



ELSEVIER

Performance Evaluation 38 (1999) 243–252

**PERFORMANCE
EVALUATION**
An International
Journal

www.elsevier.com/locate/peva

Improving data transport performance through SSCOP with rate-based flow control

Jin-Ru Chen*, Yaw-Chung Chen

*Department of Computer Science and Information Engineering, National Chiao Tung University,
Hsinchu 30056, Taiwan, ROC*

Received 5 October 1998; received in revised form 21 June 1999

Abstract

This work tries to improve the effective throughput as well as to achieve better fairness of the transport protocol with efficient selective retransmission scheme. The commonly used transport layer protocol, transmission control protocol (TCP), has been developed for more than two decades, but its retransmission scheme is hard to be satisfactory. However, service specific connection oriented protocol (SSCOP) provides an excellent retransmission mechanism with poor flow control function. When the lower layer protocol is Asynchronous Transfer Mode, using SSCOP instead of TCP/IP over AAL5 to provide transport layer service can specifically reduce the protocol overhead. We propose a scheme for SSCOP to enhance its flow control performance. This scheme combines the rate-based flow control with the queuing delay control. The simulation result shows that our enhanced SSCOP achieves both higher throughput and better fairness. It demonstrates that a proper flow control scheme with efficient selective retransmission is crucial for optimizing the performance of a transport protocol. ©1999 Elsevier Science B.V. All rights reserved.

Keywords: Selective retransmission; TCP; SSCOP; Rate-based flow control

1. Introduction

The development of asynchronous transfer mode (ATM), Ethernet and Layer3 switching techniques significantly improves the network capacity as well as reduces the packet loss caused by some multi-access contention techniques. Both wireless and satellite networks are also developed for various missions. Basically, the congestion avoidance mechanism can be divided into the flow control part and the error recovery part. When trying to improve the protocol to achieve good performance under different network environments, both two parts need to be taken care. Due to the popularity of transmission control protocol (TCP), a variety of improvements have been developed to enhance its performance. Since the selective retransmission scheme was not provided originally, thus the selective retransmission function has relatively lower efficiency comparing with the flow control function.

* Corresponding author. Tel.: +886-3-5731864; fax: +886-3-5727842
E-mail address: jzchen@csie.nctu.edu.tw (J.-R. Chen)

After examining various protocols with end-to-end transport capability, we find that service specific connection oriented protocol (SSCOP) [3] features an excellent retransmission mechanism but quite poor flow control function. There is still quite a lot of room for further improvement in its flow control part. Previously proposed scheme [2] has improved the efficiency of its selective retransmission, a more efficient flow control mechanism with simple implementation complexity is still necessary to further enhance the performance of SSCOP. Besides the performance issue, since the ATM network is widely used in the current environment, when the SSCOP is used to provide transport layer service, the TCP/IP protocol stack over AAL layer can be replaced. Therefore, a great part of protocol overhead can be avoided to improve the host side efficiency.

In this work, we use a probing scheme to evaluate the proper transmission rate, and employ the queuing delay control (QDC) to improve the buffer utilization as well as the fairness of bandwidth sharing. Since both rate probing and QDC help improve the precision of bandwidth sharing, the fairness can thus be guaranteed. As a result, the SSCOP can achieve a remarkable high throughput and fair resource sharing with minimum waste of network resources.

The SSCOP flow control used in this article is addressed in Section 2. Two essential control schemes, service rate probing and QDC, are discussed in Sections 3 and 4 respectively. In Section 5, we explain the enhancement of SSCOP for rate-based flow control. The simulation and numerical results are presented in Sections 6 and 7 concludes the work.

2. SSCOP flow control

In SSCOP protocol, a receiver grants the credit that allows the source to transmit further data units. The credit value depends on the availability of both network bandwidth and receiver's buffer space. This value restricts the number of outstanding data units allowed for the source transmission. The sequence number tagged on each data unit guarantees the data integrity. If there is any data unit lost, the receiver can detect the loss through the discontinued sequence number and use the selective retransmission to recover the missing data.

