ISSN:0975-9646

Sudhakar Murugesan et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 3 (5) , 2012,5243-5246

Avoiding Congestion Control Using Network Border Patrol

Sudhakar Murugesan, Ganesh Gangadharan, Dr. Rajan John

Computer Science & Engineering Department,
All Nations University College
P.O Box KF 1908, Ghana, West Africa

Abstract-Scalability and robustness is the important factor of the end-to-end nature of Internet congestion control. However end-to-end congestion control algorithm alone is incapable of preventing the congestion collapse and unfair bandwidth allocations created by applications which are unresponsive to network congestion. This paper proposes and investigates a new congestion avoidance mechanism called Network Border Patrol (NBP). NBP relies on the exchange of feedback between router at the border of a network in order to detect and restrict unresponsive traffic flow before they enter the network. Simulation results show that NBP effectively eliminates congestion collapse and that, when combined with fair queuing, NBP achieves approximately max-min fair bandwidth allocations for competing network flow. Network Border Patrol is a core-stateless congestion avoidance mechanism.

1. INTRODUCTION

The essential philosophy behind the internet is expressed by the scalability argument: no protocol algorithm (or) service should be introduced into the internet if does not scale well. TCP have been a critical factor in the robustness of the internet. The current internet suffers from two maladies,

1. Congestion Collapse from undelivered packets
2. Unfair bandwidth allocation

The first malady-Congestion collapse from undelivered packets-arises when bandwidth is continuously consumed by packets that are dropped before reaching their destinations. The second malady-unfair bandwidth allocation-arises in the internet for a variety of reasons, one of which is the presence of unresponsive flows. To avoid these maladies, a novel Internet traffic control protocol called network border patrol (NBP). Although NBP is capable of preventing congestion collapse but fails to provide fairness of bandwidth allocations. To avoid these type of maladies improved packet scheduling or queue management mechanism is used in network routers. There are several rate control algorithm able to prevent the congestion collapse. This algorithm designed for the ATM Available Bit Rate (ABR). But this algorithm is not suitable to the current internet, because they violate the internet design philosophy of keeping router implementation is simple and pushing complexity to the edges of the network.

2. THE PROBLEM OF UNRESPONSIVE FLOWS

Two responsive flows compete for bandwidth in a network containing two links arbitrated by a fair queuing mechanism. At the first link (R1-R2), fair queuing ensures that each flow receives half of the link's available bandwidth (750 kbps). The second link (R2-S4), much of the traffic from flow B is discarded due to the link's limited capacity (x kbps). Hence flow A achieves a throughput of 750 kbps and flow B achieves throughput of x kbps. Clearly, congestion collapse occurred, because flow B packets which are intimately discarded on the second link, unnecessarily limit the throughput of flow A across the first link.

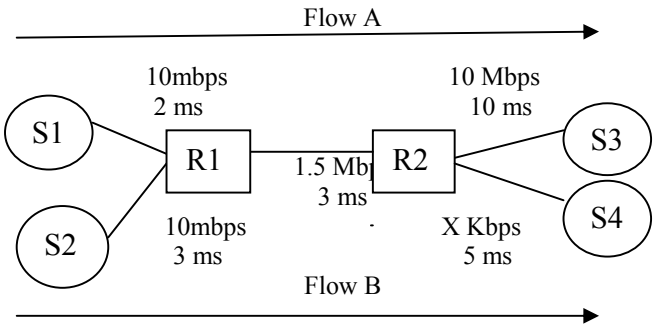

Figure 1: Example of a network which experiences congestion collapse

3. REVIEW OF LITERATURE

Several approaches came to avoid the congestion collapse. Floyd and Fall have approached the problem of congestion collapse by proposing low-complexity router mechanisms. Their suggested approach requires selected gateway router to monitor high-bandwidth flows in order to determine whether they are responsive to congestion. But they can't identify the flow rates and unresponsive flows are somewhat arbitrary and not always successful. ERICA, ERICA+ are designed for the ATM Available Bit Rate service and require all network switches to compute fair allocation of the current internet, because they violate the internet design philosophy of keeping router implementations simple and pushing complexity to the edge router.

