

Packet Loss Control Using Tokens at the Network Edge

Z. Shi, D. Ionescu and D. Zhang

Abstract— Presently the Internet accommodates simultaneous audio, video, and data traffic. This requires the Internet to guarantee the packet loss which at its turn depends very much on congestion control. A series of protocols have been introduced to supplement the insufficient TCP mechanism controlling the network congestion. CSFQ was designed as an open-loop controller to provide the fair best effort service for supervising the per-flow bandwidth consumption and has become helpless when the P2P flows started to dominate the traffic of the Internet. Token-Based Congestion Control (TBCC) is based on a closed-loop congestion control principle, which restricts token resources consumed by an end-user and provides the fair best effort service with $O(1)$ complexity. As Self-Verifying CSFQ and Re-feedback, it experiences a heavy load by policing inter-domain traffic for lack of trust. In this paper, Stable Token-Limited Congestion Control (STLCC) is introduced as new protocols which appends inter-domain congestion control to TBCC and make the congestion control system to be stable. STLCC is able to shape output and input traffic at the inter-domain link with $O(1)$ complexity. STLCC produces a congestion index, pushes the packet loss to the network edge and improves the network performance. Finally, the simple version of STLCC is introduced. This version is deployable in the Internet without any IP protocols modifications and preserves also the packet datagram.

Modern IP network services provide for the simultaneous digital transmission of voice, video, and data. These services require congestion control protocols and algorithms which can solve the packet loss parameter can be kept under control.

Congestion control is therefore, the cornerstone of packet switching networks [28]. It should prevent congestion collapse, provide fairness to competing flows and optimize transport performance indexes such as throughput, delay and loss. The literature abounds in papers on this subject; there are papers on high-level models [17][18][19] of the flow of packets through the network, and on specific network architecture [20][21][22][23][24][25][26] [27].

Despite this vast literature, congestion control in telecommunication networks struggles with two major problems that are not completely solved. The first one is the time-varying delay between the control point and the traffic sources. The second one is related to the possibility that the traffic sources do not follow the feedback signal. This latter may happen because some sources are silent as they have nothing to transmit.

Congestion control of the best-effort service in the Internet was originally designed for a cooperative environment. It is still mainly dependent on the TCP congestion control algorithm at

terminals, supplemented with load shedding [1] at congestion links. This model is called the Terminal Dependent Congestion Control case.

Although routers equipped with Active Queue Management such as RED [2] can improve transport performance, they are neither able to prevent congestion collapse nor provide fairness to competing flows.

In order to enhance fairness in high speed networks, Core-Stateless Fair Queuing (CSFQ) [3] set up an open-loop control system at the network layer, which inserts the label of the flow arrival rate onto the packet header at edge routers and drops the packet at core routers based on the rate label if congestion happens. CSFQ is the first to achieve approximate fair bandwidth allocation among flows with $O(1)$ complexity at core routers.

According to CacheLogic report, P2P traffic was 60% of all the Internet traffic in 2004, of which Bit-Torrent [4] was responsible for about 30% of the above, although the report generated quite a lot of discussions around the real numbers. In networks with P2P traffic, CSFQ can provide fairness to competing flows, but unfortunately it is not what end-users and operators really want. Token-Based Congestion Control (TBCC) [5] restricts the total token resource consumed by an end-user. So, no matter how many connections the end-user has set up, it cannot obtain extra bandwidth resources when TBCC is used.

The Self-Verifying CSFQ [6] tries to expand CSFQ across the domain border. It randomly selects a flow, re-estimates the flow's rate, and checks whether the re-estimated rate is consistent with the label on the flow's packet. Consequently Self-Verifying CSFQ will put a heavy load on the border router and makes the weighted CSFQ null and void.

In [7][8], the authors present a congestion control architecture Re-feedback, which aims to provide the fixed cost to end-users and bulk inter-domain congestion charging to network operators. Re-feedback not only demands very high level complexity to identify the malignant end-user, but also is difficult to provide the fixed congestion charging to the inter-domain interconnection with low complexity.

There are three types of inter-domain interconnection polices, the Internet Exchange Points, the private peering and the transit. In the private peering polices, the Sender Keep All (SKA) peering arrangements are those in which traffic is exchanged between two domains without mutual charge. As Re-feedback is based on congestion charges to the peer domain, it is difficult for Re-feedback to support the requirements of SKA.

In this paper a new and better mechanism for congestion control with application to Packet Loss in networks with P2P traffic is proposed. In this new method the edge and the core routers will write a measure of the quality of service guaranteed by the router by writing a digital number in the Option Field of the datagram of the packet. This is called a token. The token is read by the path routers and interpreted as its value will give a measure of the congestion especially at the edge routers. Based on the token number the edge router at the

Z. Shi, Software of Institute, Chinese Academy of Science, China, szq@iscas.ac.cn

D. Ionescu, University of Ottawa, Canada, ionescu@site.uottawa.ca

D. Zhang, University of Ottawa, Canada, dzhang@site.uottawa.ca

source's edge point will shape the traffic generated by the source, thus reducing the congestion on the path. In Token-Limited Congestion Control (TLCC) [9], the inter-domain router restricts the total output token rate to peer domains. When the output token rate exceeds the threshold, TLCC will decrease the Token-Level of output packets, and then the output token rate will decrease.