The control messages are exchanged based on the Poll Protocol Data Unit (POLL PDU), the Solicited Status PDU (STAT PDU) and the Unsolicited Status PDU (USTAT PDU). An SSCOP source transmits POLL PDU periodically to obtain the packet receiving status in the receiver, which returns a STAT PDU according to the received POLL PDU. The STAT PDU carries all previously detected but yet unresolved errors together with the control information back to the source. Whenever the receiver detects a packet loss, it sends an USTAT PDU immediately to request the source to retransmit the lost PDU. The source checks the STAT PDU or USTAT PDU to decide whether a transmitted PDU is acknowledged or not. The poll number is used to avoid duplicated retransmissions.

However, the process by which a receiver entity determines the credit is not subjected to standardization, and here we proposed a flow control scheme to improve the SSCOP performance. The flow control information supports two purposes: to ensure the data sequence integrity and to handle the error recovery, but not for flow control; to derive the proper source transmission rate at the receiver based on the probed service rate. Neither the STAT PDU nor the USTAT PDU provides the source rate information, therefore, we add a field in both STAT PDU and USTAT PDU so that source can use this information to adjust its transmission rate. As a sequel, we are able to take advantage of the SSCOP characteristics to design a flexible mechanism for both rate-based and window-based flow control.

3. Service rate probing

Obviously we need a mechanism to precisely estimate the service rate for each virtual connection (VC), and this rate can always be determined by the bottleneck switch. The Packet-Pair probing scheme proposed in [5] uses specific back-to-back transmitted packets to detect the service rate. When sending the probing packets, it generates bursty traffic, which is twofold the maximum packet size. This scheme may detect the bottleneck service rate incorrectly because the volume of probing traffic is relatively too small to effectively reflect the actual traffic characteristics.

To improve the accuracy of the service rate probing, we design a scheme, which transmits probing packets with a specific traffic pattern, and embeds the rate probing mechanism in the packet transmission. To minimize the traffic overhead needed for rate probing purpose, we restrict the maximum rate of probing traffic to twofold the current transmission rate. The probing process proceeds as long as the data transmission is active. This can avoid the incorrect rate information caused by interleaving probing packets with regular data traffic.

To keep the average transmission rate same as the derived rate, the source transmits packets with two different patterns alternatively. If the current transmission rate was λ , the average inter-cell duration is $I = 1/\lambda$. We use two patterns with different inter-packet intervals derived from I for generating probing packets. These two intervals are $I_s = I/2$ and $I_l = 3I/2$. Since the packet size varies, the source needs a mechanism to control the transmission rate, this can be accomplished through a counter. The counter value was initialized to zero, when the counter value is larger than zero, the source uses the shorter interval as the inter-packet duration. After a packet is transmitted, the counter decrements by the transmitted packet size. On the other hand, if the counter value is no greater than zero, the source transmits packets with the longer inter-packet interval. After this is done, the counter increments by the transmitted packet size. Since the source always tries to keep the counter value close to zero, its average transmission rate will converge to λ .

The receiver detects the packet interarrival time to derive the bottleneck service rate. In order to reduce the computing complexity, the derivation of the service rate is performed in two steps, the information collection and the rate calculation. Let the interarrival times of the probing packets be I'_s and I'_l , when a PDU arrives at the receiver, the inter-packet interval can be normalized to lower layer data unit, such as cell slots, according to the lower layer protocol characteristics. It is important to perform such normalization because the inter-packet interval is unable to represent the transmission rate accurately. The normalized intervals will not provide sufficient network status information unless the inter-packet time of both short and long intervals are obtained. These transmission intervals can be sorted to derive the network service rate, and it takes $O(n) = \log_2 n$ time complexity to accomplish the task.

