ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

Literature Survey on Hybrid Approach for Cloud Bandwidth and Cost Reduction System Using Predictive Acknowledgements

Syed Sajid Hussain[1], A R Arunachalam[2]

M.Tech, Department of CSE, Bharath University, Chennai, India[1]

Associate Professor, Department of CSE Bharath University Chennai, India[2]

ABSTRACT:- During the last few years there has been a tremendous and efficient improvement in the computational world and one of the area that is leaving its footprints is cloud computing. One of the main trending issue in this computing is traffic redundancy elimination (TRE).Cloud-based TRE needs to apply a judicious use of cloud resources so that the bandwidth cost reduction is optimized together with the additional cost of TRE computation . Here, I present an approach that is suitable for cloud computing customers called Hybrid PACK(predictive acknowledgements) regarding TRE systems. Unlike other techniques PACK's main advantage is its capability of offloading the cloud-server TRE effort to the end clients ,and hence minimizing the overhead and processing costs induced by the TRE algorithm. In the present scenario the server keeps track of all the end clients. But this is not the case with the PACK as in this mechanism a client can manage his own status and hence the server is offloaded. This makes the PACK suitable for pervasive computing .PACK is based on a novel TRE technique ,which allows the client as well as server to use newly received chunks to identify previously received chunk chains, which in turn can be used as reliable predictors to future transmitted chunks. I here present a hybrid implementation of the PACK in which there is an active involvement of client and server and moreover the chunking signatures are done using SHA 2 (secure hash algorithm) in order to process the any size of chunks.

KEYWORDS: Predictive Acknowledgement, Traffic Redundancy Elimination, Caching , Cloud Computing , Network Optimization.

I. INTRODUCTION

Cloud computing is emerging style of delivery in which applications, data and resources are rapidly provisioned as standardized offerings to users with a flexible price. The cloud computing paradigm has achieved widespread adoption in recent years. Its success is due largely to customer's ability to use services on demand with a pay-as-you go pricing [5] model, which has proved convenient in many respects. Low costs and high flexibility make migrating to the cloud compelling. Cloud computing is the long dreamed vision of computing as a utility, where users can remotely store their data into the cloud so as to enjoy the on-demand high quality applications and services from a shared pool of configurable computing resources. By data outsourcing, users can be relieved from the burden of local data storage and maintenance. Traffic redundancy and elimination approach is used for minimizing [5] the cost. Traffic redundancy stems from common end-users' activities, such as repeatedly accessing, downloading, 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 ex- amine 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 [2] . 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 centres and branch offices, eliminating repetitive traffic between them [4] [9] (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. First, cloud providers cannot benefit from a technology whose goal is to reduce customer bandwidth bills, and thus are not likely to invest in one. Moreover, a fixed client-side

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

and server-side middle-box pair solution is inefficient for a combination of a mobile environment, which detaches the client from a fixed location, and cloud-side elasticity which motivates work distribution and migration among data centres. 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.

In this paper we proposed a new method called hybrid approach for cloud bandwidth and cost reduction system using predictive acknowledgements.

The layout of the paper is as follows. In section 2, we address the above mentioned techniques and also give a brief on the literature being reviewed for the same. Section 3, presents a comparative study of the various research works explored in the previous section. Lastly, we concluded in section 4 and section 5 is provided references.

II. RELATED WORK

There has been an enormous and steady usage of several TRE techniques. A protocol independent technique for eliminating redundant network traffic was proposed in [1]. In this paper a technique for identifying repetitive information transfers and use it to analyze the redundancy of network traffic has been suggested. Here the insight is that dynamic content , streaming ,media and other traffic that is not caught by today's web caches is nonetheless likely to derive from similar information. Hence a similarity detection techniques have been adopted for designing a system to eliminate redundant transfers. And to identify repeated byte ranges between packets to avoid retransmitting the redundant data.