Similarly to CSFQ and TBCC, TLCC uses also the iterative algorithm to estimate the congestion level of its output link, and requires a long period of time to reach a stable state. With bad parameter configuration, TLCC may cause the traffic to fall into an oscillation process. The window size of TCP flows will always increase when acknowledge packets are received, and the congestion level will increase at the congested link. At congestion times many flows will lose their packets. Then, the link will be idle and the congestion level will decrease. The two steps may be repeated alternately, and then the congestion control system will never reach stability.

To solve the oscillation problem, the Stable Token-Limited Congestion Control (STLCC) is introduced. It integrates the algorithms of TLCC and XCP [10] altogether. In STLCC, the output rate of the sender is controlled according to the algorithm of XCP, so there is almost no packet lost at the congested link. At the same time, the edge router allocates all the access token resource to the incoming flows equally. When congestion happens, the incoming token rate increases at the core router, and then the congestion level of the congested link will also increase. Thus STLCC can measure the congestion level analytically, allocate network resources according to the access link, and further keep the congestion control system stable.

This paper is organized as follows. In section II, the architecture of Token-Based Congestion Control (TBCC), which provides fair bandwidth allocation to end-users in the same domain will be introduced. Section III evaluates two congestion control algorithms CSFQ and TBCC. In section IV, STLCC is presented and the simulation is designed to demonstrate its validity. Section V presents the Unified Congestion Control Model which is the abstract model of CSFQ, Re-feedback and STLCC. In section VI, the simple version of STLCC is proposed, which can be deployed on the current Internet. Finally, conclusions will be given.

To inter-connect two TBCC domains, the inter-domain router is added to the TBCC system as in Figure 8. To support the SKA arrangement, the inter-domain router should limit its output token rate to the rate of the other domains and police the incoming token rate from peer domains.

To limit the output token rate, three elements $tkprev$, $tkdown$ and $tkbackdown$ are inserted into the extended header $tkhead$. At the source edge router, the $tkprev$ is set to the same value as the $tklevel$ and cannot be modified by routers. The sum of $tkdown$ represents the decrements of Token-Level at all the inter-domain routers in the transmission path. When the packet arrives at the destination, the sum of $tkpath$ and $tkdown$ is the Congestion-Index of the transmission path. In the reverse packet, the $tkbackdown$ is used to return the elements of $tkdown$ in the forwarding packet header to the source edge router.

Keywords— P2P, Congestion Control, Congestion-Index, CSFQ, TBCC, Re-feedback, Inter-Domain, TLCC.

ACKNOWLEDGMENT

This work is supported by National Natural Science Foundation of China under Grants No. 60873258, and by Chinese National 973 Fundamental Research Program under Grants No.2007CB307106.

