Regular Paper

Experimental TCP Performance Study on Emulating Diffserv Assured Forwarding over ATM SBR Service

Shigehiro Ano,† Nicolas Decre*†† and Toru Hasegawa†

Recently, the deployment of Diffserv (Differentiated Services) that enables the QoS (Quality of Service) guarantee is urgently required by IP network customers. However, AF (Assured Forwarding) PHB (Per Hop Behavior) in Diffserv still has not been provided by conventional routers. It is a realistic solution that SBR3 (Statistical Bit Rate 3) of ATM emulates AF PHB, but it is not clear whether TCP traffic over AF PHB emulated by ATM is differentiated from the best effort TCP traffic over DF (Default Forwarding) PHB. To confirm the differentiation, we have experimentally studied TCP performance through the link into which TCP connections over AF PHB and DF PHB are aggregated. This paper describes the experimental results and discusses the possibility of the TCP performance differentiation between AF PHB and DF PHB over wide area ATM networks.

1. Introduction

According to the growing demand for QoS guaranteed services over IP networks, the deployment of Differentiated Service (Diffserv)[1)] is urgently required. In order to provide various types of QoS, two per hop behaviors, i.e., EF (Expedited Forwarding) PHB (Per Hop Behavior)[2)] and AF (Assured Forwarding) PHB[3)], have been specified by IETF. A PHB specifies how routers deal with packets marked as the behavior in order to differentiate the QoS of the marked packets from the other packets. Among the Diffserv PHBs, EF PHB, which may be used to transfer constant rate traffic, has been already supported by commercial routers. Many test bed networks supporting EF PHB such as QBone of Internet2 have been deployed, and performance experiments have been enthusiastically performed.

Since AF PHB is considered useful to transfer data traffic such as TCP traffic, IP networks supporting AF PHB is required to be urgently deployed. However, the early deployment is difficult due to the following reasons. TCP performances over AF PHB were not studied experimentally because AF PHB was not yet supported by commercial routers. Besides, although some studies have evaluated TCP performances over AF PHB using simulation[4),5)], the results show that the standard traffic control methods of router, i.e., RED (Random Early Detection), do not completely differentiate TCP traffic of AF PHB from best effort traffic transferred over DF (Default Forwarding) PHB.

On the other hand, the IP backbone networks may be replaced by the routers based on IP over WDM (Wavelength Division Multiplexing) of which trunk line speed may be peta bit/s order. In such situation, no traffic control may be required. However, in the access/subscriber IP networks, some traffic controls are still required to guarantee the QoS (Quality of Service) of the traffic due to the insufficient physical bandwidth. ATM traffic control can cope with the insufficient bandwidth; therefore ADSL (Asymmetric Digital Subscriber Line) over ATM[6)~8)] and ATM-PON (ATM Passive Optical Network)[9)~11)] are intensively studied recently. Furthermore, a prototype card aiming Diffserv over ATM has also been developed[12)]. However, the drawback of the card is that it requires more complexity in ATM node hardware if the scalability need to be satisfied.

In these circumstances, it is hopeful to deploy AF PHB over ATM SBR3 (Statistical Bit Rate 3)[13)] service for the early deployment of the AF PHB based service. Because the objective of AF PHB and ATM SBR3 is to provide average (sustained) bandwidth guarantee. In this approach, the traffic parameters of AF PHB are mapped to those of ATM SBR3 service with SCD (Selective Cell Discard)[14),15)]. On the contrary, DF PHB is mapped to ATM UBR (Unspecified Bit Rate) service which pro-

† KDDI R&D Laboratories Inc.
†† Ecole Nationale Supérieure des Télécommunications of Paris

* The author was engaged in this work through the internship at KDDI R&D Laboratories Inc.

vides best effort communication in ATM networks. The mapping of AF PHB to ATM SBR3 has been specified by ATM Forum and Internet2 [16),17)].