4. NETWORK BORDER PATROL

Network Border Patrol is a core-stateless congestion avoidance mechanism. The basic principle of NBP is to compare the border of networks, the rates at which packets flow entering the network and leaving the network.

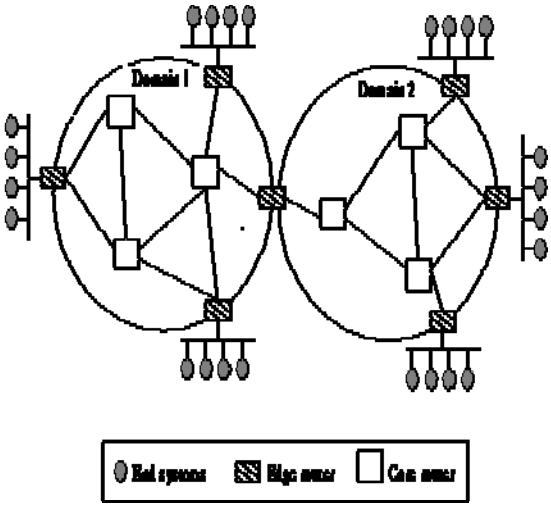

Figure 2: The core-stateless Internet architecture assumed by NBP

Figure 2 represent distinction between two types of edge routers. An edge router operation on a flow passing in to a network is called ingress router and whereas an edge route operating on a flow passing out of a network is called an egress router. NBP prevents congestion collapse through a combination of per-flow rate monitoring at egress routers and per-flow rate control an ingress routers. Rate monitoring allows an egress router to determine how rapidly each flow's packets are leaving the network; rate control allows an ingress router to police the rate at which each flow's packets in the network. Two functions are used to feedback packets exchanged between ingress and egress routers. Ingress router sends egress routers forwards feedback packets to inform them the flows are being rate controlled. NBP introduced an added communication overhead, in order for an edge router edge router to know the rate at which packets are leaving the network and must exchange feedback with other edge routers.

Three important aspects of NBP mechanism:
1) The architectural components
2) The feedback control algorithm
3) Rate control algorithm

4.1. Architectural Components

There are two types of ports used in the architectural components.

1) Input Port (Egress Router)
2) Output Port (Ingress Router)