The concept of a low-bandwidth network file system [2] was implemented. Users rarely consider running network file systems over slow or wide-area networks, as the performance would be unacceptable and the bandwidth consumption is too high. This paper presents LBFS (low- bandwidth network file system), 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.

In this paper [3] The key idea in the prediction algorithm is to treat a set of previously observed traffic matrices as "experts" and learn online the best weighted linear combination of these experts to make its prediction. With tenant VM placement using these predictive guarantees, we find that the inter-rack network utilization in certain datacenter topologies can be more than doubled.

In this paper [4] based on apparatus for reducing network traffic. At sender a data chunk is identified for transmission to a receiver, which is connected to the sender over a communication link. The sender computes a signature of the data chunk and determines whether the data chunk has been previously transmitted by looking up the signature in a sender index table. The sender index table associates the signatures of previously transmitted data chunks with unique index values. A message is transmitted to the receiver, where if the data chunk has previously been transmitted then the message includes an index value from the sender index table that is associated with the signature of the data chunk. At the receiver , the data chunk is located in a receiver cache that stores the previously transmitted data chunks by looking up the index value included in the message in a receiver index table. The receiver index table associates the unique index values with the locations in the receiver cache of the previously transmitted data chunks.

With the advent of the cloud computing [5], Developers with innovative ideas for new internet services no longer require the large capital outlays in hardware to deploy their service or the human expense to operate it. They need not be concerned about over provisioning for a service whose popularity does not meet their predictions, thus wasting costly resources, or under provisioning for one that becomes wildly popular, thus missing potential customers and revenue.

In last couple of years there huge increase in the usage cloud computing because cloud computing is emerging style of IT-delivery in which applications, data and resources are rapidly provisioned provided as Standardized offerings to users with a flexible price. But it is important to provide the convenient pricing model for the users of cloud. Hence, here a new traffic redundancy and elimination scheme [6] has been designed for reducing the cloud bandwidth and costs. This new traffic redundancy elimination approach also called as novel-TRE or receiver based TRE, which detects redundancy at the client side. However for server specific TRE approach it is difficult to handle the traffic efficiently and it doesn't suites for the cloud environment because of high processing costs. Novel-TRE matches

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

incoming chunks with a previously received chunk chain or local file and sending to the server for predicting the future data.

Cloud computing provides users to store their data remotely and enjoy the on demand high quality services and offers it as usage based pricing. Here [7] , bandwidth reduction is an important issue in this paper. Cloud service providers (CSP) are trying to reduce this bandwidth and also applying the judicious use of cloud resources by deploying a tre system to end-to-end clients. Generally cloud server maintains the client mobility and server migration to cloud elasticity along with this it also maintaining the continuous client's status of data transferring, traffic Reduction, it is again a burden to server, so in order to reduce the overhead to server in this paper TRE system is receiver based, by maintaining a chunk store and allowing the client to use newly received chunks to identify previously received chunk chains which in turn sends a predictor message for the subsequent chunks. When redundancy is detected the sender then sends only acknowledgement to the prediction instead of sending data.

Despite the exceptional prominence [8] of the cloud computing , the customers are lack of direct sense to select the cloud that delivers the best performance, due to the Performance heterogeneity of each cloud provider. Existing solutions either migrate the application to each cloud and evaluate the performance individually, or benchmark each cloud along various dimensions and predict the overall performance of the application. However, the former incurs significant migration and configuration overhead, while the latter may suffer from coarse prediction accuracy. This thesis introduces two systems to address this issue. Cloud prophet predicts the web application performance by tracing and replaying the on-premise resource demand on the cloud machines. DTRCP further predicts the performance for general applications. In particular, it addresses the execution path divergence manifested during replaying the on-premise resource demand. The results show that both systems can accurately predict the application performance.