Despite the usefulness, the performances of AF PHB over ATM service were not yet experimentally studied. Furthermore, it is not clear that the throughput in TCP level may not be guaranteed even if AF PHB provides an assured bandwidth in IP level. Therefore, in order to evaluate the feasibility, we have experimentally evaluated the performances of AF PHB and DF (Default Forwarding) PHB over ATM services [18)]. The various experiments were performed to evaluate the differentiation of PHBs from the viewpoints of TCP throughput and fairness. The objective is to know whether TCP flow control mechanisms can guarantee TCP throughput equal to the bandwidth provided by Diffserv AF PHB. The results show that Diffserv over ATM service is a practical solution for the early deployment of Diffserv services.

The rest of this paper is organized as follows. In Section 2, we describe the brief introduction of Diffserv PHBs and the relationship to ATM traffic control schemes. In Section 3, we show the overview of the experiments and the network parameter tunings. In Section 4, we show the results of the experiments that compares between each PHBs or the assured packet rates of AF PHB. In Section 5, we discuss the obtained results and consider the effectiveness of Diffserv AF PHB over ATM SBR service.

2. Diffserv AF/DF PHB over ATM

2.1 Diffserv

Service differentiation is desired to accommodate heterogeneous application requirements and user expectations, and to permit differentiated pricing of the Internet service. The main concept of Diffserv is as follows.

- IP packets are classified and marked according to their Differentiated Service field (DS field) defined in IP header.
- IP packets are handled in aggregated IP flow at a Diffserv-capable router. The forwarding behavior to the aggregated IP flow is called PHB [19)].

Three PHBs are considered so far: the EF PHB, AF PHB and DF PHB. The detailed mechanisms are described in following.

(1) **EF PHB**

The EF PHB is used to build an end-to-end service that guarantees the assured bandwidth with a low loss, low jitter and low latency. It warrants a minimum throughput (subscribed throughput).

(2) **AF PHB**

The AF PHB is used to build an assured bandwidth end-to-end service without jitter/latency guarantee. In AF PHB group, four classes are defined in terms of allocated network resources such as buffer space and bandwidth. Within each class, IP packets are marked with one of the three drop precedence values.

In the actual service implemented by AF PHB, the traffic contract may consist of assured packet rate and committed burst size. Furthermore, a token bucket based on these traffic contract values may be used as a traffic policer. The drop precedence values are changed by the token bucket depth, d. The depth is measured in bytes. The conditions are, for example, as follows.

i) Low Drop Precedence

$\text{“}\textit{Committed burst size}\text{”} < d$

ii) Medium Drop Precedence

$0 < d \leq \text{“}\textit{Committed burst size}\text{”}$

iii) High Drop Precedence

$d = 0$ (*The bucket is empty.*)

Depending on the drop precedence values, the packets are scheduled to drop or queue in the congestion periods.

(3) **DF PHB**

The DF PHB is used to accommodate the Best Effort (BE) traffic. No traffic control is applied like the current Internet during the congestion period.

2.2 ATM SBR3 and UBR

In SBR3, a traffic contract consists of PCR (Peak Cell Rate) for CLP (Cell Loss Priority) = 0+1 [PCR01], SCR (Sustainable Cell Rate) and MBS (Maximum Burst Size) values for CLP=0 [SCR0, MBS0] [13)]. Using these contract values, traffic policing, also known as Usage Parameter Control (UPC), ensures that the connections with reserved bandwidth are not exceeding their reservations. Those cells that exceed the traffic contract are tagged over SBR with SCD. The tagged cells, which are non conforming cells, are scheduled to drop or queue in the congestion periods. The conformance check whether cells exceed the contract or not is performed using Generic Cell Rate Algorithm (GCRA) that is a cell level token bucket.