The rate calculation is meaningful only when the number of records is larger than two. If this number is three, it must avoid the situation that two out of three maximum-sized packets come with long inter-packet time. We use the averaged interval of the lower one-third of total received intervals as the probed interval. The rate derivation rules are described in the following and the reciprocal of the probed interval is used as the proper source transmission rate.

- No packet loss during the poll interval. We can find k such that,

$$\sum_{i=1}^k \text{Size}(i) \leq \frac{1}{3} \sum_{i=1}^n \text{Size}(i) < \sum_{i=1}^{k+1} \text{Size}(i).$$

Then we set

$$\text{Threshold} = \frac{1}{3} \sum_{i=1}^n \text{Size}(i),$$

$$\text{Remaining} = \text{Threshold} - \sum_{i=1}^k \text{Size}(i),$$

$$\text{Threshold Interval} = 1 / \left\{ \sum_{i=1}^k \text{Size}(i) \times \text{Interval}(i) + \text{Remaining} \times \text{Interval}(k + 1) \right\},$$

where $\text{Size}(n)$ represents the size of n th packet received and $\text{Interval}(n)$ represents n th interval in lower layer protocol data unit.

- Packet loss during the poll interval. Let $(\sum_{i=1}^n \text{Size}(i) \times \text{Interval}(i)) \text{Interval} = 1 / \sum_{i=1}^n \text{Size}(i)$, where $\text{Size}(n)$ and $\text{Interval}(n)$ have the same meaning as described above.

4. Queuing delay control

Instead of controlling the queuing delay of all packets directly, we use the fixed poll interval of SSCOP to control the queuing delay of the transmitted data. Since we treat the network as a single queue under our QDC mechanism, the fixed POLL PDU is used to evaluate the queuing delay level of the network queue. The time duration from a POLL PDU transmitted to the corresponding STAT PDU received is the measured round-trip time, the queuing delay in the network can then be derived. Furthermore, the queue occupancy is also available based on the notion of virtual queue occupancy (VQO). With these collected information, the QDC mechanism can use the transmission rate reduction to maintain the queuing delay at a steady value.

4.1. VQO evaluation

VQO is an essential parameter for QDC. Calculation of VQO is based on deriving the queuing delay. This is similar to the scheme used in [5]. Basically, queuing delay is the difference between the measured PDU round-trip time and the smallest round-trip time ever occurred.

If we treat the network as a single queue, the queue occupancy can be derived from the queuing delay and the queue service rate. The formula to derive the VQO is

$$N = \mu T,$$

where N is the queue occupancy, μ the queue service rate and T the queuing delay of the detection PDU. When a detection PDU arrives at the queue in which there are already N PDUs ahead, the queuing delay expected by the detection PDU will be N/μ . Therefore, the queue occupancy is equal to the virtual queue occupancy.

4.2. Queuing delay reduction

The queuing delay reduction can be achieved by slowing down the source transmission rate. If the source transmits data with the detected rate, the queue occupancy will keep increasing unless the detection

mechanism generates a rate which becomes smaller than the available rate. Therefore, if the queue occupancy needs to be reduced and the source transmission rate is lower than the available rate, the queue occupancy reduction would have been done. Besides, since the proper transmission rate is obtained through the detection, dramatically reducing the transmission rate may affect the detection result and cause rate fluctuation. Thus, the queuing delay reduction not only has to reduce the transmission rate, but also needs to restrict the reduction ratio to avoid the rate fluctuation.

The way to calculate the proper transmission rate for reducing the buffer occupancy is to hold packet transmission for a specified period of time. This holding period is calculated using both the VQO reduction value and the proper transmission rate. But to hold transmission abruptly may affect the rate detection result, such effect can be smoothed throughout the control period. Further, to avoid the rate fluctuation, the holding period is set to a fixed portion of the control period.

4.3. Control period

The POLL PDU is used to detect the queue occupancy in the network. The initially suggested poll interval is too long for high bandwidth and low propagation delay environment. We change the definition of this value, and set the poll interval to about a round-trip time. This means that the transmission of a POLL PDU starts at the time a STAT PDU is received. A timer is required to avoid the loss of POLL PDU. The traffic overhead to transmit the POLL PDU every round-trip time is quite low relative to other flow control mechanisms.

