A PREDICTIVE CONTROL APPROACH BASED ON THE VIRTUAL COLLISION CONCEPT FOR BEST EFFORT TRAFFIC CONTROL OF IP NETWORKS

B.A.Costa*, **M.S.Nunes**†, **J.M.Lemos**‡

*†‡IST/INESC-ID
Rua Alves Redol 9; 1000-029 Lisboa-Portugal
*e-mail:bac@digitais.ist.utl.pt †e-mail:Mario.Nunes@inesc.pt ‡e-mail:jlml@inesc.pt
for European Control Conference 2003

Abstract

In this paper, a congestion control algorithm for best-effort traffic in IP networks is proposed. The aim is to avoid congestion and loss of packets, and also to share the bandwidth of a congested link among the flows that are using it. Each client provides the flow rate value that would like to obtain, and a time value (round-trip delay) that is used by routers to compute the number of active flows. One assumes that routers are allowed to use a field in the packets to propagate the reference of the flow rate, in principle, it can be done using the field option of the IP header. The clients use the information that is provided by the network and the information that is obtained by probing the Internet to adjust the servers' flow rates. In the paper on follows the principle that clients do not need to increase their flow rates to obtain information about the congestion state of the network.

Keywords: Predictive Control, Internet Traffic, Fairness Concept, Flow Concept.

1 Introduction

Telecommunication networks are considered to be a field where automatic control methodologies can improve resource sharing, flow and congestion control. Two of the most important telecommunication network technologies are ATM and Internet. In the context of ATM technology, the Available Bit Rate (ABR) Service has been used to test several types of automatic controllers.

1.1 ATM Networks

ATM technology is based on the concept of resource reservation to provide Quality of Service. Several services were defined to provide QoS based on a description of the traffic that will be generated by the connection end points. CBR- *a constant bit rate service* and VBR *a variable bit rate service* are two classes that use the resource reservation concept. Traffic belonging to such services is considered uncontrollable. ABR, *an available bit rate service*, is an ATM class of service that can be used to improve the network resource utilization by services that can adapt their traffic to the state of the network, to the available bandwidth.

ABR traffic is composed of data cells and resource management cells (RM cells). RM cells are sent every N-1 data cells by a source to carry information about the transmission rate. Switches along the communication path, between the source and receiver, can change RM cells to inform about the traffic congestion state. Several control methodologies using congestion information were proposed for controlling the source flow rate [9],[1],[8],[7]. In [9] the Smith principle was investigated for ABR traffic, [1],[8] use the LQ methodology. In [10] an explicit-flow control algorithm is developed considering loss and fairness constraints.

1.2 Internet Networks

Internet is a technology for data transmission, it is based on the IP protocol and it was developed with no QoS guarantees. In Internet, routers do not send congestion information on IP packets. Routers implement simple functions and treat packets equally, store-and-forward packets in a connection-less way. The communication end points perform complex function such as packet re-transmission, error detection on packets and flow rate adjustment. The best-effort service is used, meaning the network will try to deliver packets to the receiver but it will not compromise with. Packets have variable size and it is not assured that they go always through the same path. Internet does not have an entity for shaping the traffic behavior. Users (traffic sources) can cause congestion by injecting traffic at rates beyond the network transport capacity. Queues are used to avoid the transient lower transport capacity by delaying packets, however packets can be dropped when the queues are overflowed.

The TCP protocol (that runs on top of IP protocol) was developed to overcome those situations, enabling a connection oriented and error free services. TCP implements a mechanism that performs flow control (a window is used to control the number of packets that the receiver can handle) and congestion control (a window is used to control the number of packets that the network can handle without dropping them). Using those two windows, the sender can send packets at its own rate without waiting for the acknowledgment of the receiver. The source uses the detection of a lost packet to adjust the congestion window, and by that it decreases its flow. A timeout is generated if the receiver does not acknowledge the reception of packets. But because the source can send packets in bursts, it can cause congestion, forcing routers to drop packets.

