network model

D. Redundancy in Network Traffic: Findings and Implications

Ashok Anand, Chitra Muthukrishnan, Aditya Akella and Ramachandran Ramjee[8] found out the various issues in network design while thinking about redundancy elimination.They shows that packet-level redundancy elimination techniques can deliver average bandwidth savings of 15-60 for enterprise at packet traces collected at twelve distinct network vantage points and also links connecting busy web servers.They also identify that overall traffic was not completely reduced so that peak traffic periods was variable.they also identify that a client-server redundancy elimination solution could provide approximately similar savings as a middlebox.

The comparisons above discussion techniques are shown in Table1

Sr. No	**Topic Name**	**Technique**	**Advantages**	**Disadvantages**
1	WANAX	1>Operated By Compressing Redundant Network Traffic Multi Resolution Chunking 2>MeshPeering Protocol 3>Intelligent load shedding technique	1>High throughput 2>High Compression 3>Higher speed local peers	hardware installation cost is very expensive.
2	LBFS	1>designed for Low bandwidth network. 2>Indexing at client & server 3>Protocol based on NFS vr.3	1>Avoid sending redundant data. 2>Require magnitude less bandwidth 3>keep track on file status	Not suitable for high bandwidth
3	SmartRE	1>Application-independent Redundancy Elimination (RE) 2>Based On 2 idea:-1.packet cache for redundancy elimination 2.Csamp Uses Network-wide optimization Frameworks	1>Allow more utilization of resourcs 2>Manage heterogeneous resource constraints traffic patterns	design problem in Dynamic network model
4	Redundancy in Network Traffic: Findings and Implications	1>Based on Rabin fingerprint algorithm 2> Redundancy in n/w traffic eliminated 2 way 1.Redundancy suppression 2.Data Compassion	1>75-90% middlebox bandwidth saving 2>Redndncy in n/w packet can eliminated	Enterprise traffic was not reduced.
5	PACK: Prediction-Based Cloud Bandwidth and Cost Reduction System	1>end-to-end traffic redundancy elimination (TRE) system 2> PACK algorithm 3>SHA-1 Algorithm	1>To maintain Load balance at both server and Client 2>Improve efficiency 3>Reduce Computational Cost 4>Eliminate Redundancy	1>There is a security problem 2>SHA-1 is slower computational algorithm 3.SHA-1 can be broken by Brute-Force Attack

III. PROPOSED SYSTEM MODEL

We have presented pack, a receiver-based, Cloud-friendly, end-to-end TRE that is based on speculative fiction the theory is that the latency and reduce operating costs to maintain a consistent Server pack required Customer status thus enabling cloud elasticity and mobility, While long-term redundancy protection. Also, Pack Content-based access enables eliminating redundancy a three-way without having to implement multiple server clients the handshake.

PACK Algorithm along with (Cryptographic algorithm)
Following is step that shows how algorithm works

1. At PACK receiver side ,stream of data received which is parse in sequence of variable size.

2. chunk are then compared to receiver local storage also called chunk store. If matching chunk is found in local chunk store, receiver retrieves sequence of chunk referred as chain which follow LRU scheduling.

3. Using constructed scheduling, receiver send prediction to sender for subsequent data. Prediction sent by receiver include predicted data, hint and signature of chunk.

4. sender identifies predicted range in its buffered data and verifies Hint for range, If result matches the received Hint, it continue to perform the more computationally SHA-1 signature operation.

5. Upon signature match sender send a confirmation message to receiver.

In this message transmission TCP wire protocol is used which help to reduce redundant data and receiver identifies that currently received chunk is identical to a chunk in its chunk store. for making more secure transmission AES cryptographic algorithm may apply so that transmission become more secure against attacker.

.

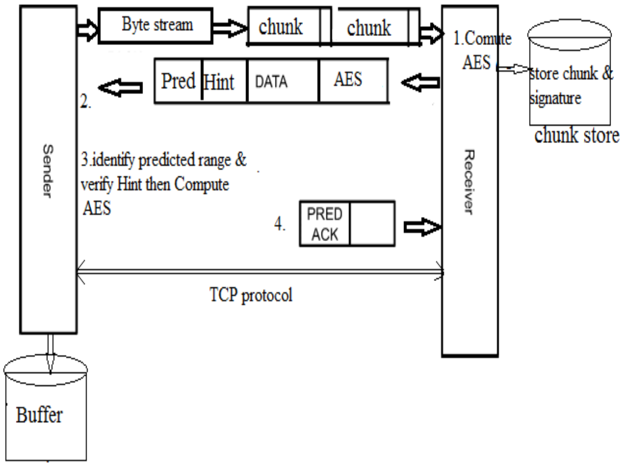

Figure 1.Working of PACK Algorithm

IV. OPTIMIZATION

For more clarification, we have to describe the additional options and optimization.