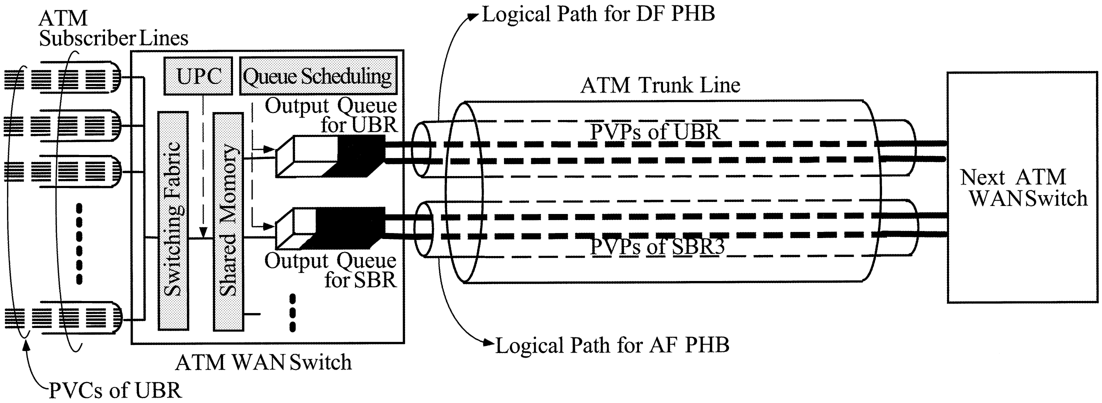

Fig. 1 Mapping strategy of AF/DF PHBs to ATM PVP connections.

UBR is used to accommodate the best effort service that is the same as DF PHB in Diffserv.

2.3 Mapping of AF/DF PHB and ATM Services

(1) AF PHB over SBR3

It is a natural way to map AF PHB to SBR3 with SCD in ATM, as discussed by the ATM Forum [16]. The traffic contracts of AF PHB and ATM SBR3 are mapped in the following way: SCR0 and MBS0 correspond to "Assured Packet Rate" and "Committed Burst Size", respectively. The CLP of SBR3 is mapped to AF drop precedence values. CLP=0 and CLP=1 correspond to medium/low drop precedence value and high drop precedence values, respectively.

The marking and dropping of AF PHB is emulated by ATM switches in the following way: when the queuing traffic exceeds the upper boundary ratio (SCD threshold) of ATM switch buffer size, SCD function starts the discard of cells with CLP=1 in advance to the cell discard by the buffer overflow. In case of no congestion, the exceeded cells are only tagged to CLP=1.

(2) DF PHB over UBR

DF PHB is a best effort forwarding; so, it is mapped to ATM UBR (Unspecified Bit Rate) service.

3. Overview of Experiments

3.1 Testing Methods

It is important to evaluate the differentiation of TCP throughput level to confirm the effectiveness of the approach described in Section 2.3. Especially, it is not clarified whether the assured bandwidth are guaranteed or not during the congestion. Therefore, we focus on the differentiation between AF PHB and DF PHB during the congestion periods. For this purpose, a PVP (Permanent Virtual Path) is applied to aggregated flow of a PHB. **Figure 1** illustrates the mapping strategy. The main features are as follows.

(1) PVP (Permanent Virtual Path) of SBR3 is used to transfer an aggregated traffic forwarded by AF PHB. A PVP of UBR is used to transfer an aggregated traffic of DF PHB. An UPC function is performed at each ATM Wide Area Network (WAN) switch according to the VP level traffic contract.

(2) In ATM WAN switch, SBR3 traffic and UBR traffic use one shared memory and two output queues. UPC is performed just before cells enter the shared memory. SCD is triggered when the shared memory is occupied by the cells at the SCD threshold as described in Section 2.3. Each output queue is used for SBR3 traffic and UBR traffic, respectively. The Queue scheduling function to send out to the trunk line uses the Weighted Fair Queuing (WFQ).

(3) PVCs (Permanent Virtual Channels) at the ATM subscriber lines are used to accommodate TCP connections.

(4) The PVPs are concentrated on a trunk line between ATM WAN switches, and the congestion will occur at the output port to the trunk line where AF PHB is carried out.

(5) Each PVP accommodates multiple PVCs (i.e., TCP/IP connection) that has the same type of traffic contract (i.e., SBR3 or UBR).

(6) The ATM traffic contract values of PVP for AF PHB are determined as follows.