UDP is another transport protocol used in Internet without mechanisms for flow control, or congestion control. This is explored by some applications to get more resources (bandwidth) from the network, than they would get if they were using the TCP, they behave as "TCP-unfriendly" [4]. This behavior has been criticized [12],[3] and suggests that routers must have detection mechanisms to avoid the problem.

1.3 Internet with Quality of Service

DiffServ [2],[6] defines a model to provide QoS in the Internet. Flows are treated as aggregates. A field in the IP header named DS-Differentiated services field is inspected by routers and is used to select the type of behavior that packets should be treated with [11],[5]. For the moment there are three types of differentiated services, EF (*Expedited Forwarding*), AF (*Assured Forwarding*) and BE (*Best Effort*). The EF service is defined so that packets leave a router with at least a minimum departure rate ensuring that the aggregated flow has a maximum latency rate, and a very low dropped packet value and jitter. The AF service is divided in four classes each having three precedence orders of dropping packets. A router drops packets with the lowest levels when it is under congestion or when is preventing congestion. The BE service corresponds to the present Internet service.

1.4 Aims of the paper

In this paper one tries to tackle the congestion control problem and fair share of resources in the context of Internet, which has a direct impact on the Quality of Service provided by the network. Note however that Internet does no provide explicit information about resource utilization. Connections induce congestion in links to generate resource sharing information, namely bandwidth sharing. In the paper on follows the Internet's principle of putting complex tasks at end points and to reduce complexity of tasks at routers, and also the principle that clients do not need to increase their flow rates to obtain information about the congestion state of the network.

The aims are: Maximization of data transfer rate (throughput) using the available network bandwidth; Minimization of packets loss and network congestion level; Minimization of packet delay and jitter;

2 Service Model Concept

To develop the service model one uses the client-server concept. A client and a server are the end points and they use network's links to exchange packets. The client uses packets to request data packets from the server. The server when receiving a packet from the client sends the data packets that are requested by the client. The client also provides to the network the flow rate value that would like to obtain and a time value (round-trip delay) which is used by routers to compute the number of active flows. Routers are allowed to propagate the flow rate references by changing the field in packets that clients use to request a flow rate. The reference signal that is received by a client gives the flow rate value that it must use to request packets from a server. In addition the client uses the round-trip delay to control the number of packets that are in the network. Fig.1 shows a simple network. Routers are the network elements that enable the routing of packets. A router has a routing table which is used to select the output port for a packet based on the destination address of the packet. One assume that routers have buffers located at the output ports to store packets waiting to be sent. To tackle the problem one assumes that all packets have the same size. Links have the same bit rate but they have different lengths. Each router keeps information about active flows in a table (database) named collision table which is used to compute the fair bandwidth share. The process of computing the bandwidth share named virtual collision concept.

Note that Internet routers use a simple scheduling strategy like round-robin, and a network has an hierarchy (of routers). If each server (or client) can use the total band-

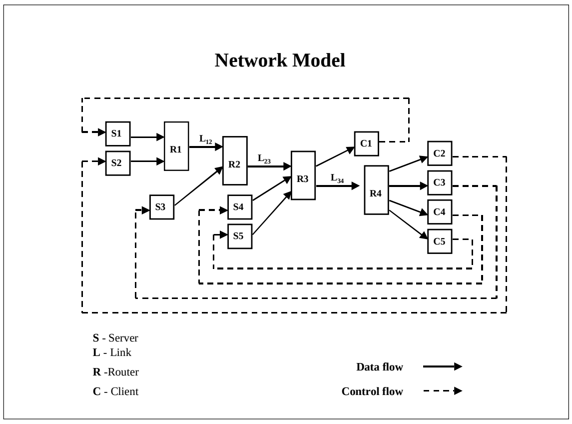

Figure 1: Network model and Service model.

width of a link, during a long time interval, the server (client) that is near the congested link gets more bandwidth than the other servers (clients) that are at a different level of the network hierarchy. For the network in fig.1, and assuming that S4 and S5 are off, server S3 gets 50% of the bandwidth of link (L_{23}) between R2 and R3 and the others, S1 and S2, get 25% each.