4.4. Control procedure

The QDC is achieved via VQO reduction because it is hard to determine an appropriate queuing delay for various transmission rates and round-trip times. Therefore, the QDC specifies an VQO threshold, α , and the reduction control keeps the VQO close to α . The way to reduce the VQO is through decreasing the transmission rate.

Up on receiving a STAT PDU, the VQO can be calculated. The control process is based on the relationship between the VQO value and α . The algorithm is as follows:

1. The proper transmission rate is read from a STAT PDU.
2. The network service rate is counted as the total number of packets successfully transmitted between the previous control period and the current calculation time.
3. Update the minimum round-trip time and calculate the queuing delay based on the detected round-trip time minus the smallest round-trip time.
4. Calculate the VQO by multiplying the queuing delay with the network service rate.
5. Calculate the control period.
6. Compare VQO with α .

If $VQO > \alpha$

Calculate the Reduction_Interval, $Reduction_Interval = (VQO - \alpha) / Proper_Transmission_Rate$.

If $Reduction_Interval > Control_Period/2$, $Reduction_Interval = Control_Period/2$.

$Transmission_Rate = Proper_Transmission_Rate \times (Control_Period - Reduction_Interval) / Control_Period$.

If $VQO > 1$

$Transmission_Rate = (Control_Period \times Proper_Transmission_Rate + PDU_size) / Control_Period$.

Else

$Transmission_Rate = Proper_Transmission_Rate$.

The way to decide the transmission rate can be considered in three situations, $VQO > \alpha$, $\alpha \geq VQO > 1$ and else. When $VQO > \alpha$, it means that the queue occupancy needs to be reduced, and the reduction rate is calculated as discussed before. When $\alpha \geq VQO > 1$, it represents that the VQO is larger than a PDU size and the link utilization achieves the highest level. This is the situation that the network is able to store a little more data in the intermediate node and improve the fairness. The queue occupancy in the FIFO reflects the bandwidth sharing. The rate increment policy is to increase the queue occupancy by one PDU every control cycle. When all sources have equal queue occupancy in the FIFO, their link bandwidth sharing will be identical. If the queue occupancy is lower than one PDU, it indicates that the link bandwidth is not fully utilized and the rate probing result will be used to allocate more bandwidth.

5. Modification to SSCOP

The modification to SSCOP involves both source and destination functions. A source transmits data packets with the probing traffic pattern, and it operates under variable packet size environment. Besides, a source also performs the QDC function, which is based on the periodic transmission of POLL PDUs. The destination is responsible for estimating the service rate of the bottleneck which is an intermediate node located somewhere along the source–destination path.

5.1. The source behavior

An SSCOP source transmits data packets and POLL PDUs, as well as changes its interval variable according to the derived transmission rate. In our modification, the source transmits data with two different rates, one is twofold, and the other is two-thirds, of the returned rate. The transmission rate is updated upon receiving an USTAT PDU or an STAT PDU. Beside the decision of the normal rate, the QDC is also performed at the source. The rate calculation is performed every control period and the transmission rate of the next control period is determined after the QDC procedure is finished. Since the poll interval is set to the round-trip time, it needs to be updated whenever an STAT PDU is received.

5.2. The destination behavior

The destination performs the calculation of the rate probing result, it derives the proper rate upon receiving a POLL PDU. Besides, when a PDU loss is detected, the destination not only generates USTAT to report the lost PDU but also calculates the PDU receiving rate. The service rate probing consists of two processes. The first process translates the receiving inter-packet intervals into a predefined format that performs the sorting, and the second process computes the proper service rate according to the sorted result whenever a packet loss is detected or a start POLL PDU of a control period is received.