PVP emulating AF PHB (SBR3 with SCD):

$$PCR01 = \text{“}\mathit{Trunk\ line\ speed}\text{”} \quad (1)$$

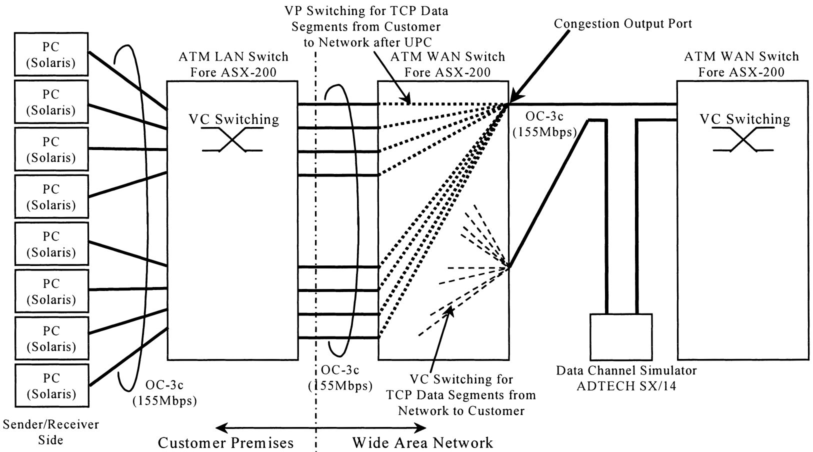

Fig. 2 Experimental configuration.

$$SCR0 = \frac{\text{“}Trunk\ line\ speed\text{”}}{\text{“}Number\ of\ PVPs\text{”}} \tag{2}$$

$$MBS0 \geq \sum_{i=1}^{\gamma} \lambda_i \tag{3}$$

Here,

γ : = *Number of accommodated TCP connections into this PVP connection*
λ_i : = *Number of cells when the data corresponding to each TCP send/receive socket buffer size is consecutively transferred by ATM.*
Equation (3) comes from the experimental results in reference [20].

3.2 Experimental Configuration

Figure 2 shows the configuration of the experiments. Eight PCs (Personal Computers: Pentium III 500 MHz and Solaris 7) with an ATM NIC (Fore PCA-200) are used. These are connected to an ATM LAN switch (Fore ASX-200BX) and an ATM WAN switch (Fore ASX-200BX) via eight OC-3c lines. Each PC establishes TCP connections with different PCs. The number of TCP connections established by one PC is six and each PC pair has two TCP connections with the different TCP port numbers. Therefore, each of eight VPs which correspond to eight OC-3c lines has six TCP connections and the total number of TCP connections for this testing is 48. Each TCP connection is mapped to one VC. At the ATM LAN switch, the VCs with the same destination are switched into the same output OC-3c line. It should be noted that cell loss due to the buffer overflow does not occur at ATM LAN switch. At the ATM WAN switch, eight input lines are multiplexed into one OC-3c output line handled as the ATM WAN trunk line. The output buffer size for each PHB in the ATM WAN switch is set to 10,000 cells. In order to emulate the propagation delay we use the data channel simulator (ADTECH SX/14). The VPs are maintained between the output ports of the ATM LAN switches and the input port of the second ATM WAN switch. VP level UPCs are performed at the first ATM WAN switch. We need to say that the second ATM WAN switch is introduced just because of the limitation of number of VPs supported by Fore ASX-200BX.

A free software module, ttcp, for TCP throughput measurement is used in the TCP/IP communication between PCs. It can calculate TCP throughput in the case that a greedy transmitter like ftp is used, by varying the values of various TCP parameters, such as TCP window size and the user data size. The TCP window size is set to 48 kbyte. The user data size and MSS (Maximum Segment Size) is fixed to 8,192 byte. Based on the principle described in section 2.1, ATM traffic contract values are set as follows.

$$PCR01 = 149.76\,\text{Mbit/s}$$

$$SCR0 = 18.72\,\text{Mbit/s}$$

$$MBS0 = 8256\,\text{cell}$$

SCD threshold is fixed to 90% through the testing. During the ttcp execution, we also measure the packet queuing delay using 2,048 byte ICMP (Internet Control Message