In this paper, flows (clients) are classified in two classes. A client is classified at a router as being satisfied if it requests a flow rate value lower or equal to the value that is computed/stored at the collision table in a router. The fair bandwidth share is obtained when the available bandwidth of the congested link is equally distributed among the unsatisfied flows that are using the congested link. Note that, if all flows, (c) are not satisfied then each flow will get $1/c$ of the output bandwidth. A router does not police flows, it compute a reference signal that clients use to adjust their flows. The policing is performed by the controllers at clients. Note that controllers can be located at the server, but because a server must handle requests from several clients that will increase the amount of work that a server must execute.

To compute the fair bandwidth share (the flow rate reference signal), routers keep track of active flows, where source and destination addresses of packets, time values and flow rates are stored. Routers decrease the time values in the collision table and renewal the information about a flow when a packet is received. In a router, flows are active if they have positive time values. The information about a flow is discarded when the time value reaches zero. Each router is allowed to decrease the flow rate reference signal in each packet, but it is not allowed to increase it. This enables a client to get the reference signal of the congested link that has the highest value of active flows.

Using this approach, the reference signal is not dependent on the number of packets that are waiting in queues and clients do not need to increase their real flow rate to probe the state of the network, this helps to eliminate the congestion problem.

2.1 The controller

The controller is located at the client. One assumes that the client must send one "control" packet to obtain one data packet from the server. One can uses others strategies to reduce the number of control packets in the network, for example by describing, in a control packet, a micro flow (number of packets and flow rate) that the server must send, but this approach will increase the action delay of a client. Note that if a packet of the client is lost then the server will not receive it and as a consequence the server will not inject data packets on the network.

The controller has two modes of operation. Mode S for startup, and mode N for normal operation.

In mode S the controller sends a packet and waits for a data packet from the server. The client's packet is seen by routers which detect the presence of one more active flow. The packet from the server can be used to inform the client about the total amount of data that the server has available, such as the size of a file. The server must copy (or decrease) the reference signal value and the time value to data packets.

The controller changes to mode N as soon it receives the first data packet. With the reception of a data packet, the client obtains a new value for the reference signal (flow rate) r_{fr}. From this point in time the client requests packets to the server using the reference that is provided by the network.

2.1.1 The control law of mode N

The control law is based on the prediction and control of the number of packets that a client can have in the network. Note that small errors in the rate at which the client is requesting packets from the server can cause packets being stored in routers. Knowing the reference signal and the round-trip delay due to propagation R_{tdp}, the number of packets that a client can have in the network is given by $N_r = r_{fr} * R_{tdp}$. However the client is only able to measure the round-trip delay R_{TD} for each request/data packet, which includes the round-trip delay due to propagation and the waiting time in buffers. To estimate R_{tdp}, the following formula is used $\hat{R}_{tdp}(t) = min\ R_{TD}(i)$ with $0 \leq i \leq t$ where t denotes the current time instant.

Let $n(.)$ denote the number of packets a client has in the network,

$$n(t+1) = n(t) + u(t) - y(t) \tag{1}$$

where $y(.) \in \{0,1\}$ denotes a data packet coming out of the network, $u(.) \in \{0,1\}$ denotes a request of the client.

The control law for requesting data packets by the client is given by

$$u(t) = \begin{cases} 1 & \text{if } n(t) < n_r(t) \text{ and } u_r(t) \geq 1.0 \\ 0 & \text{if } n(t) < n_r(t) \text{ and } u_r(t) < 1.0 \\ 0 & \text{if } n(t) \geq n_r(t) \end{cases} \tag{2}$$

where $n_r(t)$ is the reference signal for the number of packets (control/data) that a client has in the network, $u_r(t)$ is obtained from the flow rate reference signal $r_{fr}(t)$ by $u_r(t) = u_r(t-1) + r_{fr}(t) - d(t)$ where

$$d(t) = \begin{cases} 0 & \text{if } u_r(t-1) + r_{fr}(t) < 1.0 \\ 1.0 & \text{if } u_r(t-1) + r_{fr}(t) \geq 1.0 \end{cases} \tag{3}$$

