International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 7 Issue 02, February-2018

TRE Approach to Reduce Cloud Bandwidth and Cost Based on Prediction

[1]Dipali N. Kalantri, [2]Priyanka R. Watekar,
[3]Manaswi R. Devtale, [4]Nikesh R. Wasekar
B.E. Student
Department of CSE,
Ballarpur Institute of Technology,
Ballarpur,Chandrapur District, Maharashtra, India.

Prof. Payal Badole
Assistant Professor
Department of CSE,
Ballarpur Institute of Technology,
Ballarpur, Chandrapur District, Maharashtra, India.

***Abstract*—Cloud computing offers the customers an economical as well as convenient pay-as-you-go service model, also known as usage-based pricing. Cloud customers pay only for the actual use for computing resources, storage, and bandwidth, according to their changing needs, utilizing the cloud's scalable and elastic computational capabilities. In particular, data transfer costs (i.e., bandwidth) is also an important issue when trying to minimize costs. Consequently, cloud customers applying a judicious use of the cloud's resources. Also, are motivated to use various traffic reduction techniques, in a particular traffic redundancy elimination (TRE), and for reducing bandwidth costs. Traffic redundancy are stems from common end users activities, such as repeatedly accessing, downloading, uploading (i.e., backup), distributing, and modifying the same or similar information items (documents, data, Web, and video). TRE is used to eliminate the transmission of redundant content and therefore to significantly reduce the network cost. In most common TRE solutions, both the sender and the receiver examine and compare signatures of data chunks. They parsed according to the data content, prior to their transmission. When redundant chunks are detected the sender replaces the transmission of each redundant chunks with its strong signature.**

***Index Terms*—Traffic Redundancy Elimination (TRE), Cloud computing, Signatures, Bandwidth.**

I. INTRODUCTION

Commercial TRE solutions are popular at enterprise network and also involve the deployment of two or more proprietary-protocol, state synchronized middle boxes at both the intranet entry points of data centers and branch offices, eliminating repetitive traffic between them (e.g., Cisco, Riverbed, Quantum, Juniper, Blue Coat, Expand Networks, and F5). While proprietary middle-boxes are popular point solutions within enterprises, they are not attractive in a cloud environment. Cloud providers can't benefit from a technology whose goal is to reduce customer bandwidth bills, and they are not likely to invest in one. The rise of "on-demand" work spaces, meeting rooms, and also work-from-home solutions detaches the workers from their offices. In such dynamic work environment, fixed-point solutions that require a client-side and a server-side middle-box pair become ineffective. On the other hand, cloud-side elasticity motivates work distribution among servers and migration among data centers. Therefore, it is commonly agreed that a universal, software-based, end-to-end TRE is crucial in today's pervasive environment. This enables the use of a standard protocol stack and makes a TRE within end-to-end secured traffic (e.g., SSL) possible. Current end-to-end TRE solutions are sender-based. In the case where the cloud server is the sender, these solutions require that the server continuously maintain clients' status. We show here that cloud elasticity calls for a new TRE solution. First, cloud load balancing and power optimizations may lead to a server-side process and data migration environment, in which TRE solutions that require full synchronization between the server and the client. They are hard to accomplish or may lose efficiency due to lost synchronization. Second, the popularity of rich media that consume high bandwidth, motivates content distribution network (CDN) solutions. In which, the service point for fixed and mobile users may change dynamically according to the relative service point locations and loads.

II. LITERATURE SURVEY

A. The power of prediction: Cloud bandwidth and cost reduction

In this paper we present PACK (Predictive ACKs), a novel end-to-end Traffic Redundancy Elimination (TRE) system, designed for cloud computing customers. Cloud-based TRE needs to apply a judicious use of cloud resources so that the bandwidth cost reduction combined with the additional cost of TRE computation and storage would be optimized. PACK's main advantage is its capability of offloading the cloud server TRE effort to end-clients, thus minimizing the processing costs induced by the TRE algorithm. Unlike previous solutions, PACK does not require the server to continuously maintain clients' status. This makes PACK very suitable for pervasive computation environment that combine client mobility and server migration to maintain cloud elasticity. PACK is based on a novel TRE technique, which allows the client to use newly received chunks to identify previously received chunk chains, which in turn can be used as reliable predictors to future transmitted chunks. We present a fully functional PACK implementation, transparent to all TCP-based applications and network devices. Finally, we analyze PACK benefits for cloud users, using traffic traces from various sources.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 7 Issue 02, February-2018

B. Content-based segmentation scheme for data compression in storage and transmission including hierarchical segment representation

In a coding system, input data within a system is encoded. The input data might include sequences of symbols that repeat in the input data or occur in other input data encoded in the system. The encoding includes determining a target segment size, determining a window size, identifying a fingerprint within a window of symbols at an offset in the input data, determining whether the offset is to be designated as a cut point and segmenting the input data as indicated by the set of cut points. For each segment so identified, the encoder determines whether the segment is to be a referenced segment with a reference label and storing a reference binding in a persistent segment store for each referenced segment, if needed. Hierarchically, the process can be repeated by grouping references into groups, replacing the grouped references with a group label. The number of levels of hierarchy can be fixed in advance or it can be determined from the content encoded.

III. EXISTING SYSTEM

Traffic redundancy stems from common end-users activities, such as repeatedly accessing, downloading, uploading (i.e., backup), distributing, and modifying the same or similar information items(documents, data, web, and video). TRE is used to eliminate the transmission of redundant content and, therefore, to significantly reduce the network cost. In most common TRE solutions, both the sender and the receiver examine and compare signatures of data chunks, parsed according to the data content, prior to their transmission. When redundant chunks are detected, the sender replaces the transmission of each redundant chunk with its strong signature. Commercial TRE solutions are popular at enterprise networks, and involve the deployment of two or more proprietary-protocol, state synchronized middle-boxes at both the intranet entry points of data centers.