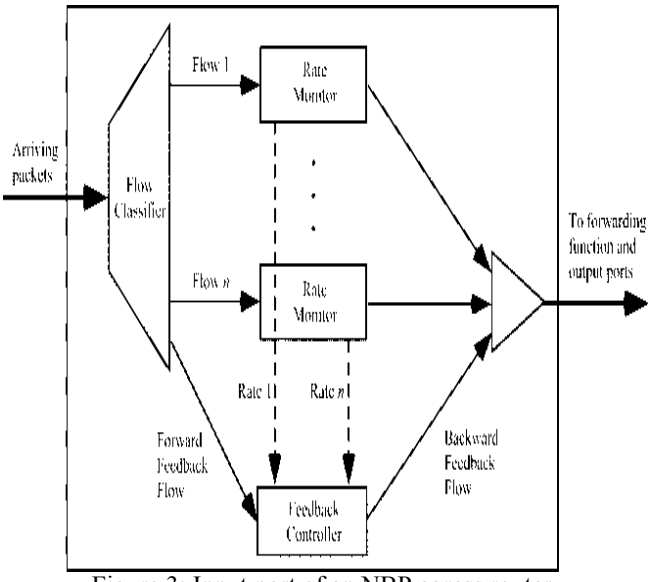

Figure 3: Input port of an NBP egress router

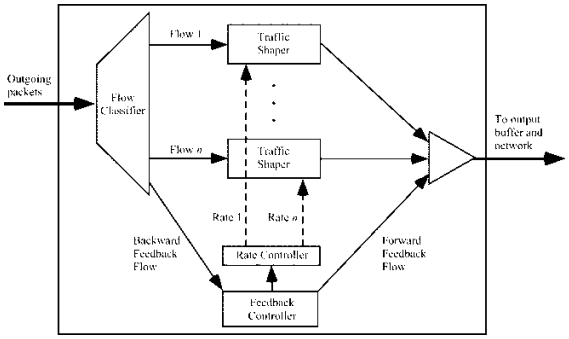

Figure 4: Output port of an NBP ingress router

The input port of egress router is used to perform per-flow monitoring of bit rates, and output ports of ingress router is used to perform per-flow rate control. The both ingress router and egress router is used to exchange and handle the feedback. Figure 3 shows that packets send by ingress routers arrive at the input port of the egress router and first classified by flow. IPV6 used to store the packet header flow label, and IPV4 used to store the packet source and destination address and port numbers. Rate monitoring algorithm such as Time Sliding Window (TSW) used to monitor the each flows bit rate. These rates are collected by a feedback controller. In Figure 4, the flow classifier classifies packets into flow, and the traffic shapers limit the rates at which packets from individual flows enter the network. The feedback controller receives backward feedback packets returning from egress routers and passes their contents to the rate controller. It also generates forward feedback packets and it periodically transmits to the networks egress routers.

4.2. Feedback Control Algorithm

The NBP feedback control algorithm determines how and when feedback packets are exchanged between edge routers. Feedback packets take the form of ICMP packets for three

reasons. First, they allow egress routers to discovers which ingress routers are acting as source for each flows they are monitoring. Second, they allow egress routers to communicate per-flow bit rates to ingress routers. Third, they allow ingress routers to detect network congestion and control their feedback generation intervals by estimating edge-to-edge round trip time. Forward feedback packet is a time stamp and a list of flow specifications for flows originating at the ingress router.

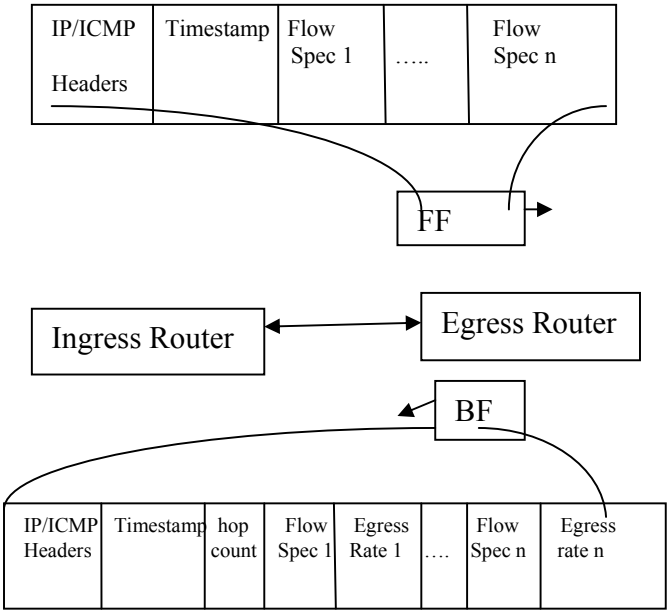

Figure 5: Forward and Backward feedback packets exchanged by edge routers

The time stamp is used to calculate the round trip time between two edge routes and the list of flow specifications indicates to an egress router the identities of active flow originating at he ingress router. When egress router receives a forward feedback packet, it immediately generates backward feedback packets and returns it to the ingress router. It contains within the backward feedback packets are the forward feedback packets original time stamp, round hop count, and a list of observed bit rates. The round hop count is used by the ingress routers rate control algorithm, which indicates how many routers are in the path between ingress router and egress router. The egress router determines the hop count by examining time to live (TTL) field of arriving forward feedback packets. When the backward feedback packets arrive at the ingress router, its contents are passed to the ingress routers rate controller. Egress router does not receive a forward feedback from an ingress router within a fixed interval of time. It generates and transmits backward feedback packets to the ingress router.