2.1.2 The reference signal for the number of packets in the network

The network provides the reference signal for the flow rate $r_{fr}(t)$, this reference signal represents the fair share at a congested link along the shared path. In order to compute a reference for the number of packets $n_r(t)$, the client estimates $\hat{R}_{tdp}(t)$ and computes $N_r(t) = r_{fr}(t) * \hat{R}_{tdp}(t)$. The reference signal $n_r(t) = N_r(t)$ if $r_{fr}(t-1) < r_{fr}(t)$, and decreases linearly at the rate $(N_r(t) - N_r(t-1))/\hat{R}_{tdp}(t)$ if $r_{fr}(t-1) \geq r_{fr}(t)$ Using this approach the client can request packets at the rate given by the network but avoiding requesting gaps caused by a sudden decrease in r_{fr}.

2.1.3 Requesting additional bandwidth

To probe the network for additional bandwidth, a client specifies the new flow rate value based on the current flow rate reference by $B_n(t) = r_{fr}(t) * \alpha$, where $\alpha > 1.0$. $B_n(t)$ is used by routers to assign bandwidth and to classify the flow as satisfied or not satisfied. As a consequence routers that are located before a congested link are assigning flow rates values above $r_{fr}(t)$. Small values for α enable a better use of unused bandwidth but the adjustment of the flow rate reference will be slow for flows that can use the available bandwidth. If α has a very large value, then a larger bandwidth value is assigned to downstream congested flows preventing other flows of using that bandwidth. The simulation results that are shown use $\alpha = 1.1$.

2.2 Comparison with TCP protocol

The control law that is implemented by the TCP protocol has a congestion window which is used to control the number of packets that the connection has in the network. The congestion window is tuned based on the packet loss rate. The TCP protocol does not provide fair share of resources and does not prevent bursts of traffic which can cause buffer overflow. The control algorithm that is described in this paper tries to solve the fair share of resources with the help of routers, and tries to avoid bursts in traffic by sending packets at constant rate, depending of the reference that is provided by the network and of the number of packets that a connection has in the network. This also helps in the identification of the round-trip propagation delay which is used to compute the reference for the number of packets that a connection can have in the network.

3 Simulation Results

In this section simulation results with the control algorithm are shown.

3.1 Network model

The network model is shown in fig.1. A discrete time simulator was used with a simulation step equal to the duration of a packet. Servers, clients and routers delay packets by one discrete time unit. The round-trip delays, in discrete time units, due to propagation observed by client i ($RtoC_i$) are, $RtoC_1 = 24$, $RtoC_2 = 26$, $RtoC_3 = 20$, $RtoC_4 = 12$, $RtoC_5 = 11$.

3.2 Simulation Results

A computer simulation experiment was done where clients C1, C2, C3, C4, C5 send their first packet to servers S1, S2, S3, S4, S5 respectively at discrete time instants $t_1 = 1$, $t_2 = 20$, $t_3 = 40$, $t_4 = 60$ and $t_5 = 75$. Client C1 begins by getting the maximum rate of 1 packet per time instant. An important point to note is that when C1, C2 and C3 are on, and C4 and C5 are off the link L_{23} is congested and its bandwidth is equally distributed among C1, C2, C3. However when C1, C2, C3, C4, C5 are on, the link L_{34} is the one that is congested forcing clients C2, C3, C4, C5 to get an equal flow rate reference of 0.25. This has a consequence that the link L_{23} is no longer congested and the available bandwidth can be used by client C1. Fig.s 2, 3 and 4 show the number of packets

