PCS :PACK BASED CO-OPERATIVE END TO END TRE TECHNIQUE WITH SCTP TO REDUCE CLOUD BANDWIDTH

K.S.R College of Engineering
Thokkavadi
Tiruchengode 638 215
Mr.C.Thirumalai Selvan M.Tech.,[Ph.D.] ctsn027@gmail.com
Miss.B.Deepthi, optimisticdeepthi@gmail.com
Miss. R.SharmilaManju, greatmanju93@gmail.com
Miss. D.Thamilarasi, tamilarasi.d27@gmail.com
Miss.M.Umamaheswari, umaksrce@gmail.com

ABSTRACT:

Cloud computing technology provides immense benefits for the cloud customers, since it is a pay as you go service as well as a measured billing service. Cloud customers suffer due to high traffic in the internet, these traffic arise due to client activities such as repeated accessing of a page, or often uploading or downloading. To reduce the cloud traffic there are many TRE techniques available. Even though the PACK is a latest technology, along with TCP protocol, it has the greatest disadvantage over the system is it does not synchronized fully and the time will also be increased. To overcome this we use a proposed system with the new computational technique named co-operative end to end TRE along with SCTP protocol to solve those problems more effectively. Finally we analyze the cooperative end to end technique along with the SCTP protocol benefits the cloud users.

KEY TERMS: caching, Traffic redundancy elimination, chunks, chunk store.

INTRODUCTION: Cloud computing is the emerging technology which is a utility based computing which means the payment is paid only for the usage. The cloud customers can use the actual computing resources, storage, and bandwidth, based on their needs, thus they can use the cloud scalability, cloud elastic computational capabilities[2]. There are several parameters available in cloud environment they are efficiency, speed, cost, bandwidth, time, performance among these parameters data transfer costs i.e. Bandwidth of cloud plays a major role in order to reduce the cost. Generally traffic arises due to replicated or reduplicated request from the client side. These replicated request cause Congestion thus the cost of the system increased. The access privileges for each user is set through ACL(access control list) provided by the cloud server. There are several TRE techniques are used in previous systems they are EndRE,protocol independent technique, server side TRE technique, Rabin finger printing scheme. Those techniques do not give better results for mass chunk redundancy content. In most of the TRE solutions, the sender and the receiver will scrutinize and correlate the unique identification of the data chunks. While checking if they encounter a redundant data chunk, the sender will restore the transmission of replicated chunk with the strong signature. If the server or client receives more chunks the cloud elasticity will adapt to changes by provisioning and deprovisioning chunks. A Cloud is an upcoming technology mainly in the enterprises network.

In normal network transfer they use middle boxes as the redundancy elimination protocol, but in the cloud environment the normal middle boxes does not support, so we use proprietary protocol in both server and client side. Thus this protocol can reduce the traffic redundancy in the initial level. In middle boxes the state of synchronization is not much efficient when compared to proprietary protocol. The biggest disadvantage of the middle boxes is it can be used only in the intranet surrounding, thus it does not support the cloud environment. After middle boxes there was an another protocol namely proprietary middle boxes, they are considered as the popular entry point solutions within the enterprise networks, and it is also not so efficient like proprietary protocol in the cloud environment. Thus the customers are not satisfied because of these protocols, so they move on with the proprietary protocol for the benefits of the cloud providers whose goal is to reduce the cloud bandwidth cost. The TRE solutions will maintain client status in a regular basis. Obviously, a TRE solution that places most of the computational effort on the client side this leads cloud to be less cost effective one. Our experiment shows that current end to end TRE solution does not require client to maintain server status. In our system we include co-operative end to end TRE solution that focused on

eliminating redundant chunks. In this solution, every receiver notifies the upcoming stream and matches the current chunks with the previously received chunks to check whether it is a part of local file. A pointer is maintained for a sequence of a stream of chunks.

On receiving upon chunks, the chunk signature is identified using SHA-1 algorithm. The corresponding chunk signature is stored in Meta data and the chunk data is stored in chunk store. Here in Meta data table we include a concept of LRU fashion. The sender often examines the hint of the received chunk with the previous chunk if the hint of the chunks match then the system will performs the TRE operations. To reduce the search time of the Meta data table we use a concept based on time, i.e. here the Meta data table is split in to two halves termed as long term chunk table and short term chunk table. These techniques can easily detect the required chunk for the table and this will reduce the searching time. In out experiment data loss or chunk loss will be finding in the initial stage.