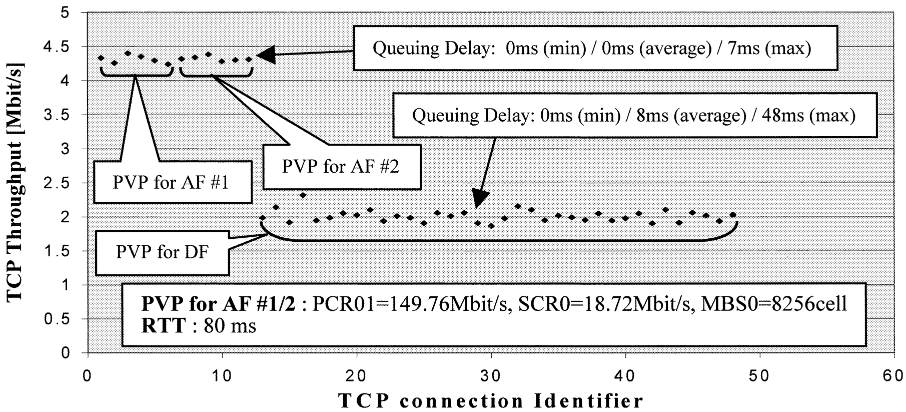

Fig. 3 Each TCP throughput under PVPs for AF and DF PHBs with RTT = 80 ms.

Protocol) packet by ping command over the route of TCP connections.

As for PCR shaping at the PCs, 35 Mbit/s including the cell header is adopted. The duration of each TCP throughput measurement is fixed to 180 seconds. In this experimental configuration, the RTT (Round Trip Time) value is almost determined by the propagation delay when the network congestion does not occur. Therefore, we insert the same propagation delay to the both directions of the trunk line by the data channel simulator. The propagation delay is set to the half of the RTT value in the no congestion case. In this paper, RTT value in the no congestion case is set to 20 ms, 80 ms or 160 ms. From the following sections, we simply use RTT as the RTT in the no congestion case.

4. Results of TCP Performance Measurement

4.1 Differentiation between AF PHB and DF PHB

Under the configuration of Fig. 2, we measured the throughput of each TCP connection under SBR3 with SCD for AF PHB and UBR for DF PHB. Two PVPs are devoted to AF PHB and other PVPs to DF PHB.

(1) RTT = 20 ms/80 ms

When RTT is 20 ms or 80 ms, the total estimated TCP throughput when the physical line speed is unlimited is much larger than the physical bandwidth of the ATM trunk line. This means the heavy congestion occurs at the output port of the sender ATM switch. **Figure 3** shows each TCP throughput in the case of RTT = 80 ms. TCP connections whose identifier are from #1 to #12 use PVPs for AF PHB and TCP connections whose identifier are from #13 to #48 use PVPs for DF PHB. The case of RTT = 20 ms shows the similar result.

As shown in the figure, the AF and DF are differentiated from the viewpoint of TCP throughput. To analyze the throughput values quantitatively, estimated TCP throughput values for AF PHB and DF PHB are calculated based on the following assumption.

Assumptions:

- The residual bandwidth which exceeds the subscribed traffic rate for PVPs for AF PHB is equally shared by the all PVPs for AF and DF PHBs.
- According to TCP behavior and ttcp software, the size of all transmitted TCP segments is MSS (=8,192 byte).
- An 8,192 byte TCP segment is transferred by 172 ATM cells. It is because the TCP segment is encapsulated by TCP/IP header (=40 byte), LLC/SNAP header (=8 byte) and AAL5 (ATM Adaptation Layer type 5) padding field (=8 byte) and AAL5 trailer (=8 byte) based on the standardized method [21]. The size of AAL5 padding field is determined by the result of the alignment to ATM user information size (=48 byte). When the TCP segment size is 8,192 byte (=MSS), the size of the padding field is 8 byte and the segment is divided into 172 cells. We also confirm the number of cells in one segment with MSS by the ATM tester. **Figure 4** illustrates the way of the mapping.
- Total bandwidth in ATM level is 149.76 Mbit/s and traffic contract value of AF in ATM level is 18.72 Mbit/s as described in