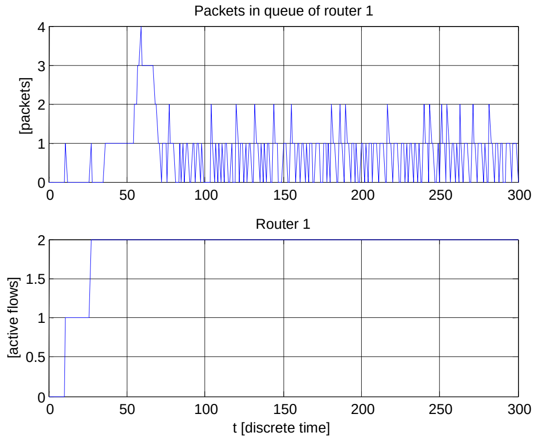

Figure 2: Number of packets waiting in queue of router R1 and the active flows in router R1.

waiting in queues and the number of active flows in routers R1, R2 and R3. Fig.5 shows the flow rate reference of

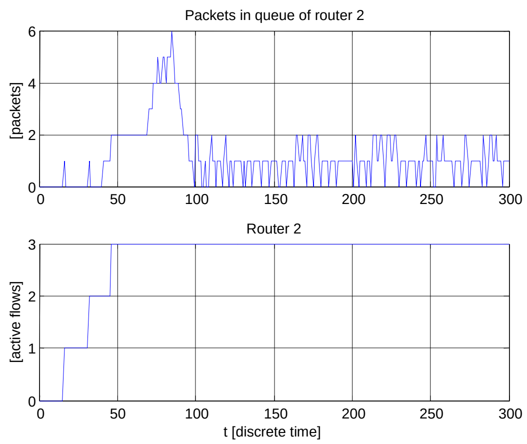

Figure 3: Number of packets waiting in queue of router R2 and the active flows in router R2.

the network for client C1 and reference and number of packets in the network of the flow of client C1. Note that the flow rate reference increases after $t = 100$ because clients C2 and C3 are restricted by the bandwidth share in link L_{34} to 0.25 each, and C1 can use the available bandwidth of L_{23}. In this case the maximum flow rate reference value for C1 is 0.45 above of 0.333. Figure 6 shows the round-trip delay observed from client C1 and its minimum value, packets requested and received by client C1. The simulation results for C2, C3, C4 and C5 are shown in fig.s 7, 8, 9, 10, 11, 12, 13, 14. From the simulation results one can observe that the number of packets waiting in queues is bounded to a lower value, and the bandwidth is equally shared among the flows that are using a congested link.

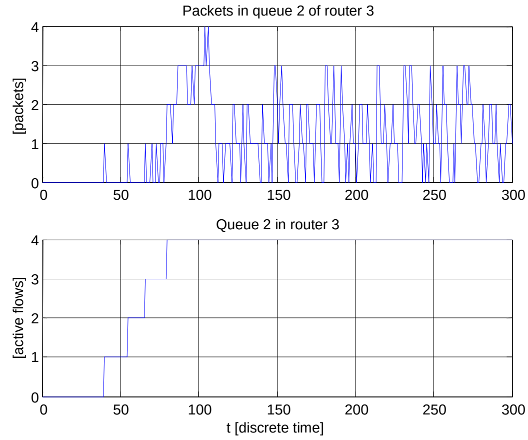

Figure 4: Number of packets waiting in queue 2 of router R3 and active flows in queue 2 of router R3.

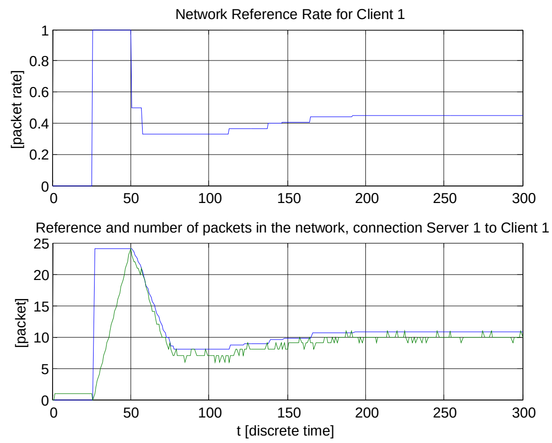

Figure 5: Reference from the network for client C1 and reference and number of packets in the network of the flow client C1/server S1.