In the current system TCP plays a major role in transferring chunks [1]. Here it transfers the data chunks in a single stream, thus if the flow of the TCP layer is disconnected, it will take time to establish the connection and transfer the chunks. To overcome this issue we are stepping up in to the next generation of the TCP protocol named as Stream Control Transmission Protocol (SCTP), here this protocol serves for the multi streaming concept. Compared with TCP and UDP, SCTP provides additional security. SCTP will reduce the data loss or packet loss, and when the packet loss is more the SCTP loss percentage will be less when compared to TCP. Mainly we are moving to SCTP for reducing the cloud Bandwidth, so that we can reduce the cost factor for cloud users. Presently SCTP is implemented within Firefox and apache to add SCTP; it requires upgrading the IP stacks and applications.

RELATED WORK:

There are several TRE techniques available. Some of them are EndRE, protocol independent technique, server side TRE technique, Rabin finger printing scheme. EndRE (End-System Redundancy Elimination Service) is a way of approach, which eliminates Redundancy in redundant network system. It is mostly used in Enterprises network [1]. Identifying and removing redundancy is typically accomplished by two steps Finger printing [2], matching and encoding. The main drawbacks of this EndRE technique is, it does not supported in cloud environment rather it has been used in normal network transfer[4].

Protocol independent technique, which is used for eliminating redundant traffic in the network. As redundancy elimination techniques become more widely deployed and more tightly integrated into network infrastructure and protocols, it becomes crucial to understand the benefits, trade-offs and design issues in these systems. This technique generally eliminates the web caching and data compression [8]. This technique has the greatest disadvantage over networks since it is not used in cloud because it has least memory storage. Internet and intranet traffic has been evolved to the shipment of large application files and rich data content. This shift has motivated the development and deployment of proprietary, middlebox based TRE solutions that address the need of large corporations. A major progress in handling redundant data, followed several seminal papers that set the foundation of modern TRE.

Rabin fingerprinting is a scheme based on arithmetic modulo. It contains an irreducible polynomial with coefficients. Data processing speed is less when compared to other TRE schemes [3]. PACK is meant for predictive acknowledgement and it has been the recent technique of TRE technique. It mainly concentrates on pervasive computing than cloud computing, this is a major drawback of the current system. PACK generally deals with chunks transfer in a transport layer. Normally chunk will be present in a third layer. For transferring the chunks it uses TCP protocol along with the proprietary algorithm, since the middle box pairs cannot be used in cloud. PACK is fully a received based, since the sender does not put more efforts so that the receiver process will be substantially increased. It does not support for the mobile devices but CORE technique supports mobile devices because it has both sender and receiver based applications.TCP present in TRE technique does not support fault tolerance. PACK along with the TCP protocol does not provides more efficiency for data loss when compared to SCTP protocol. On examining the previous TRE techniques PACK yields a better results. But PACK is less efficient when compared to upcoming TRE technique.

TRAFFIC REDUNDANCY:

Traffic is the major issue that has been occurred while transmission of data. This has been occurred due to the repeated request from a same client to the different servers for a several times. This repeated request results the system to hang as well as attackers can attack the system when the traffic is

increased. This may leads to packet loss. May some packets have the life time .it may lose due to the heavy traffic over the network It tends to discard the packet. So that the packet or data or chunk loss will be increased. If the receiver receives the chunk more than its capacity then congestion will pay the way for the traffic to occur. For e.g. checking results in the portal during the publishing time. In this situation the normal network will not have the efficient outcome. But if we consider the cloud, it can provide various services to the server, so that it can reduce the traffic.

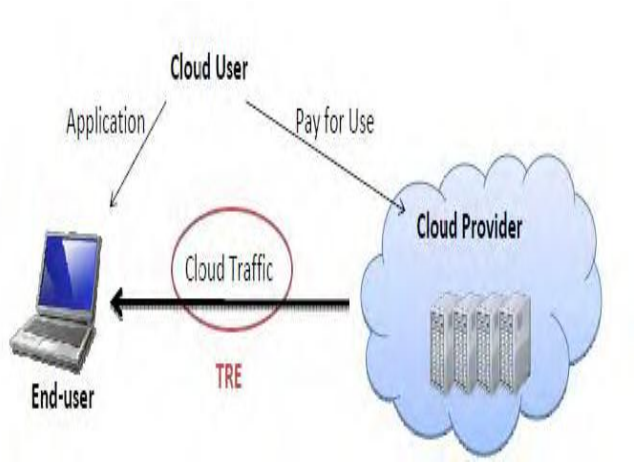


Fig: 1 Cloud TRE Environment

This shows how the traffic encounters in the cloud computing between end user and the cloud server by transferring the same content repeatedly. 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 the data in a secure manner.