The research [9] focuses on application mapping strategy required and financial cost benefits gained for MVNO's (Mobile Virtual Network Operator) if their systems are implemented on cloud. The thesis research involves delphi method based study in which expert interviews are conducted in two rounds to get expert views on the subject. The responses and related analyses are drawn while introducing a proof-of-concept hybrid cloud architecture to implement MVNO systems. An innovative solution of billing-asa-service on cloud is discussed in the thesis as well. The implications and future prospects are presented evaluating the thesis results. In the end, conclusion is derived to summarize the thesis research.

This paper [10] is summarized as follows. 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 or an unreferenced segment, replacing the segment data of each referenced segment with a reference label and storing a reference binding in a persistent segment store for each referenced segment, if needed.

A new and innovative method was proposed in [11]. In this paper, apart from other TRE techniques a new algorithm has been suggested which is entirely based on predictions. This algorithm is PACK (Predictive Acknowledgement). PACK is based on 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 chunk chains.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

III. COMPARTIVE STUDY

We have analyzed the various research works on several parameters and presented their comparison in the table below.

Table1. COMPARISON OF VARIOUS RESEARCH WORKS

S.No.	TITLE	AUTHOR	ISSUES	METHOD USED	TOOLS	ADVANTAGE/ DISADVANTAGE
1	A protocol independent technique for eliminating redundant network traffic	N.T. Spring and D. Wetherall	SIGCOMM, 2000,vol.30, PP.87-95	Identification of similar files in a file system,Duplication,Suppression	Rabin Fingerprinting	1.Protocol independent 2.Able to identify fine-grianed sharing 3.Doesn't need to be updated **Disadvantages** 1.Tools used are outdated.
2	A low bandwidth network file system	A.Muthitacharoen, B. Chen and D. Mazieres	SOSP,2001, pp.174-187	Usage of large,persistent file cache at client	SHA-1(hash function)	1.Indexing file chunks by their hash values 2. secure **Disadvantages** 1. requires custom protocol semantics for accessing ?les managed through an LBFS server.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

3	Cicada: Predictive Gaurantees for cloud network bandwidth	Katrina LaCurts, Jeffrey C. Mogul, Hari Balakrishnan, and Yoshio Turner	MIT-CSAIL-TR-2014-004,March 24,2014	Prediction algorithm usable by a cloud provider to suggest an appropriate bandwidth guarantee to a tenant	Cicada, a system that implements predictive guarantees	1. With tenant VM placement using these predictive guarantees, we find that the inter-rack network utilization in certain datacenter topologies can be more than doubled. 2.the predictive guarantee abstraction supports timevarying and space-varying demands.
4	Method and apparatus for reducing network traffic over low bandwidth links	E.Lev-Ran, I.Cidon and I.Z.Ben-Shaul	US Patent 7636767,Nov.2009	Syncronization of data between sender and receiver for reducing network traffic.	Rabin fingerprinting by random polynomials	1.This network infrastructure allows for greater sharing of information for users throughout the enterprise.
5	A view of cloud computing	M.Armbrust, A. Fox,R. Griffith,A.D. Joseph,R.Katz,A.Konwinski	Commun. ACM, vol. 53,no. 4, pp.50-58,2010	Clearing the clouds away from the true potential and obstacles posed by this computing capability.	Online provisioning of resources.	1.Elimination of an up-front commitment by Cloud users 2. Simplify operation and increase utilization via resource virtualization
6	Design of traffic redundancy and elimination approach for reducing cloud bandwidth and cost	Suresh Chougala and Sharavana K	IJIRCCE, Vol. 2,issue 2, February 2014	Receiver based TRE system for networks	Novel TRE	1.efficiently prediction of redundant data 2.Network cost is reduced