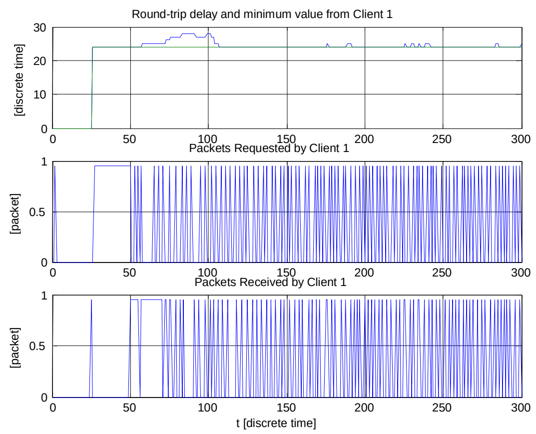

Figure 6: Round-trip delay observed from client C1, packets requested and received by client C1.

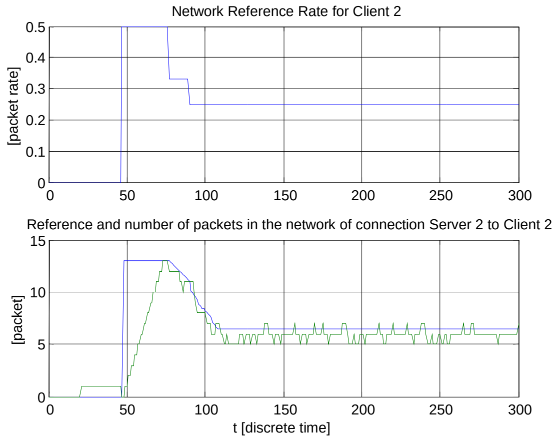

Figure 7: Reference from the network for client C2 and reference and number of packet in the network of the flow server S2/client C2.

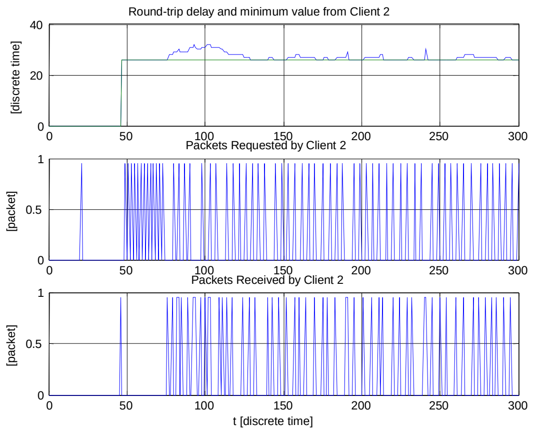

Figure 8: Round-trip delay observed from client C2, packets requested and received by client C2.

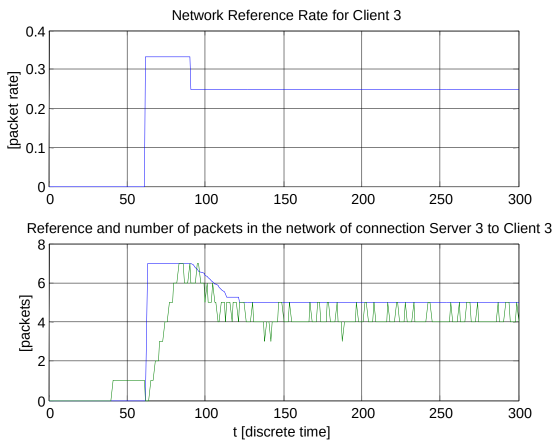

Figure 9: Reference from the network for client C3 and reference and number of packet in the network of the flow server S3/client C3