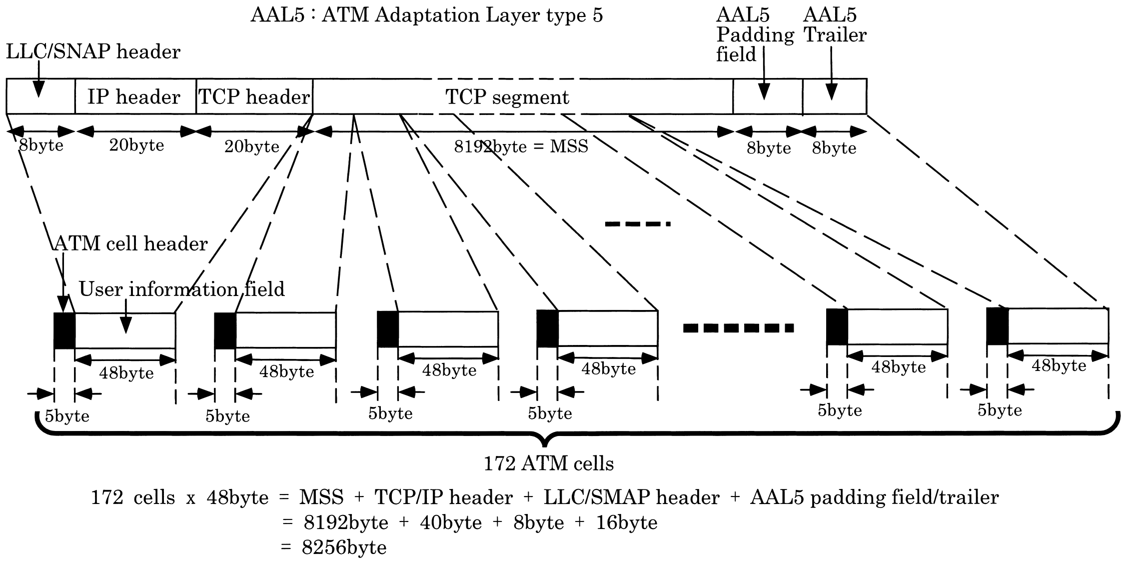

Fig. 4 Mapping of TCP segment with MSS to ATM cells.

Section 3.2.

Calculation:

(DF PHB)

"Estimated PVP bandwidth"

$$= \frac{149.76\,\text{Mbit/s} - 18.72\,\text{Mbit/s} \times 2}{8\ PVP\ connections}$$

$$= 14.04\,\text{Mbit/s}$$

"Estimated TCP throughput"

$$= \frac{14.04\,\text{Mbit/s} \times \dfrac{8,192\,\text{byte}}{172\,\text{cell} \times 53\,\text{byte}}}{6\ TCP\ connections}$$

$$= 2.10\,\text{Mbit/s}$$

(AF PHB)

"Estimated PVP bandwidth"

$$= 14.04\,\text{Mbit/s} + 18.72\,\text{Mbit/s}$$

$$= 32.76\,\text{Mbit/s}$$

"Estimated TCP throughput"

$$= \frac{32.76\,\text{Mbit/s} \times \dfrac{8,192\,\text{byte}}{172\,\text{cell} \times 53\,\text{byte}}}{6\ TCP\ connections}$$

$$= 4.91\,\text{Mbit/s}$$

The average TCP throughput values of Fig. 3 are almost close to the estimated values. The differences may be due to cell loss and the consequence of TCP retransmission bandwidth.

On the other hand, the average queuing delay of PVP for AF PHB is 0 ms and the maximum queuing delay is 7 ms. These results are much smaller than 8 ms average delay and 48 ms maximum delay in PVPs for DF PHB.

(2) RTT = 160 ms

In the case of RTT = 160 ms, the total traffic of all TCP connections is not always enough large to fill the trunk line with TCP data. In other words, the link is not always utilized due to the small window sizes of TCP. In this case, the estimated TCP aggregated throughput is calculated by the following equation independent of the kind of PHBs.

"Estimated TCP aggregated throughput"

$$= \frac{\mu}{RTT + \delta} \times \alpha \qquad (4)$$

Here,

μ := "TCP send/receive socket buffer size set in PVP for AF#2"

δ := "Data sending out time to ATM line"

α := "Number of the accommodated TCP connections"

The measured and calculated values are as follows.