4.3. Rate Control Algorithm

The NBP rate control algorithm regulates the rate at which flow enter the network. The goal of rate control algorithm is to set the flow transmission rates that prevent congestion collapse from undelivered packets. The NBP rate control algorithm, shown in Figure 6, a flow may be in one of two phases, slow start or congestion avoidance, which are similar to the phase of TCP congestion control. New flows enter the network in the slow start phase and proceed to the congestion avoidance phase only after the flow has experienced congestion. The rate control algorithm is invoked whenever a backward feedback (BF) packet arrives at an ingress router. Recall that egress routers send two types of BF packets to ingress router: normal BF packets, which are generated when an egress router receives a forward feedback (FF) packet, and asynchronous BF packets, which egress routers generate without any prompting from an ingress router. Both types of BF packets contain a list of flows arriving at the egress router from the ingress router as well as the monitored egress rate for each flow. Only the normal BF packets contain meaningful time stamps which are copied from arriving FF packets. If the arriving BF packet is a normal BF packet, then the algorithm calculates the current round trio time and updates the base round trip time. It calculates delta RTT, which is the difference between the current round trip time (e.currentRTT) and the base round trip time (e.baseRTT). A deltaRTT value greater than zero indicates that packets are requiring a longer time to traverse the network. NBP rate control algorithm decides that a flow is experiencing congestion whenever it estimates that the network has buffered the equivalent of more than one of the flow's packets at each round hop.

```
on arrival of BF packet p from egress router e
if (p.asynchronous == FALSE)
   e.currentRTT = cur_time - p.timestamp;
   if (e.currentRTT < e.baseRTT)
       e.baseRTT = e.currentRTT;
   deltaRTT = e.currentRTT - e.baseRTT;
   for each flow f listed inp
 f.mrc = min (MSS / e.currentRTT, f.egress_rate / MF);
if (f.phase == SLOW_START)
   if (deltaRTT□□ f.ingress_rate < MSS□□ e.hopcount)
      f.ingress_rate = f.ingress_rate□□ 2;
  else
      f.phase = CONG_AVOID;
if (f.phase == CONG_AVOID)
   if (deltaRTT□□ f.ingress_rate < MSS□□ e.hopcount)
       f.ingress_rate = f.ingress_rate□□ f.mrc;
else
.ingress_rate = f.egress_rate - f.mrc;
     else /* p.asynchronous TRUE */
       for each flow f listed in p
            if (f.phase == SLOW_START)
              f.egress_rate□□□□)
                  f.ingress_rate = f.egress_rate - f.mrc;
                  f.phase = CONG_AVOID;
            else /* f.phase == CONG_AVOID */
if (f.ingress_rate > f.egress_rate +□□□□ f.mrc)
                  f.ingress_rate = f.egress_rate - f.mrc;
```

Figure 6: Pseudocode for ingress router rate control algorithm

4. CONCLUSION

Unlike existing internet congestion control approaches, which rely on end-to-end control, NBP is able to prevent the congestion collapse from undelivered packets. NBP requires no modifications to core routers nor to end systems. Buffering of packets in carried out in the edge routers rather than in the core routers. The packets are sent into the network based on the capacity of the network and hence there is no possibility of any undelivered packets present in the network. Only edge routers are enhanced so that they can perform the requisite per-flow monitoring, per-flow rate control and feedback exchange operations. The feedback-based traffic control mechanism, stability is an important performance concern in NBP. Fair allocation of bandwidth is ensured using the Network Border Patrol and this avoiding the congestion in the network.