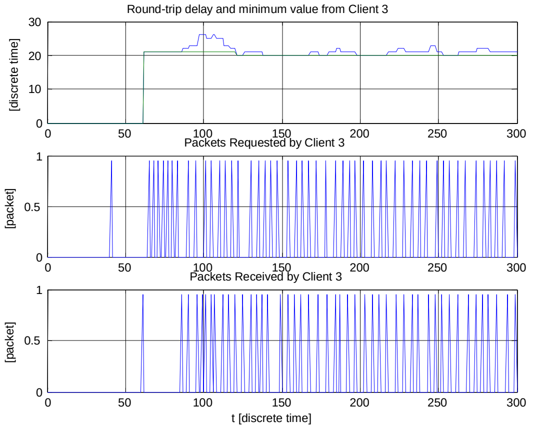

Figure 10: Round-trip delay observed from client C3, packets requested and received by client C3.

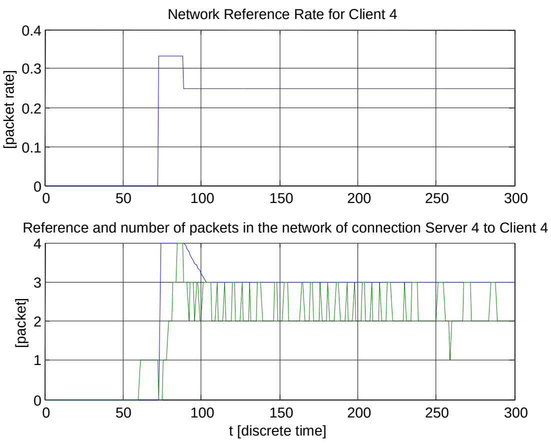

Figure 11: Reference from the network for client C4 and reference and number of packet in the network of the flow server S4/client C4

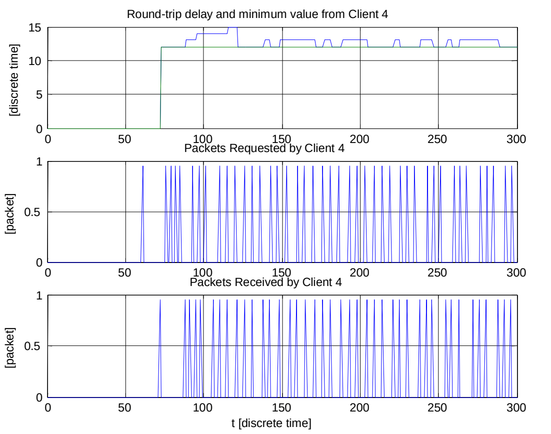

Figure 12: Round-trip delay observed from client C4, packets requested and received by client C4.

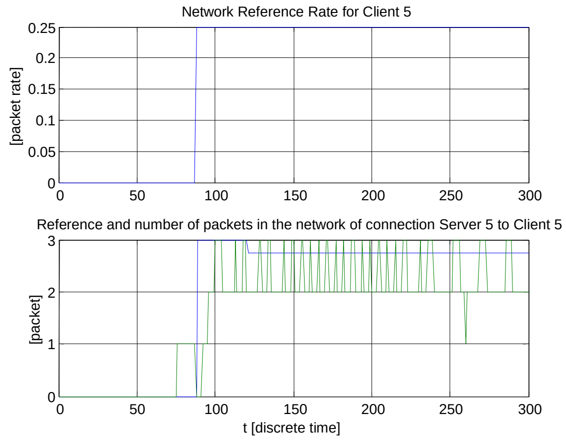

Figure 13: Reference from the network for client C5 and reference and number of packet in the network of the flow server S5/client C5

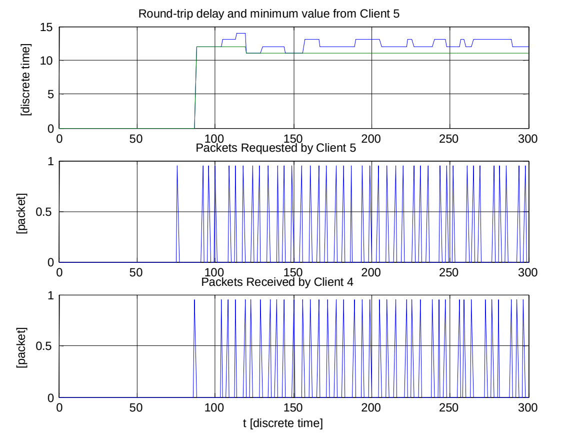

Figure 14: Round-trip delay observed from client C5, packets requested and received by client C5.