PVPs for AF : 2.30 Mbit/s (measured)
: 2.41 Mbit/s (calculated)

PVPs for DF : 2.28 Mbit/s (measured)
: 2.41 Mbit/s (calculated)

Figure 5 shows each measured TCP throughput in the case of RTT = 160 ms. From these results, the measured values are almost the

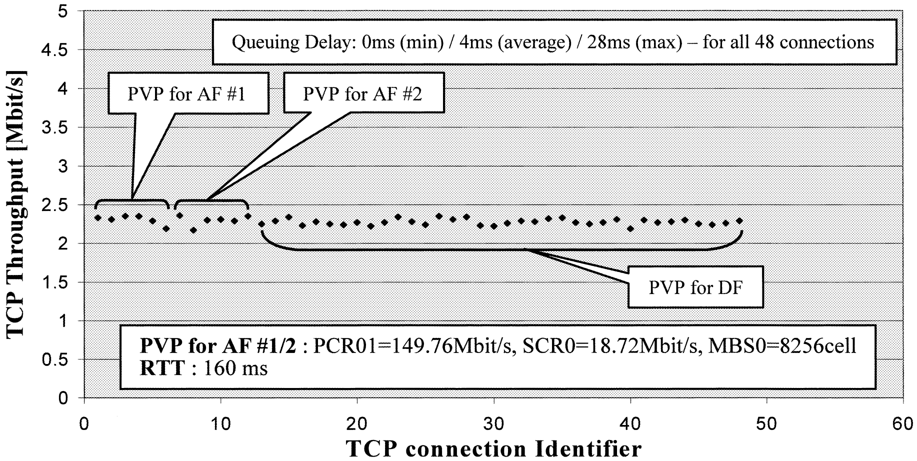

Fig. 5 Each TCP throughput under PVPs for AF and DF PHBs with RTT = 160 ms.

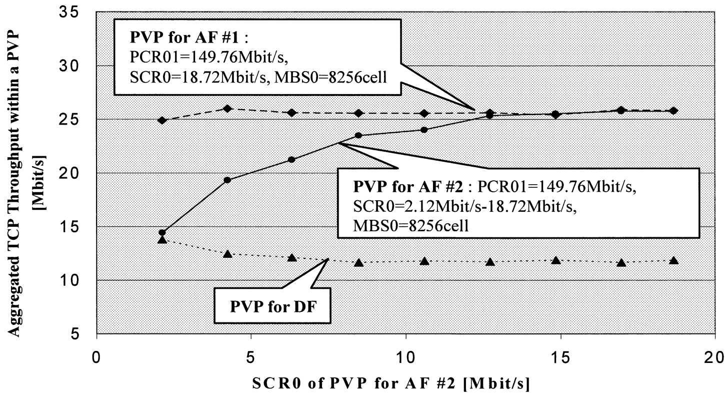

Fig. 6 Effect of SCR0 with RTT = 80 ms.

same as the calculated values. This means that the Diffserv PHBs do not affect the TCP throughput if the total estimated TCP throughput without the limitation of the physical line speed is lower than the physical bandwidth of the trunk line. In other words, this trunk line is not congested.

4.2 Different Assured Packet Rates of AF PHB

The experiments changing SCR0 values were performed to study the differentiation between the traffic contracts of different assured packet rates. Besides, the experiments changing MBS0 values were performed to study the differentiation between the traffic contracts of different committed burst sizes. The following PVPs are adopted for testing.

- The traffic contract values of one PVP for AF PHB is the same as those in Section 3.2. We call this PVP AF #1.
- The traffic contract values of the other PVP for AF PHB is the same as PVP AF #1 except for SCR0 or MBS0. Either SCR0 or MBS0 is changed during the experiments. We call this PVP AF #2.
- The other six PVPs are used for DF PHB. UBR is applied to those PVPs.

(1) Effect of SCR0

Figure 6 shows the results of changing SCR0 when RTT is set to 80 ms. To evaluate the bandwidth allocation for AF #2, we estimated the TCP aggregated throughput based on the assumptions in Section 4.1. **Table 1** 1 shows the estimated values along with the measured values in the case of RTT = 80 ms. The measured values are almost the same as the estimated values even if SCR0 value is changed. This means that the TCP aggregated throughput of AF #2 is allocated the fair share of the residual bandwidth. The similar results are

Table 1 TCP aggregated throughput in PVP for AF #2 changing SCR0 values.