A. *Adaptive Receiver Virtual Window*

In this process,we terms the receiver fetching local data as a virtual data. It means the advanced algorithm performed at the receiver's side.PACK behavior when a data segment arrives after its prediction was sent and the virtual window is doubled.the reception of a successful acknowledgement message (PRED-ACK) from thesender. The receiver reads the data from the local chunk store. It then modifies the next byte sequence number to the last byte of the redundant data that has just been read plus one, and sends the next TCP ACK, piggybacked with the new prediction. Finally,the virtual window is doubled. The size increase of the virtual window introduces a tradeoff in case the prediction fails from some point on. then receiver's behavior when the arriving data does not match the recently sent predictions. The new received chunk may, of course, start a new chain match. Following the reception of the data, the receiver reverts to the initial virtual window until a new match is found in the chunk store.

B. *Cloud Server as a Receiver*

Because of metadata cloud storage is becoming a dominant player from backup and sharing services. In many of these services, the cloud is often the receiver of the data. If the sending client has no power limitations, PACK can work to save bandwidth on the upstream to the cloud. In these cases, the end-user acts as a sender, and the cloud server is the receiver. The PACK algorithm need not change. It does require, however, that the cloud server like any PACK receiver maintain a chunk store.

V. SECURE_PACK

In this paper, a security maintained receiver driven operation of Predictive acknowledgement protocol is described. The incoming stream of data received at receiver side is parsed to a sequence of variable-size, content based signed chunks. The comparison between the incoming chunk data and the previously arrived chunks which is present in the local storage termed as chunk store takes place. If a matching chunk is found in the chunk store, the receiver retrieves sequence of subsequent chunks referred to as a chain. This is done by traversing the sequence of LRU chunk pointers that are included in the chunk's metadata. Thus a chain of matched data will be constructed and using this chain the receiver sends a prediction to the sender for subsequent data. Part of each prediction termed as a hint, is an easy-to-compute function with a small enough false positive value. The prediction sent by the receiver which is present in the prediction queue includes the length of predicted data, the hint and the hashed value of the predicted data. The data owner identifies the length of the data and verifies the hint. If the result matches the received hint, it continues to perform SHA-1 [1] signature operation.

Depending on the signature match, data owner sends a confirmation message to the receiver, thus permitting it to copy the matched data from its chunk store. The system uses a new chunk chain scheme, in which chunks are linked to other chunks according to their last received order. To efficiently maintain and retrieve the stored chunks, their predicted chunk chain, caching and indexing techniques are used. Each chunk's signature is computed using SHA-1 when the new data are received and parsed to the chunks. At this point, the chunk is added to the chunk store. Thus the newly received chunk is placed after the previously received chunk in a least recently used manner. After the identification of the non redundant chunk, encryption of data takes place using AES.

A. *Receiver Algorithm*

The storage overhead is special issue of the access control scheme in cloud storage system. We compare storage overhead on each entity in system. When the new data arrives, respective signature for each chunk is computed by the receiver and then match is being looked out in its chunk store. If the chunk's signature is found, receiver finds out whether it is a part of previously received chunk chain using the chunk's metadata. Thus the receiver sends a prediction to the data owner indicating the next expected chunk chain. The prediction contains a starting point in the byte stream, the total length of the chunk and identity of subsequent chunks. Fig 2 illustrates the receiver operation and describes how the data gets encrypted

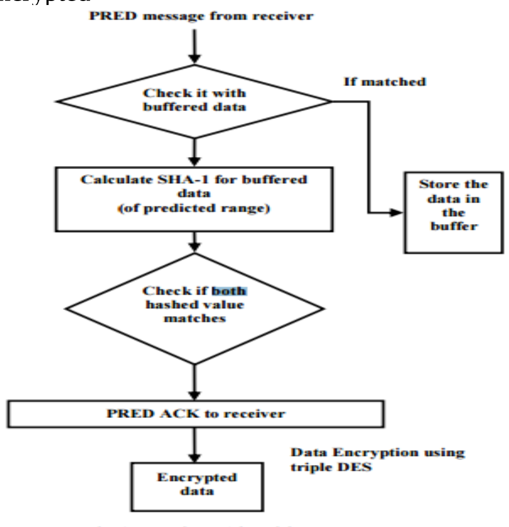
PRED message from receiver
Check it with buffered data
If matched
Calculate SHA-1 for buffered data (of predicted range)
Store the data in the buffer
Check if both hashed value matches
PRED ACK to receiver
Data Encryption using triple DES
Encrypted data