Disadvantages of Existing System:

- Cloud providers cannot benefit from a technology whose goal is to reduce customerv bandwidth bills, and thus are not likely to invest in one.
- The rise of “on-demand” work spaces, meeting rooms, and work-from-home solutionsv detaches the workers from their offices. In such a dynamic work environment, fixed-point solutions that require a client-side and a server-side middle-box pair become ineffective.
- cloud load balancing and power optimizations may lead to a server-side process and datav migration environment, in which TRE solutions that require full synchronization between the server and the client are hard to accomplish or may lose efficiency due to lost synchronization.
- Current end-to-end solutions also suffer from the requirement to maintain end-to-endv synchronization that may result in degraded TRE efficiency.

IV. PROPOSED SYSTEM

In this paper, we present a novel receiver-based end-to-end TRE solution that relies on the power of predictions to eliminate redundant traffic between the cloud and its end-users. In this solution, each receiver observes the incoming stream and tries to match its chunks with a previously received chunk chain or a chunk chain of a local file. Using the long-term chunks’ metadata information kept locally, the receiver sends to the server predictions that include chunks’ signatures and easy-to-verify hints of the sender’s future data. On the receiver side, we propose a new computationally lightweight chunking (fingerprinting) scheme termed PACK chunking. PACK chunking is a new alternative for Rabin fingerprinting traditionally used by RE applications.

Advantages of Proposed System:

- Our approach can reach data processing speeds over3 Gb/s, at least 20% faster thanϖ Rabin fingerprinting.
- The receiver-based TRE solution addresses mobility problems common to quasi-mobileϖ desktop/ laptops computational environments.
- One of them is cloud elasticity due to which the servers are dynamically relocated aroundϖ the federated cloud, thus causing clients to interact with multiple changing servers.
- We implemented, tested, and performed realistic experiments with PACK within a cloudϖ environment. Our experiments demonstrate a cloud cost reduction achieved at a reasonable client effort while gaining additional bandwidth savings at the client side.
- Our implementation utilizes the TCP Options field, supporting all TCP-basedϖ applications such as Web, video streaming, P2P, e-mail, etc.
- We demonstrate that our solution achieves 30% redundancy elimination withoutϖ significantly affecting the computational effort of the sender, resulting in a 20% reduction of the overall cost to the cloud customer.

V. CONCLUSION AND FUTURE SCOPE

Cloud computing is expected to trigger high demand for TRE solutions as the amount of data exchanged between the cloud and its users is expected to dramatically increase. The cloud environment redefines the TRE system requirements, making proprietary middle-box solutions inadequate. Consequently, there is a rising need for a TRE solution that reduces the cloud’s operational cost while accounting for application latencies, user mobility, and cloud elasticity. In this paper, we have presented PACK, a receiver-based, cloud-friendly, end-to-end TRE that is based on novel speculative principles that reduce latency and cloud operational cost. PACK does not require the server to continuously maintain clients’ status, thus enabling cloud elasticity and user mobility while preserving long-term redundancy. Moreover, PACK is capable of eliminating redundancy based on content arriving to the client from multiple servers without applying a three-way handshake. Our evaluation using a wide collection of content types shows that PACK meets the expected design goals and has clear advantages over sender-based TRE, especially when

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 7 Issue 02, February-2018

the cloud computation cost and buffering requirements are important.

Moreover, PACK imposes additional effort on the sender only when redundancy is exploited, thus reducing the cloud overall cost. Two interesting future extensions can provide additional benefits to the PACK concept. First, our implementation maintains chains by keeping for any chunk only the last observed subsequent chunk in an LRU fashion. An interesting extension to this work is the statistical study of chains of chunks that would enable multiple possibilities in both the chunk order and the corresponding predictions. The system may also allow making more than one prediction at a time, and it is enough that one of them will be correct for successful traffic elimination. A second promising direction is the mode of operation optimization of the hybrid sender–receiver approach based on shared decisions derived from receiver's power or server's cost changes.

VI. REFERENCES

[1] E. Zohar, I. Cidon, and O. Mokryn, "The power of prediction: Cloud bandwidth and cost reduction," in *Proc. SIGCOMM*, 2011, pp. 86–97.
[2] 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," *Commun. ACM*, vol. 53, no. 4, pp. 50–58, 2010.
[3] U. Manber, "Finding similar files in a large file system," in *Proc. USENIX Winter Tech. Conf.*, 1994, pp. 1–10.
[4] N. T. Spring and D. Wetherall, "A protocol-independent technique for eliminating redundant network traffic," in *Proc. SIGCOMM*, 2000, vol.30, pp. 87–95.
[5] A. Muthitacharoen, B. Chen, and D. Mazières, "A low-bandwidth network file system," in *Proc. SOSP*, 2001, pp. 174–187.
[6] E. Lev-Ran, I. Cidon, and I. Z. Ben-Shaul, "Method and apparatus for reducing network traffic over low bandwidth links," US Patent 7636767, Nov. 2009.
[7] S.Mccanne andM. Demmer, "Content-based segmentation scheme for data compression in storage and transmission including hierarchical segment representation," US Patent 6828925, Dec. 2004.
[8] R. Williams, "Method for partitioning a block of data into subblocks and for storing and communicating such subblocks," US Patent 5990810, Nov. 1999.
[9] Juniper Networks, Sunnyvale, CA, USA, "Application acceleration," 1996 [Online]. Available: http://www.juniper.net/us/en/products-services/application-acceleration/
[10] Blue Coat Systems, Sunnyvale, CA, USA, "MACH5," 1996 [Online].
[11] Available: http://www.bluecoat.com/products/mach5