International Journal of Science Engineering and Advance Technology, IJSEAT, Vol 2, Issue 10, October - 2014
ISSN 2321-6905

International Journal of Science Engineering and Advance Technology

Forecast-Founded Haze Bandwidth and Cost Decrease Scheme

S. Srinivasulu[1], **K. Manoj Kumar**[2]
Department of CSE, Visvodaya Engineering College,Kavali.

Abstract—In this paper, we present PACK (Predictive ACKs), a novel end-to-end traffic redundancy elimination (TRE) Scheme, designed for Haze computing customers. Haze-Founded TRE needs to apply a judicious use of Haze resources so that the bandwidth cost Decrease combined with the additional cost of TRE computation and storage would be optimized. PACK's main advantage is its capability of offloading the Hazeserver TRE effort to endclients, 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 Haze elasticity. PACK is Founded 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-Founded applications and network devices. Finally, we analyze PACK benefits for Haze users, using traffic traces from various sources.

IndexTerms—Caching, Haze computing, network optimization, traffic redundancy elimination.

I. Introduction

HAZE computing offers its customers an economical and convenient pay-as-you-go service model, known also as usageFounded pricing [2]. Haze customers[1] pay only for the actual use of computing resources, storage, and bandwidth, according to their changing needs, utilizing the Haze's scalable and elastic computational capabilities. In particular, data transfer costs (i.e., bandwidth) is an important issue when trying to minimize costs [2]. Consequently, Haze customers, applying a judicious use of the Haze's resources, are motivated to use various traffic Decrease techniques, in particular traffic redundancy elimination (TRE), for reducing bandwidth costs. 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.

While proprietary middle-boxes are popular point solutions within enterprises, they are not as attractive in a Haze environment. Haze 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 [13] 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, Hazeside elasticity motivates work distribution among servers and migration among data centers. Therefore, it is commonly agreed that a universal, softwareFounded, end-toend TRE is crucial in today's pervasive environment [14], [15]. 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 senderFounded. In the case where the Haze server is the sender, these solutions require that the server continuously maintain clients' status.

II. Related Work

Several TRE techniques have been explored in recent years. A protocol-independent TRE was proposed in [4]. The paper describes a packet-level TRE, utilizing the algorithms presented in [3].

Several commercial TRE solutions described in [6] and [7] have combined the sender-Founded TRE ideas of [4] with the algorithmic and implementation approach of [5] along with protocol specific optimizations for middle-boxes solutions. In particular, [6] describes how to get away with threeway handshake between the sender and the receiver if a full state synchronization is maintained.

References [17] and [18] present redundancyaware routing algorithm. These papers assume that the routers are equipped with data caches, and that they search those routes that make a better use of the cached data.

A large-scale study of real-life traffic redundancy is presented in [19], [20], and [14]. In the latter, packetlevel TRE techniques are compared [3],[21]. Our paper builds on their finding that "an end to end redundancy elimination solution, could obtain most of the middle-box's bandwidth savings," motivating the benefit of low cost software end-toend solutions. Wanax [22] is a TRE Scheme for the developing world where storage and WAN bandwidth are scarce. It is a softwareFounded middle-box replacement for the expensive commercial hardware. In this scheme, the sender middle-box holds back the TCP stream and sends data signatures to the receiver middlebox. The receiver checks whether the data is found in its local cache. Data chunks that are not found in the cache are fetched from the sender middle-box or a nearby receiver middle-box. Naturally, such a scheme incurs a three-way-handshake latency for noncached data.

EndRE [15] is a sender-Founded end-to-end TRE for enterprise networks. It uses a new chunking scheme that is faster than the commonly used Rabin fingerprint, but is restricted to chunks as small as 32– 64 B. Unlike PACK, EndRE requires the server to maintain a fully and reliably synchronized cache for each client. To adhere with the server's memory requirements, these caches are kept small (around 10 MB per client), making the Scheme inadequate for medium-to-large content or longterm redundancy.

EndRE is server-specific, hence not suitable for a CDN or Haze environment.

end for

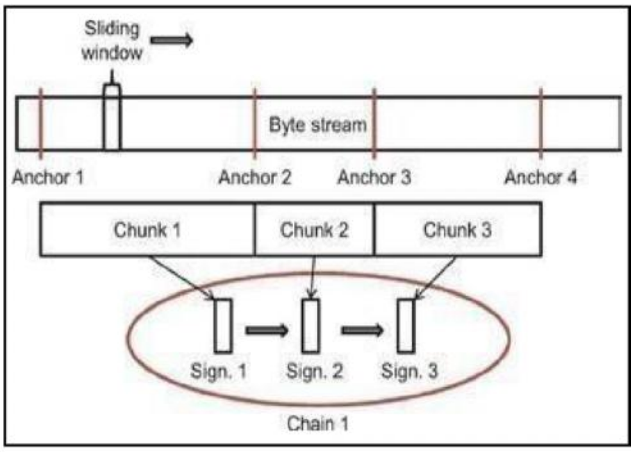

Fig. 1: From Stream to Chain

Fig. 1. From stream to chain.

To the best of our knowledge, none of the previous works have addressed the requirements for a Hazecomputingfriendly, end-to-end TRE, which forms PACK's focus.