Network traffic exhibits large amount of redundancy when different users on the Internet access same or similar content. Many diverse systems have explored how to eliminate this redundant content from network links and improve network efficiency. Most of the large scale industries or organizations like Amazon, Google are now a day's they are moving to cloud. In the current system the TCP is the major layer 4 protocol for transferring the data chunks here the data chunks will be transferred through the transport layer. It transfers the data chunk in the single stream. TCP is a connection oriented protocol so that the chunks will be transferred in a secured way[10]. The reliability of the system will also high. TCP generally transfer the data in the byte stream. The flow control of the TCP protocol is high since it is a connection oriented protocol. This protocol can eliminate the congestion control occur on the system while transferring the data. Data is delivered in the concept of stream of bytes[4][5]. TCP deals with some disadvantages such as all the work is done by operating system, if there is any bugs in operating system we have to tackle the problems, as surfing and downloading the content from the internet. It cannot be used for broadcast and multicast connection. If the system receives extra loads the transmission gets slower.

STREAM CONTROL TRANSMISSION PROTOCOL :

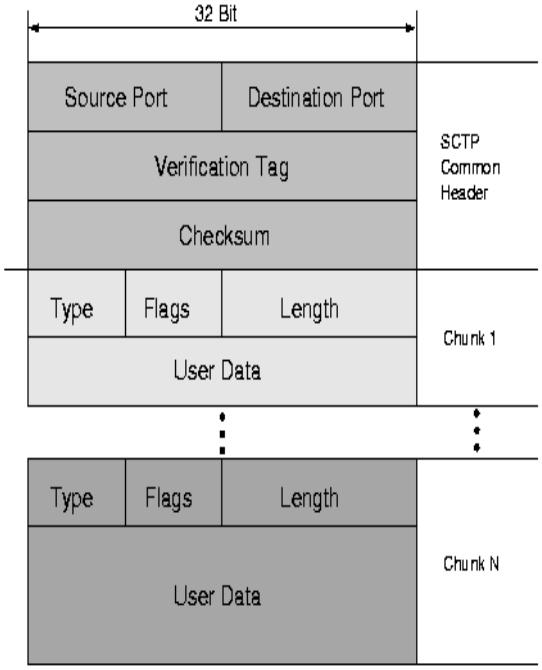
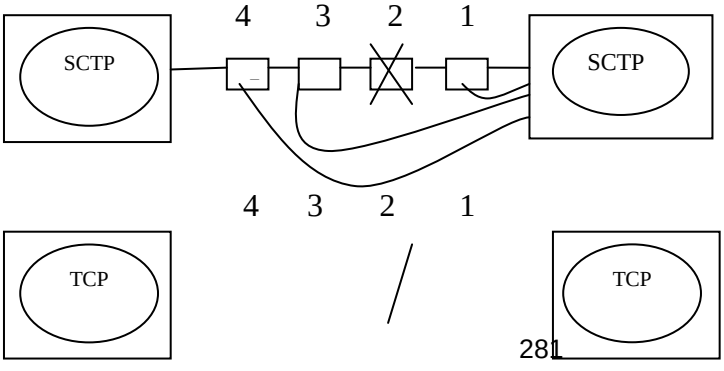


Fig.1 SCTP packet format

SCTP is the stream control transmission protocol which sends the data in multi streaming concept. SCTP is considered as the next generation of the TCP protocol. It is the unicast protocol that supports the end to end data delivery exactly within the two end points; the end points may have more than one IP addresses.It is a reliable transport protocol operating on top of a connectionless packet network such as IP (same level as TCP)[7]. It is a full duplex transmission protocol with some features like retransmission, flow control and sequence maintenance.