6. System simulation and numerical results

6.1. Simulation model

Assume that the lower layer data transfer is handled by ATM UBR service, which provides a common-FIFO queue in the intermediate switch. We will compare the performance of SSCOP with TCP Vegas

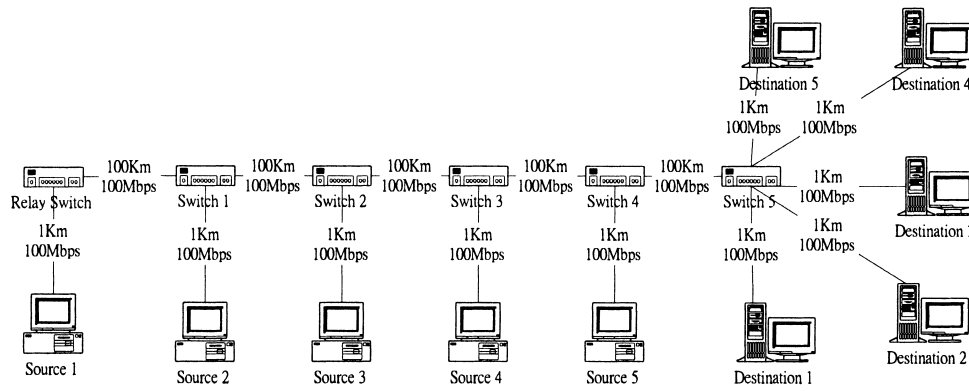


Fig. 1. The simulation model.

[1], which is nowadays the TCP congestion avoidance scheme with the highest throughput. The network model used throughout this work is a parking-lot configuration, which has been used to illustrate the fairness problem. A configuration with five switches is shown in Fig. 1. The capacity of each link is set to 100 Mbps. Various sources experience different round-trip times. This configuration is derived from [6].

The simulation parameters of the modified SSCOP and TCP Vegas are listed in Table 1. Since in the simulation, we want to compare the behavior either when the available bandwidth is reduced or when it is increased, the link bandwidth between SW4 and SW5 is reduced from 100 to 50 Mbps during the time period from 1 to 2 s.

The performance of SSCOP is measured using two well-known performance metrics based on the throughput and the fairness index. Jain’s fairness index [4] is used to evaluate the fairness among sources with different round-trip times, it is described as follows. Given a set of throughputs (x_1, x_2, \dots, x_n) , the function below assigns a fairness index to the set:

$$f(x_1, x_2, \dots, x_n) = \frac{(\sum_{i=1}^n x_i)^2}{n \sum_{i=1}^n x_i^2}.$$

Since the throughput is non-negative, the fairness index always results in values between 0 and 1, the latter value represents that all sources share the bandwidth equally.

Table 1
Simulation parameters

Parameter	Value setting
α	3 PDU
PDU size	512 bytes
TCP Vegas α	1
TCP Vegas β	3
Queuing space	128 PDUs