III. Motivating A Receiver-Founded Approach

The objective of this section is twofold: evaluating the potential data redundancy for several applications that are likely to reside in a Haze, and to estimate the PACK performance and Haze costs of the redundancy elimination process.

Our evaluations are conducted using: 1) video traces captured at a major ISP; 2) traffic obtained from a popular social network service; and 3) genuine data sets of real-life workloads. In this section, we relate to an average chunk size of 8 kB, although our algorithm allows each client to use a different chunk size.

A. Traffic Redundancy 1) Traffic Traces: We obtained a 24-h recording of traffic at an ISP's 10-Gb/s PoP router, using a 2.4GHz CPU recording machine with 2 TB storage (4×500 GB 7 200 RPM disks) and 1-Gb/s NIC. We filtered YouTube traffic using deep packet inspection and mirrored traffic associated with YouTube servers IP addresses to our recording device. Our measurements show that YouTube traffic accounts for 13% of the total daily Web traffic volume of this ISP. The recording of the full YouTube stream would require 3 times our network and disk write speeds. Therefore, we isolated 1/6 of the obtained YouTube traffic, grouped by the video identifier (keeping the redundancy level intact) using a programmed load balancer that examined the upstream HTTP requests and redirected downstream sessions according to the video identifier that was found in the YouTube's URLs, to a total of 1.55 TB. For accurate reading of the true redundancy, we filtered out the client IP addresses that were used too intensively to represent a single user and were assumed to represent a NAT address.

Conclusion

Haze computing is expected to trigger high demand for TRE solutions as the amount of data exchanged between the Haze Our evaluation using a wide collection of content types shows that PACK meets the expected design goals and has clear advantages over sender-Founded TRE, especially when the Haze computation cost and buffering requirements are important. Moreover, PACK imposes additional effort on the sender only when redundancy is exploited, thus reducing the Haze 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 Forecasts. The Scheme may also allow making more than one

Forecast 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 Founded on shared decisions derived from receiver's power or server's cost changes.

References

[1] E. Zohar, I. Cidon, and O. Mokryn, “The power of Forecast: Haze bandwidth and cost Decrease,” 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 Haze computing,” Commun. ACM, vol. 53, no. 4, pp. 50–58, 2010.

[3] U. Manber, “Finding similar files in a large file Scheme,” 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 Scheme,” in Proc. SOSP, 2001, pp. 174–187.

[6] E. LevRan, 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 and M. Demmer, “ContentFounded 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/productsservices/application-acceleration/

[10] Blue Coat Schemes, Sunnyvale, CA, USA, “MACH5,” 1996 [Online]. Available: http://www.bluecoat.com/products/mach5

[11] Expand Networks, Riverbed Technology, San Francisco, CA, USA, “Application acceleration and WAN optimization,” 1998 [Online]. Available: http://www.expand.com/technology/ application-acceleration.aspx

[12] F5, Seattle, WA, USA, “WAN optimization,” 1996 [Online]. Available:http://www.f5.com/solutions/acceleration/ wan-optimization/

[13] A. Flint, “The next workplace revolution,” Nov. 2012 [Online]. Available: http://m.theatlanticcities.com/jobsandeconomy/2012/ 11/ nextworkplacerevolution/3904/

[14] A. Anand, C. Muthukrishnan, A. Akella, and R. Ramjee, “Redundancy in network traffic: Findings and implications,” in Proc. SIGMETRICS, 2009, pp. 37–48.

[15] B. Aggarwal, A. Akella, A. Anand, A. Balachandran, P. Chitnis, C. Muthukrishnan, R. Ramjee, and G. Varghese, “EndRE: An endScheme redundancy elimination service for enterprises,” in Proc. NSDI, 2010, pp. 28–28.

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

[17] 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.

[18] A. Anand, V. Sekar, and A. Akella, “SmartRE: An architecture for coordinated network-wide redundancy elimination,” in Proc. SIGCOMM, 2009, vol. 39, pp. 87–98.

[19] A. Gupta, A. Akella, S. Seshan, S. Shenker, and J. Wang, “Understanding and exploiting network traffic redundancy,” UWMadison, Madison, WI, USA, Tech. Rep. 1592, Apr.2007.

[20] M. Zink, K. Suh, Y. Gu, and J. Kurose, “Watch global, cache local: YouTube network traffic at a campus network—Measurements and implications,” in Proc. MMCN, 2008, pp. 1–13.

[21] S. Schleimer, D. S. Wilkerson, and A. Aiken,“Winnowing: Local algorithms for document fingerprinting,” in Proc. SIGMOD, 2003, pp. 76–85.

[22] S. Ihm, K. Park, and V. Pai, “Wide-area network acceleration for the developing world,” in Proc. USENIX ATC, 2010, pp. 18–18.

[23] H. Stevens and C. Pettey, “Gartner says Haze computing will be as influential as e-business,” Gartner Newsroom, Jun. 26, 2008.

Authors:

S.Srinivasulu , Asst. Professor, Department Of Cse, Visvodaya Engineering College, Kavali.

K. Manoj Kumar, M.Tech(Cse) Student,Visvodaya Engineering College,Kavali