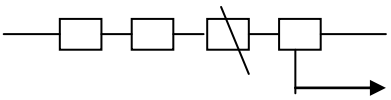


Figure 2-A head of line blocking concept in SCTP and TCP

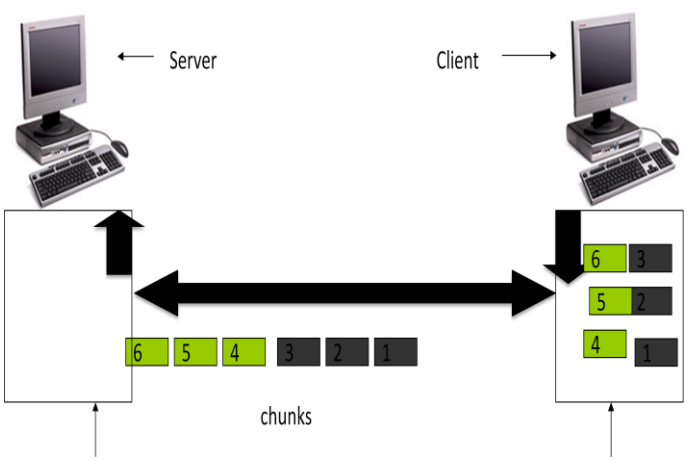
In TCP, if any data chunk gets lost means then the entire stream will get lost. As in the case of SCTP, it allows for unordered data delivery and since it has multiple streams, only the one affected is temporarily blocked and sends the remaining data chunks with an improved security than TCP.

Protection against denial of service: SCTP connection setup involves 4 messages (unlike 3 for TCP) and avoids depositing any state at the "called" endpoint until it has ensured that the other end is genuinely interested in setting up an association. This makes SCTP less prone to DoS attacks. This feature may not be very useful *inside* the data centre – however, since most connections within the data centre are long-lived, the additional overhead of this mechanism is not detrimental. SCTP has the ability to selectively acknowledge receipt of missing, disordered, or duplicated messages. SCTP is originally designed to support SS7 messaging system over IP networks[9].

We propose a novel end to end sender and receiver based Cooperative TRE solutions (CoRE) for efficiently identifying and removing both short-term and long-term chunks redundancy. Through a two-layer redundancy detection design and one single pass algorithm for chunking and fingerprinting, CoRE efficiently carries out cooperative operations between the sender and the receiver. On the evaluation with several real traces, we show that CoRE is able to identify both short-term and long term redundant chunk with optimized cost.This cooperative end to end technique is used in meta data table which is present in both sides of the sender and receiver.It is used for reducing traffic redundant chunks.

CO OPERATIVE END TO END TECHNIQUE:

Cooperative end-to-end TRE solution (CoRE) for efficiently identifying and removing both short-term and long-term redundancy. Through a two-layer redundancy detection design and one single pass algorithm for chunking and fingerprinting, CoRE efficiently carries out cooperative operations between the sender and the receiver.



SCTP based meta data table

SCTP based meta data table

Fig 3. CO OPERATIVE END TO END TECHNIQUE

To assess the bandwidth saving potential of long-term caching, we measured the social network redundancy while changing cache sizes at the client-side. We set different cache sizes by changing the maximal age of data kept in the client's storage. The redundancy level as a function of cache sizes. The upper line represents the redundancy for TRE schemes that identify returning clients despite changes of their IP addresses. The lower line is the upper bound for TRE schemes that do not track returning clients that change their IP addresses (e.g., [4][2]). Clearly, the larger the client's cache is, the larger is the amount of traffic that can be eliminated.

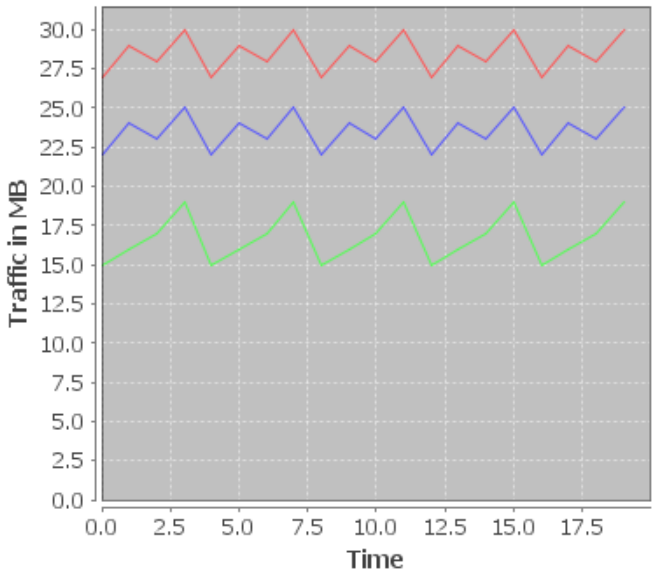
The Cooperative end to end TRE receiver maintains a chunk store, which is a large size cache of chunks and their associated metadata. Chunk's metadata includes the chunk's signature and a (single) pointer to the successive chunk in the last received stream containing this chunk. When the new data are received and parsed to chunks, the receiver computes each chunk's. At this point, the chunk and its signature are added to the chunk store.