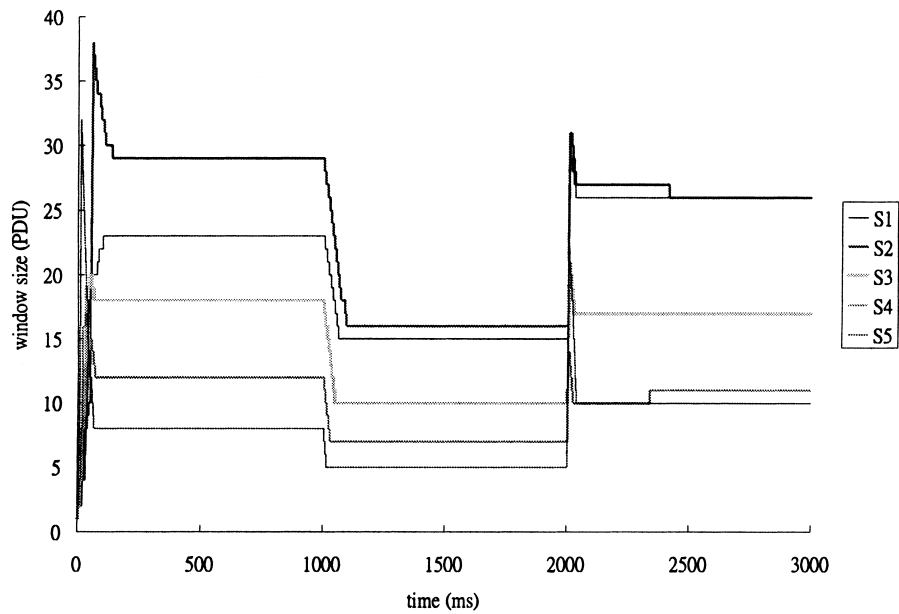


Fig. 2. The window size change in Vegas.

6.2. Numerical results

The simulation condition is used to demonstrate the control function of the proposed SSCOP congestion avoidance scheme. The following figures demonstrate the window size change for TCP Vegas and SSCOP.

Figs. 2 and 3 demonstrate the window size variation during the simulation. The throughput of SSCOP is similar to Vegas. However, if the bandwidth sharing is fair, five sources with different round-trip times should experience different stable window sizes. Vegas does not guarantee to achieve the fairness because it uses two parameters. Instead, our single parameter policy achieves better fairness. The performance comparison between TCP Vegas and SSCOP are shown in the following table.

As Table 2 shows, both TCP Vegas and SSCOP achieve extremely high throughput. However, the fairness of SSCOP is still better than Vegas. Therefore, this improvement allows SSCOP to have very high throughput and nearly optimal fairness.

The queue occupancy at the bottleneck switch is shown in Fig. 4, in which the efficiency of QDC is demonstrated. The queue occupancy can be reduced, but it will never be reduced to zero so that the high throughput can be maintained. As a result, the queue occupancy can be maintained around a specified value and the queuing delay can be guaranteed. Since the queue occupancy stays at a relative low level, a larger free space can absorb bursts from other traffic.

Table 2
Performance comparison

	Throughput (%)	Fairness index
TCP Vegas	99	0.984
SSCOP	99	0.998

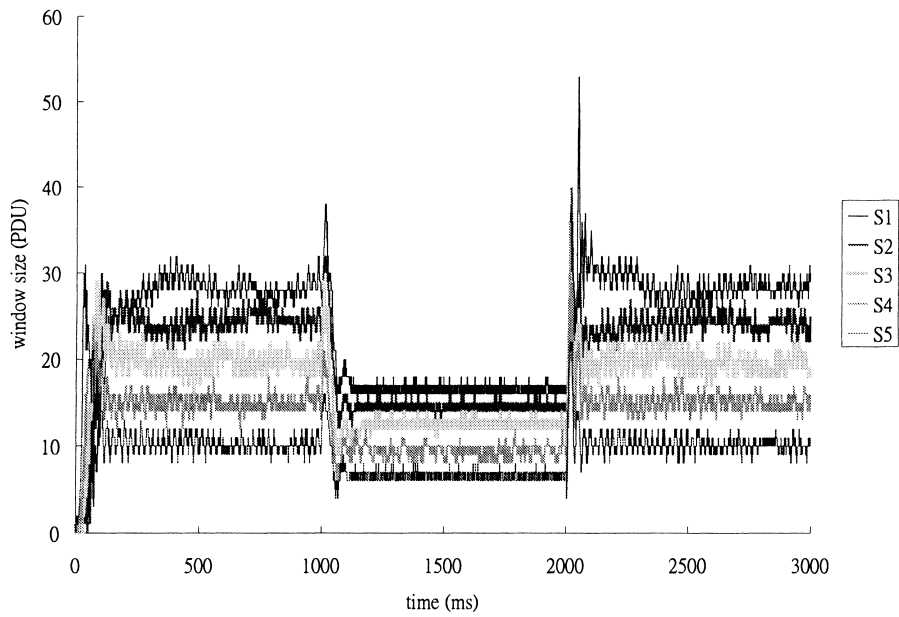


Fig. 3. The window size change in SSCOP.