4 Conclusion

In this paper a service model for the Internet was proposed, that uses the client-server concept, and the propagation of information about resource utilization based on the active flow and virtual collision concept. Routers mark packets with a reference signal that gives the flow rate at which a client can request packets. In principle, the field option of the IP header can be used to carry the flow rate reference value. Routers are allowed to decrease (but not to increase) the reference signal value of incoming packets depending on the number of active flows that are using an output link. Routers have a collision database to keep track of active flows. Routers decrease the time values in the databases and renewal the information about a flow when a packet is received. In a router, flows are active if they have positive time values. The information about a flow is discarded when the time value reaches zero. The control law that is implemented at clients, generates requests based on the reference from the network and also on a prediction of the number of packets that are allowed in the network. This value is computed at the client based on estimates of the round-trip delay of packets.

Computer simulation were done to evaluate the concept. From the simulation results one can observe that the number of packets waiting in queues is bounded to a lower value, and the bandwidth is equally shared among the flows that are using a congested link. The lower values that were obtained is due to behavior of the clients, before sending packets they get information about the state of the network and use it to shape the request of packets. Also to probe the state of the network a client does not need to increase their real flow rate. This does not happens in Internet with the TCP protocol. The results can be improved if the round-trip time values of competing flows are incorporated in the control law. Further work must be done to improve the selection of parameter α used to assign the available bandwidth at routers. The estimation of round-trip delay due to propagation must be improved to handle a possible change in the path of packets due to an update of routing information in routers.

References

[1] E. Altman, T.Basar, and R. Srikant. Congestion control as a stochastic control problem with action delays. *Automatica*, 35:1937–1950, 1999.

[2] D. Black, S. Blake, M. Carlson, E. Davies, Z. Wang, and W. Weiss. RFC2475 - An Architecture for Differentiated Services. Technical report, URL http://www.ietf.org, Dec-1998.

[3] S. Floyd and K. Fall. Promoting the Use of End-to-End Congestion Control in the Internet. *IEEE/ACM Trans. on Networking*, May 1999.

[4] P. Gevros, J. Crowcroft, P. Kirstein, and S. Bhatti. Congestion Control Mechanisms and the Best Effort Service Model. *IEEE Network*, 15 No3:16–26, May/June 2001.

[5] J. Heinanen, F. Baker, W. Weiss, and J. Wroclawski. RFC2597 - Assured Forwarding PHB Group, June-1999.

[6] G. Hustom. RFC2990 - Next Steps for IP QoS Architecture. Technical report, URL http://www.ietf.org, Nov-2000.

[7] O. C. Imer and T. Basar. Control of Congestion in High-Speed Networks. *European Journal of Control*, 7:132–144, 2001.

[8] O.C. Imer, S. Compans, T. Basar, and R. Srikant. Available Bit Rate Congestion Control in ATM Networks. *IEEE Control System Magazine*, 21:38–56, Fev-2001.

[9] S. Mascolo. Congestion control in high-speed communication networks using the Smith principle. *Automatica*, 35:1921–1935, 1999.

[10] C-F. Su and G. Veciana. Explicit Rate Flow Control for ABR Services in ATM Networks. *IEEE/ACM Trans. on Networking*, 8 N.3:359–361, June-2001.

[11] K. Nicols K. Poduri V. Jacobson. RFC2598 - An Expedited Forwarding PHB. Technical report, URL http://www.ietf.org.

[12] J. Widmer, R. Denda, and M. Mauve. A Survey on TCP-Friendly Congestion Control. *IEEE Network*, 15 N.3:28–37, May/June 2001.