© 2017 IJSRST | Volume 3 | Issue 7 | Print ISSN: 2395-6011 | Online ISSN: 2395-602X
Themed Section: Science and Technology

Cost Reduction through TRE System Based on Prediction

Naushin Parveen N Sheikh[1], Sagar M. Bhakre[2]

[1]M. Tech Scholar, Department of Computer Science & Engineering, Ballarpur Institute of Technology, Ballarpur, Maharashtra, India
[2]Assistant Professor, Department of Computer Science & Engineering, Ballarpur Institute of Technology, Ballarpur, Maharashtra, India

ABSTRACT

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 environments 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.

Keywords: Caching, Cloud Computing, Network Optimization, Traffic Redundancy Elimination

I. INTRODUCTION

Cloud computing offers its customers an economical and convenient pay-as-you-go service model, known also as usage-based pricing. Cloud customers1 pay only for the actual use of 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 an important issue when trying to minimize costs. Consequently, cloud customers, applying a judicious use of the cloud's resources, are motivated to use various traffic reduction techniques, in particular traffic redundancy elimination (TRE), for reducing bandwidth costs. 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 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 as attractive in a cloud environment. Cloud providers cannot benefit from a technology whose goal is to reduce customer bandwidth bills, and thus are not likely to invest in one. The rise of "on-demand" work spaces, meeting rooms, and work-from-home solutions 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. 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 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. Moreover, if an end-to-end solution is employed, its additional computational and storage costs at the cloud side should be weighed against its bandwidth saving gains. Clearly, a TRE solution that puts most of its computational effort on the cloud side2may turn to be less cost-effective than the one that leverages the combined client-side capabilities. Given an end-to-end solution, we have found through our experiments that sender-based end-to-end TRE solutions add a considerable load to the servers, which may eradicate the cloud cost saving addressed by the TRE in the first place. Our experiments further show that current end-to-end solutions also suffer from the requirement to maintain end-to-end synchronization that may result in degraded TRE efficiency.

II. LITERATURE REVIEW

Athicha Muthitacharoen, Benjie Chen, David Mazieres "A Low-bandwidth Network File System"

Users rarely consider running network file systems over slow or wide-area networks, as the performance would be unacceptable and the bandwidth consumption too high. Nonetheless, efficient remote file access would often be desirable over such networks—particularly when high latency makes remote login sessions unresponsive. Rather than run interactive programs such as editors remotely, users could run the programs locally and manipulate remote files through the file system. To do so, however, would require a network file system that consumes less bandwidth than most current file systems. This paper presents LBFS, a network file system designed for low-bandwidth networks. LBFS exploits similarities between files or versions of the same file to save bandwidth. It avoids sending data over the network when the same data can already be found in the server's file system or the client's cache. Using this technique in conjunction with conventional compression and caching, LBFS consumes over an order of magnitude less bandwidth than traditional network file systems on common workloads.

LBFS 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 on small regions of the file to determine chunk boundaries. It indexes file chunks by their hash values, and subsequently looks up chunks to reconstruct files that contain the same data without sending that data over the network. Under common operations such as editing documents and compiling software, LBFS can consume over an order of magnitude less bandwidth than traditional file systems. Such a dramatic savings in bandwidth makes LBFS practical for situations where other file systems cannot be used. In many situations, LBFS makes transparent remote file access a vi able and less frustrating alternative to running interactive programs on remote machines.

Ankita Kotalwar , Dr. Sadhana Chidrawar "Prediction Analysis for Cloud Bandwidth and Cost Reduction" International Journal of Advance Research in Computer Science and Management Studies Volume 3, Issue 9, September 2015.

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 environments 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 net-work devices. Finally, we analyze PACK benefits for cloud users, using traffic traces from various sources

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 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.

Athicha Muthitacharoen, Benjie Chen ,David Mazier "A Low-bandwidth Network File System"

Network file system often best way to access data Copying files back and forth threatens consistency. Remote login frustrating given latency or packet loss. Most file systems too bandwidth-hungry for WAN LBFS exploits file commonality to save bandwidth. Break files into variable-size chunks based on contents. Index chunks in file system and client cache. Avoid sending chunks already present in other files. LBFS works where other file systems impractical

T. Navin Kumar, K. V. Manoj Kumar, M. R. Srinivasulu and Mrs. Girija "Cost Reduction Technique Using Cloud Computing" International Journal of Technical Research and Application Volume 3, Issue 2 (Mar-Apr 2015)

In this paper, we are using cloud computing for eliminating traffic redundancy and reducing cost for a benefit of cloud customers. Here, we introduce the technique as bandwidth prediction through synchronization over user and server. The user's bandwidth were predicted by the server and proceed with the acknowledgement process, server of cloud provide the bandwidth to the end user of which customer needs. From this, cloud providing different bandwidth for different user which automatically eliminating end-to-end traffic redundancy and cost beneficial for an every customer accessing cloud. So easily cloud customer obtains their task and pay only for the usage in the cloud.

It does not require the server to continuously to maintain client's status, thus enabling cloud elasticity and user mobility to preserving long-term eliminate traffic redundancy. Moreover, this method is capable of eliminating the redundancy based on content arriving from the cloud to the client from multiple servers without applying a three-way handshake. The Multiple predictions are used to more accurate delivery of service to consumers from provider. This system is more reliable and reducing cost for eliminating the traffic problem from existing techniques.

G. Krishnaveni, D. SriLakshmi "Advanced Prediction-Based System for Cloud Bandwidth and Cost Reduction" International Journal of Computer Science and technology IJCST Vol. 4, Issue Spl - 4, OCT - DEC 2013

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 environments 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 net-work devices. Finally, we analyse PACK benefits for cloud users, using traffic traces from various sources.

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 the cloud computation cost and buffering requirements are important. More-over, 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 sub-sequent 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 de-rived from receiver's power or server's cost changes.

III. PROPOSED WORK

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.

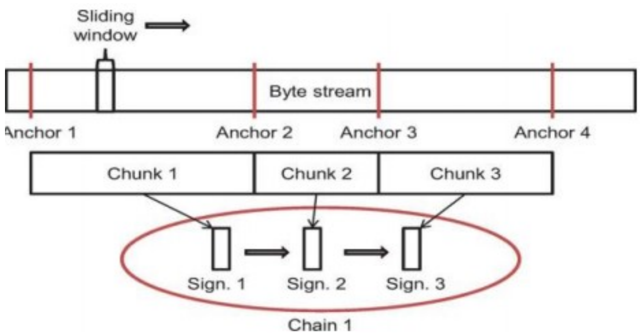

Figure 1 System Architecture

The stream of knowledge received at the PACK receiver is parsed to a sequence of variable -size, content-based signed chunks similar to . The chunks are then compared to the receiver native storage, termed chunk store. If a matching chunk is found within the native chunk store, the receiver retrieves the sequence of later chunks, referred to as a chain, by traversing the sequence of LRU chunk pointers that are enclosed within the chunks' data.

IV. CONCLUSION

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.

V. 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. Rabin, 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].