Figure 2. Receiver Algorithm

Upon a successful prediction, data owner sends a PREDACK confirmation message. Ones receiver confirms that the chunk is not redundant, the resulted data is encrypted using the triple AES algorithm and the encrypted data is sent to the cloud server by the receiver .Thus this improves the security level. The receiver copies the corresponding encrypted data from the chunk store to its cloud's buffer after the reception of the PRED-ACK message. If the data chunk is a redundant content, then the corresponding data ID is sent to the cloud. Thus traffic is avoided. At this point, receiver sends a TCP Acknowledgement with the next expected TCP sequence number.

B. Data Owner Algorithm

After the reception of PRED message from the receiver, it compares the prediction message with the data present in the data owner side. The data owner determines the range and verifies the hint for each prediction. Upon a hint match, the SHA-1 signature for the predicted data is being found out and the result is compared with the signature in the PRED message. If the both SHA-1 signature matches, data owner confirms that the receiver's prediction is correct. Thus the data owner sends a PRED-ACK message to the receiver. If suppose the hint does not match, then a computationally expansive operation is saved, which is thus used as a future predictor. The non redundant data is encrypted using triple AES algorithm and is sent to the cloud server. Fig 3 illustrates the sender operation and describes how the sender tries to match a predicted range to its outgoing data.

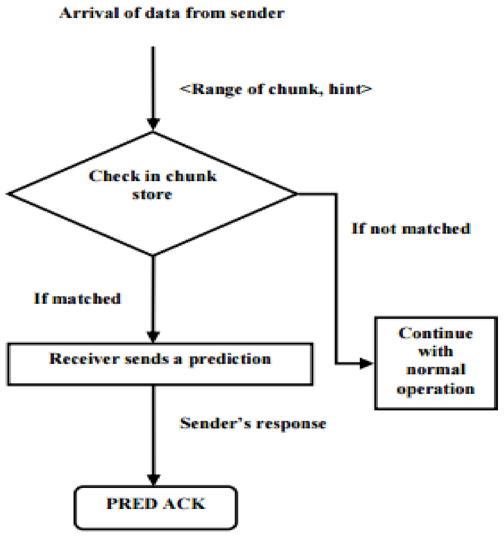

Figure 3. Data Owner Algorithm

C. Cloud Server Module

Cloud infrastructure is a "pay-as-you-go model". Data owner stores their applications on cloud. The delivery of computing service is over the internet. The data to be stored in the cloud is encrypted using the triple AES algorithm,which makes the data more secure. The procedure for encryption is exactly the same as regular AES ,except that it passes three times through the AES engine. The first pass is a AES encryption , the second pass is a AES decryption of the first AES Ciphertext result and the third pass is a AES encryption of the second pass result. This produces the resultant Triple AES Ciphertext. The encrypted data is finally placed in the cloud. Cloud server module can view the registered clients and their transaction details. Due to the existence of predictive acknowledgement, redundant traffic is eliminated. Thus non-redundant data is only being stored in the cloud. This reduces the cloud bandwidth , the overall operational cost and the security is also being maintained.