This model for understanding steady-state cache behavior in a bandwidth-constrained long-term prefetching environment. We then present and evaluate a long term prefetching policy based on both chunks request rates and lifetimes. Our hypothesis is that by prefetching chunks that have reasonably high probabilities of access before they are updated-those that are both long-lived and popular, we can significantly improve hit rates for moderate bandwidth costs. We define the prefetch quotient of an chunk as the probability that it would be accessed before it is updated. We place a

threshold on the prefetch quotients of chunks and select them for prefetching if they cross the threshold.

COMPARISION GRAPH :

This graph is implemented for comparing the TRE techniques along with their protocols.



Co-pack PACK SCTP

CONCLUSION:

The cloud computing triggers the massive demand on the TRE solutions, as the amount of data transfer is also increased in cloud environment. Here we use an advanced TRE technique to reduce the cloud transmission. This will reduce the cloud cost to the cloud customers.

In this paper we present PCS, a novel both sender and receiver based Cloud friendly protocols and algorithms. To reduce the latency and cloud operational cost. Here PCS server does not maintains the client status continuously. The cloud elasticity and user mobility are the added advantage to the cloud customers.

FUTURE ENHANCEMENTS :

The SCTP protocol used in this system is implemented only for low bandwidth range and in future it can be extended for higher bandwidth range. So far we have seen many TRE techniques and we have use advance TRE techniques. Further this project can be implemented by using SCTP protocol along with the upcoming TRE techniques.

REFERENCES :

[1] E. Zohar, I. Cidon, and O. Mokryn, "PACK: Prediction-Based Cloud Bandwidth and Cost Reduction System. IEEE/ACM Transactions on Networking, 01 February 2013.
[2] E. Zohar, I. Cidon, and O. Mokryn, "The power of prediction: Cloud bandwidth and cost reduction," in *Proc. SIGCOMM*, 2011, pp. 86–97 .
[3] Z. Broder, "Some applications of Rabin's fingerprinting method, "in Sequences II: Methods in communications, Security, and Computer Science. New York, NY, USA: Springer-Verlag, 1993, pp. 143–152.
[4] B. Aggarwal, A. Akella, A. Anand, A. Balachandran,P. Chitnis, C. Muthukrishnan, R. Ramjee, and G. Varghese.EndRE: An End-System Redundancy Elimination Service for Enterprises. In *Proc. of NSDI, 2010*.
[5] A. Anand, C. Muthukrishnan, A. Akella, and R. Ramjee, "Redundancy in network traffic: Findings and implications," in Proceedings of the eleventh international joint conference on Measurement and modeling of computer systems. ACM New York, NY, USA, 2009, pp. 37–48.
[6] P. Kulkarni, F. Douglis, J. LaVoie, and J. Tracey, "Redundancy Elimination Within Large Collections of Files," in USENIX Annual Technical Conference, General Track, 2004, pp. 59–72.
[7] A. Jungmaier et al., "SCTP: A Multi-link End-to-End Protocol for IP-based Networks," *Int'l J. Electronics and Comm.*, vol. 55, no. 1, Jan. 2001.
Randall Stewart is a technical leader with Cisco Systems.
[8] N. T. Spring and D. Wetherall, "A protocol-independent technique for eliminating redundant network traffic," in Proc. SIGCOMM, 2000, vol. 30, pp. 87–95.
[9] R. Stewart et al., "Stream Control Transmission Protocol,"IETF RFC 2960, Oct. 2000; http://ietf.org/rfc/rfc2960.txt.
[10] P. Amer, S. Iren, and P. Conrad, "The Transport Layer: Tutorial and Survey," *ACM Computing Surveys*, vol. 31, no. 4, Dec. 1999.