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

7	Periodically predicting client's bandwidth and cost acknowledgements sends to server to optimize resource usage	Prakash E J, A Nageshwara Rao	IJIRCCE, Vol.2, special issue 4,September 2014	Deployment of TRE system to END-to-END clients.	Receiver based TRE	1.Judicious use of resources 2.Network traffic gets reduced **Disadvantages** 1.No continuous synchronization of the system.
8	Predicting application performance in the cloud	Xaunran Zong	Duke university,2011	It addresses the execution path divergence manifested during replaying the on-premise resource demand	Cloud Prophet,DTRCP	1. promising prediction results for web applications. **Disadvantages** 1.DTRCP can handle only applications with a few threads.
9	Mobile virtual network operator systems on cloud: An architectural and cost-benefit study	Rushil Dave	Master's thesis,august 8,2011	Delphi method based study	Techno-Economic analysis	1.Reduction in ARPU(average revenue per user) 2.Gain in the economy
10	Content based segmentation scheme for data compression in storage and transmission including hierarchical segment representation	S. Maccane and M. Demmer	US Patent 6828925,Dec. 2004	Encoding the input data and segmenting the input data as indicated by the set of cut points.	Data compression and segmentation	1.Useful for low bandwidth networks 2.secure **Disadvantages** 1.needs complex system for its implementation
11	The power of Prediction,cloud bandwidth and cost reduction system	Eyal Zohar, Israel Cidon,and Osnat Mokryn	IEEE/ACM transactions on networking,vol.22 issue.1,feb. 2014	Receiver based traffic redundancy elimination.	PACK: predictive acknowledgements with SHA-1 fingerprinting	1.Receiver enjoys the deemed status 2.Mobility problems of the receiver are eliminated **Disadvantages** 1.server remains idle

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

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 middlebox 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.

Here, PACK a hybrid-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.

The evaluation using a wide collection of content types shows that PACK meets the expected design goals and has clear advantages over sender-based TRE, as well as receiver based 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, the 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.

REFERENCES

1. N. T. Spring and D. Wetherall, "A protocol-independent technique for eliminating redundant network traffic," in Proc. SIGCOMM, 2000, vol.30, pp. 87–95.
2. A.Muthitacharoen, B. Chen, and D. Mazières, "A low-bandwidth network file system," inProc. SOSP, 2001, pp. 174–187.
3. Katrina LaCurts, Jeffrey C. Mogul, Hari Balakrishnan, and Yoshio Turner, "Cicada: Predictive Guarantees for cloud network bandwidth," MIT-CSAIL-TR-2014-004,March 24,2014
4. 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.
5. 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.
6. Suresh Chougala and Sharavana K," Design of traffic redundancy and elimination approach for reducing cloud bandwidth and cost," IJIRCCE, Vol. 2,issue 2, February 2014.
7. Prakash E J, A Nageshwara Rao," Periodically predicting client's bandwidth and cost acknowledgements sends to server to optimize resource usage," IJIRCCE, Vol.2, special issue 4,September 2014.
8. Xaunran Zong, "Predicting application performance in the cloud," Master's thesis , Duke university,2011.
9. Rushil Dave, "Mobile virtual network operator systems on cloud: An architectural and cost-benefit study," Master's thesis, august 8,2011.
10. S.Mccanne and M.Demmer, "Content-based segmentation scheme for data compression in storage and transmission including hierarchical segment representation," US Patent 6828925, Dec. 2004.
11. E. Zohar, I. Cidon, and O. Mokryn, "The power of prediction: Cloud bandwidth and cost reduction," in *Proc. SIGCOMM*, 2011, pp. 86–97.

BIOGRAPHY

Syed Sajid Hussain is currently pursuing his master's degree (M.Tech) in computer science and engineering at Bharath university, Chennai, India. He completed his B.Tech degree (2009-2013) in the same subject from Islamic University of science and technology Awantipora, Jammu and Kashmir, India. His areas of interest are Neural Networks, Mobile computing, Database Management and Cloud Computing.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

AR Arunachalam received his B.E degree in Computer science and engineering from Madras University in 2002 and received his M.Tech degree in Computer science and engineering from Bharath university in 2007.He is currently pursuing his Ph.D in Computer science and engineering at Bharath university, Chennai.He has 10 years of teaching experience and has guided many B.Tech and M.tech projects.