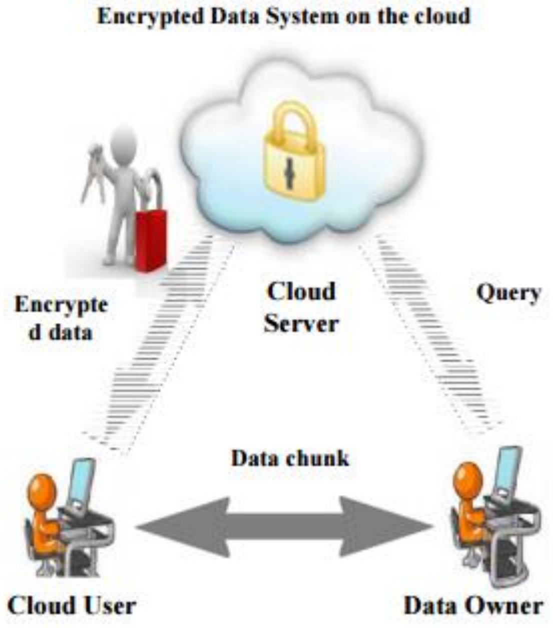

Figure 4. Secure_PACK

Fig. 4 shows the system architecture of the Secure_PACK . After the non redundant data is being identified, the data is encrypted using the AES algorithm and is sent to the cloud server for storage. Since in cloud computing a distributed computing takes place, it is not secure to place the raw data in the cloud. Hence for maintaining security, data is encrypted using triple AES . Thus specific data owner can only view his data , which created privacy. In the previous work , even though bandwidth and cost were reduced , security level was not at all maintained. In our work the security level is maintained thus this overcomes the disadvantage of the existing system.

VI. PERFORMANCE EVALUATION

The section AEScribes about the experimental evaluation. The AES and SHA-1 are used to show the experimental results. Advantages

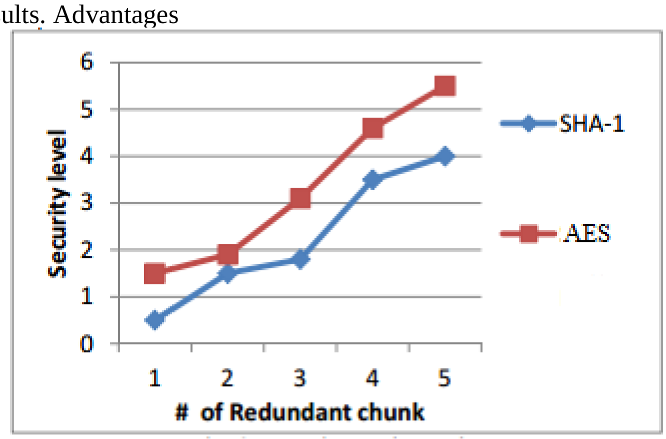

Figure 5.Experimental Result

The fig. 5 shows that security level increases even when the number of redundant chunk increases. The blue line shows the performance of the system when SHA-1 alone is applied. The red line shows the performance of the system when AES is used. In the proposed system triple AES is being used for encrypting the data, hence security is maintained compared to the existing system. It clearly shows that the proposed system performs better than the previous methods in terms of security level.

VII. CONCLUSION

The Traffic redundancy eliminate over network.TRE is also used to Proprietary middle box solution inadequate that reduces a growing cloudy needs is operational the cost accounting application latencies, while user dynamics, and elasticity. The main advantage of the Pack Cloud-server is its ability to span end clients TRE effort, thus minimizing processing costs prompted by the PACK Algorithm Limitations is that there is a security problem while sending a data in chunk for over a network so for solving this problem AES cryptographic algorithm which provide much more security against attacker .

ACKNOWLEDGMENT

I would like thanks to Prof. H. A. Hingoliwala professor of Computer Engineering at J.S.P.M .Hadapsar ,who guided through this paper.

REFERENCES

[1] E. Zohar, I. Cidon, and O. Mokryn, The power of prediction: Cloud bandwidth and cost reduction, in Proc. SIGCOMM, 2011, pp. 8697.

[2] N. T. Spring and D. Wetherall, A protocol-independent technique for eliminating redundant network traffic, in Proc. SIGCOMM, 2000, vol. 30, pp. 8795.

[3] Suresh Anak Agung Putri Ratna, Ahmad Shaugi, Prima Dewi Purnamasari, “Analysis and comparison of MD5 and SHA-1 algorithm implementation in authentication based security system”, 2013 International Conference on Computer Science and Electronics Engineering, pp. 99-104

[4] Joan Daemen and Vincent Rijmen, "The Design of Rijndael: AES-The Advanced Encryption Standard," Springer-Verlag, 2002

[5] Sunghwan Ihm, KyoungSoo Park, and Vivek S. Pai, Wide-area Network Acceleration for the DevelopingWorld:, in Department of Computer Science,Princeton University.

[6] Athicha Muthitacharoen, Benjie Chen, and David Mazieres, A Lowbandwidth Network File System.

[7] K. C. Lan and C. M. Chou, SmartRE: An Architecture for Coordinated

Network-wide Redundancy Elimination

[8] Ashok Anand, Chitra Muthukrishnan, Aditya Akella and Ramachandran Ramjee, Redundancy in Network Traffic: Findings and Implications.