REFERENCES

[1] S. Floyd and K. Fall, "Promoting the use of End-to-End Congestion Control in the Internet," IEEE/ACM Transactions on Networking, August 1999, to appear.

[2] B.Suter, T.V. Lakshman, D. Stiliadis, and A.Choudhury, "Design Considerations for Supporting TCP with Per-Flow Queueing," in Proc. Of IEEE Infocom '98, March 1998, pp.303-314

[3] I. Stoica, S.Shenker, and H.Zhang, "Core-stateless Fair Queueing: Achieving Approximately Fair Bandwidth Allocation in High Speed Networks," in proc. Of ACM SIGCOMM, September 1988, pp. 118-130.

[4] S. Floyd and K. Fall. "Router Mechanisms to Support End-to-End Congestion Control".

[5] V. Jacobson "Congestion Avoidance and Control", in proceeding of SIGCOMM '88 (Stanford, CA, Aug.1988), ACM

[6] Jain, R., Ramakrishnan, K., and CHIC, D.M. Congestion Avoidance in Computer Networks with a Connectionless Network Layer. Tech.Rep.DEC-TR-506, Digital Equipment Corporation, Aug. 1987.

[7] Nagle, J. "Congestion Control in IP/TCP Internetworks", ARPANET Working Group Request for comment, DDN Network Information Center, SRI International, Menlo Park, CA, Jan. 1984. RFC-896.

[8] Mills, D. "Internet Delay Experiments", ARPANET Working Group Requests for Comment, DDN Network Information Center, SRI International, Menlo Park, CA, Dec.1983. RFC-889.

[9] Jain, R. "A Timeout-Based Congestion Control Scheme for Window Flow-Controlled Networks", IEEE Journal on Selected Areas in Communications SAC-4, 7(Oct. 1986).

[10] K. Thompson, G. Miller, and R. Wilder. "Wide-Area Internet Traffic Patterns and Characteristics", IEEE Network, 11(6):10-23, Nov.1997.

[11] G. Varghese "On Avoiding Congestion Collapse", viewgraphs, Washington University Workshop on the Integration of IP and ATM, Nov.19 1996

[12] B.Brsden et al, "Recommendations on queue management and Congestion Avoidance in the Internet, " RFC 2309, IETF, April 1998.

[13] A. Parekh and R. Gallager, "A Generalized Processor Sharing Approach to Flow Control- the single node case," IEEE/ACM Transactions on networking, vol. 1, no. 3, pp. 344-357, June 1993.

[14] W. Stevens, "TCP Slow Start, Congestion Avoidance, Fast Retransmit, and Fast Recovery Algorithms," RFC 2001, IETF, January 1997.

[15] W.K. Tsai and Y. Kim, "Re-Examining Maximum Protocol: A Fundamental study on Convergence, Complexity, Variations and Performance," in Proc. Of IEEE Infocom, April 1999, pp.811-818.

[16] D. Bertsekas and R. Gallager, Data Networks, Second Edition, Prentice Hall, 1987.

[17] D. Lin and R. Morris, "Dynamics of Random Early Detection," in proc. Of ACMSIGCOMM, September 1997, pp.127-137.

[18] B. Vandalore, S. Fahmy, R. Jain, R. Goyal, and M. Goyal, " A definition of generalized fairness and its support in switch algorithms," AT Forum, Doc. 98-0151, Traffic Management WG, Feb.1998.

[19] C. Partridge, J. Bennett, and N. Shectman, "packet recording is not pathological network behavior," IEEE/ACM Trans. Networking, Vol. 7.

[20] R. Pan, B. Prabhakar, and K. Psounis, "Choke-A stateless active queue management scheme for approximating fair bandwidth allocation. " in proc. IEEE INFOCOM, Mar.2000, pp.942-951.

[21] U. Hengartner, J. Bolliger, and Th. Gross, "TCP Vegas revisited," in Proc. IEEE INFOCOM, Mar.2000, pp.1546-1555.