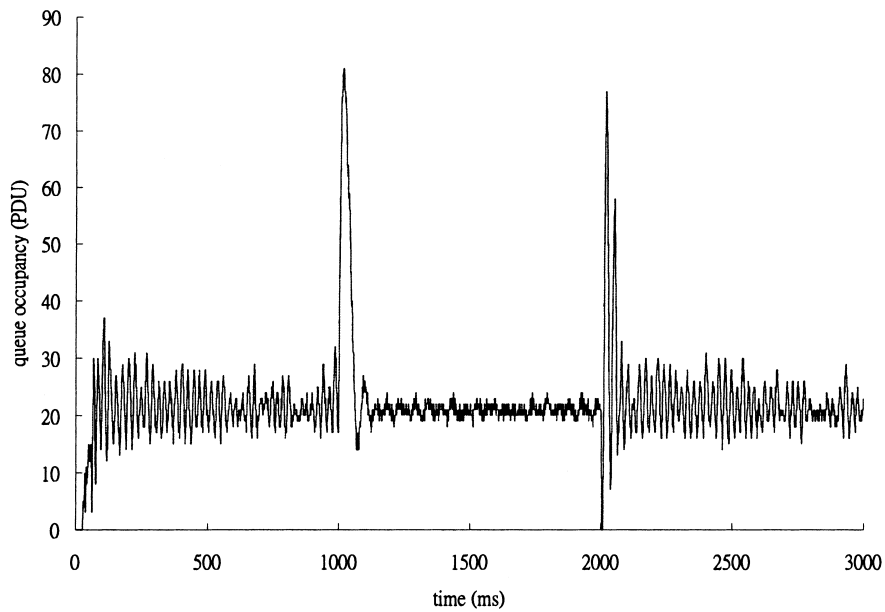


Fig. 4. The queue occupancy at the bottleneck switch (Switch 4).

7. Conclusion

In this work, we developed a flow control scheme based on the characteristic of basic SSCOP flow control. It allows SSCOP to achieve both higher performance and more efficient resource utilization. Through the simulation, the improved SSCOP demonstrates same throughput performance with TCP

Vegas under error free transmission situation, however, it outperforms TCP Vegas in the aspect of fairness. With its efficient selective retransmission mechanism, SSCOP may be a better choice of transport protocol for various networks, especially the ATM network.

References

- [1] L.S. Brakmo, L.L. Peterson, TCP Vegas: end to end congestion avoidance on a global internet, *IEEE J. Select Areas Commun.* 13 (8) (1995) 1465–1480.
- [2] R. Cohen, An improved SSCOP-like scheme for avoiding unnecessary retransmissions and achieving ideal throughput, *INFOCOM 96*, vol. 2, pp. 855–862.
- [3] ITU-T, Final draft rec. Q.2110–B-ISDN signaling ATM adaptation layer overview description, December 1993.
- [4] R. Jain, *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling*, Wiley, New York, 1991.
- [5] S. Keshav, A control-theoretic approach to flow control, *ACM SIGCOMM*, August 1991, pp. 189–201.
- [6] K.K. Romakrishnan, R. Jain, A binary feedback scheme for congestion avoidance in computer networks, *ACM Trans. Comput. System* 8(2) (1990) 158–181.



Jin-Ru Chen received his B.S. degree in Computer Science and Information Engineering from National Chiao-Tung University in 1996. He is currently pursuing the Ph.D. degree in Computer Science, National Chiao-Tung University. He is a student member of IEEE.



Yaw-Chung Chen received his Ph.D. degree in Computer Science from Northwestern University, Evanston, Illinois in 1987. During 1987–1990, he worked at AT & T Bell Laboratories as a member of technical staff. Since 1990, he has been an Associate Professor in the Department of Computer Science and Information Engineering of National Chiao Tung University. His research interests include multimedia communications, broadband internet, and ATM traffic management. He is a member of IEEE and ACM.