REFERENCES

- [1] Andrew S. Tanenbaum, Computer Networks, Prentice-Hall International, Inc.
- [2] S. Floyd and V. Jacobson. Random Early Detection Gateways for Congestion Avoidance, ACM/IEEE Transactions on Networking, August 1993.
- [3] Ion Stoica, Scott Shenker, Hui Zhang, "Core-Stateless Fair Queueing: A Scalable Architecture to Approximate Fair Bandwidth Allocations in High Speed Networks", In Proc. of SIGCOMM, 1998.
- [4] D. Qiu and R. Srikant. Modeling and performance analysis of BitTorrent-like peer-to-peer networks. In Proc. of SIGCOMM, 2004.
- [5] Zhiqiang Shi, Token-based congestion control: Achieving fair resource allocations in P2P networks, Innovations in NGN: Future Network and Services, 2008. K-INGN 2008. First ITU-T Kaleidoscope Academic Conference.
- [6] I. Stoica, H. Zhang, S. Shenker, Self-Verifying CSFQ, in Proceedings of INFOCOM, 2002.
- [7] Bob Briscoe, Policing Congestion Response in an Internetwork using Refeedback, In Proc. ACM SIGCOMM05, 2005.
- [8] Bob Briscoe, Re-feedback: Freedom with Accountability for Causing Congestion in a Connectionless Internetwork, http://www.cs.ucl.ac.uk/staff/B.Briscoe/projects/e2ephd/e2ephd_y9_cutdown_appxs.pdf
- [9] Zhiqiang Shi, Yuansong Qiao, Zhimei Wu, Congestion Control with the Fixed Cost at the Domain Border, Future Computer and Communication (ICFCC), 2010.
- [10] Dina Katabi, Mark Handley, and Charles Rohrs, "Internet Congestion Control for Future High Bandwidth-Delay Product Environments." ACM Sigcomm 2002, August 2002.
- [11] Abhay K. Patekh, "A Generalized Processor Sharing Approach Flow Control in Integrated Services Networks: The Single-Node Case", IEEE/ACM Trans. on Network, Vol. 1, No.3, June 1993.
- [12] Sally Floyd, Van Jacobson, Link-sharing and Resource Management Models for Packet Networks, IEEE/ACM Transactions on Networking, Vol.3, No.4, 1995.
- [13] John Nagle, RFC896 congestion collapse, January 1984.
- [14] Sally Floyd and Kevin Fall, Promoting the Use of End-to-End Congestion Control in the Internet, IEEE/ACM Transactions on Networking, August 1999.
- [15] V. Jacobson. "Congestion Avoidance and Control". SIGCOMM Symposium on Communications Architectures and Protocols, pages 314–329, 1988.
- [16] <http://www.isi.edu/nsnam/ns/>
- [17] L. Benmohamed and S. M. Meerkov, "Feedback control of congestion in packet switching networks: The case of a single congested node," IEEE/ACM Trans. Networking, vol. 1, pp. 693–708, Dec. 1993.
- [18] E. Altman, T. Basar, and R. Srikant, "Control methods for communication networks," in Proc. 36th Conf. Decision and Control, San Diego, CA, 1997, pp. TA 31774–1809, TM3 2368–2404, TP3 2903–2945.
- [19] S. H. Low and D. E. Lapsley, "Optimization flow control—I: Basic algorithms and convergence," IEEE/ACM Trans. Networking, vol. 7, pp. 861–874, Dec. 1999.
- [20] S. Floyd and K. Fall, "Promoting the use of end-to-end congestion control in the Internet," IEEE/ACM Trans. Networking, vol. 7, pp. 458–472, Aug. 1999.
- [21] S. J. Golestani and S. Bhattacharyya, "A class of end-to-end congestion-control algorithms for the Internet," in Proc. Int. Conf. Network Protocols (ICNP), Austin, TX, Oct. 1998, pp. 137–150.
- [22] A. Kolarov and G. Ramamurthy, "A control theoretic Approach to the design of explicit rate controller for ABR service," IEEE/ACM Trans. Networking, vol. 7, pp. 781–753, Oct. 1999.

- [23] N. Ghani and J. W. Mark, "Enhanced distributed explicit rate allocation for ABR services in ATM networks," *IEEE/ACM Trans. Networking*, vol. 7, pp. 710–723, Oct. 1999.
- [24] S. Kalyanaraman, R. Jain, R. Goyal, S. Fahmy, and B. Vandalore, "The ERICA switch algorithm for ABR traffic management in ATM networks," *IEEE/ACM Trans. Networking*, vol. 8, pp. 87–98, Feb. 2000. [12] T. V. Lakshman, P. P. Mishra, and K. K. Ramakrishnan, "Transporting compressed video over ATM networks with explicit-rate feedback control," *IEEE/ACM Trans. Networking*, vol. 8, pp. 71–86, Feb. 2000.
- [25] R. Muthukrishnan, S. Dasgupta, A. Varma, L. Kalampoukas, and K. K. Ramakrishnan, "Design, implementation and evaluation of an explicit rate allocation algorithm in an ATM switch," in *Proc. IEEE INFOCOM*, Tel Aviv, Israel, Apr. 2000, pp. 1313–1322.
- [26] Y. Zhao, S. Q. Li, and S. Sigarto, "A linear dynamic model for design of stable explicit-rate ABR control schemes," in *Proc. IEEE INFOCOM*, Kobe, Japan, Apr. 1997, pp. 283–292.
- [27] S. Mascolo, D. Cavendish, and M. Gerla, "ATM rate-based congestion control using a Smith predictor: An EPRCA implementation," in *Proc. IEEE INFOCOM'96*, San Francisco, CA, 1996, pp. 569–576.



Zhiqiang Shi is an associate professor of Institute of Software Chinese Academy of Sciences (ISCAS). He received his Ph.D. in Computer Engineering at ISCAS in 2001. He has done extensive research in areas of congestion control, mobile computing, Quality of Service and network management in the Internet. His contributions to the area include Token-based Congestion Control, which provides closed loop congestion control at the Network Layer and protects traditional TCP applications from the abuse of P2P streams.



Dr. Dan Ionescu is a full Professor with the University of Ottawa. He is the Director of the Network Computing and Control Technologies research laboratory since 1999. Former director of Computer Engineering at the above university from 1996 to 2000, Dr. Dan Ionescu is a senior member of various IEEE, IFIP, and IFAC groups. His research at the University of Ottawa spans over a few domains such as Artificial Intelligence, Machine Vision, Distributed Computing, Network Computing and Control, Internetworking Technologies, Web Collaboration platforms and others.



Dongli Zhang received the B.Sc. and M.A.Sc. degrees in Computer Science and Engineering, NanKai University, TianJin, China, in 1996 and 1999, respectively. He got his Ph.D. degree in Electrical and Computer Engineering at the University of Ottawa, ON, Canada. He is currently working with Cisco Canada. His research interests include distributed computing and web service, VPN and security, data communication and network, network management, QoS and control technology.