RTT	SCR0	18.72 Mbit/s	14.84 Mbit/s	10.60 Mbit/s	6.32 Mbit/s	2.12 Mbit/s
80 ms	Estimated	29.44 Mbit/s	26.39 Mbit/s	23.05 Mbit/s	19.69 Mbit/s	16.39 Mbit/s
	Measured	25.72 Mbit/s	25.52 Mbit/s	24.01 Mbit/s	21.23 Mbit/s	14.33 Mbit/s

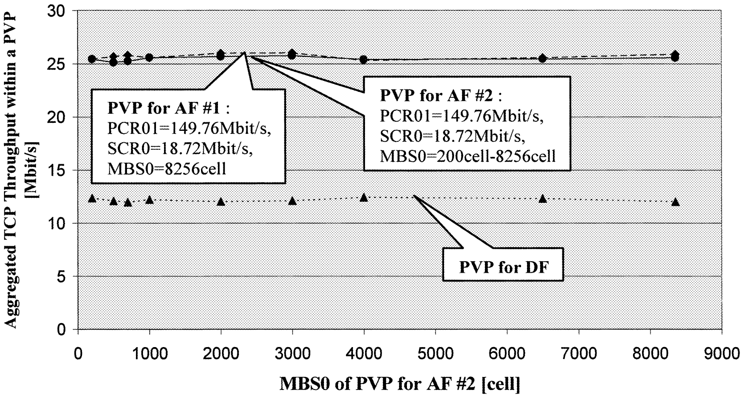

Fig. 7 Effect of MBS0 with RTT = 80 ms.

observed in the case of RTT = 20 ms. However, the measured throughput becomes a little larger than the estimated throughput when SCR0 becomes smaller value (≤10 Mbit/s). In this condition, the standard deviation of TCP throughput becomes large.

(2) Effect of MBS0

Figure 7 shows the results of changing MBS0 when RTT is set to 80 ms. As shown in the figure, the TCP aggregated throughput does not change independent of MBS0. The same results are obtained when the other RTT values (= 20 ms and 160 ms) are set. The differentiation to committed burst size for the TCP aggregated traffic is not achieved because the traffic becomes constant due to the statistical multiplexing.

5. Discussions

The following results are made clear from the above experiments.

(1) The TCP throughput differentiation between AF PHB and DF PHB over ATM networks is successfully realized during the congestion. ATM SBR3 with SCD can assign the nearly ideal throughput to each TCP connection belonging to PVPs for AF PHB. In other words, the assured bandwidth allocated to AF PHB is shared equally by TCP connections over AF PHB, and the residual bandwidth is shared equally by all TCP connections over AF PHB and DF PHB. On the other hand, the average queuing delay of PVP for AF PHB is smaller than that of DF PHB.

(2) Even when AF PHBs of different assured packet rates are mixed, each AF PHB over ATM networks can assign nearly ideal throughput according to the assured packet rates. This is achieved by using independent PVPs for the AF PHBs with the different assured packet rates.

(3) The TCP throughput fairness in AF PHB is better than that in DF PHB. In other words, the variance of TCP throughputs in AF PHB is much less than that in DF PHB. In order to clarify the reason, we analyze TCP communication sequences using the *TCP Analyzer* [22]. The TCP analyzer, which was developed by KDDI R&D Labs., emulates TCP protocol behaviors using TCP segments captured from a communication link, and it can estimate internal procedures of TCP flow control and internal values such as a congestion window (*cnwd*).

In the case of DF PHB, timeout retransmissions frequently occur as shown in **Fig. 8**, which is the result of the TCP analyzer. In the figure, the dotted line represents the TCP sequence numbers. TCP sequence numbers do not increase frequently for longer duration than one second (e.g., 18,000 ms, 40,500 ms, 58,500 ms, 76,500 ms and 81,000 ms in Fig. 8). At these times, a TCP sender detects lost TCP segments

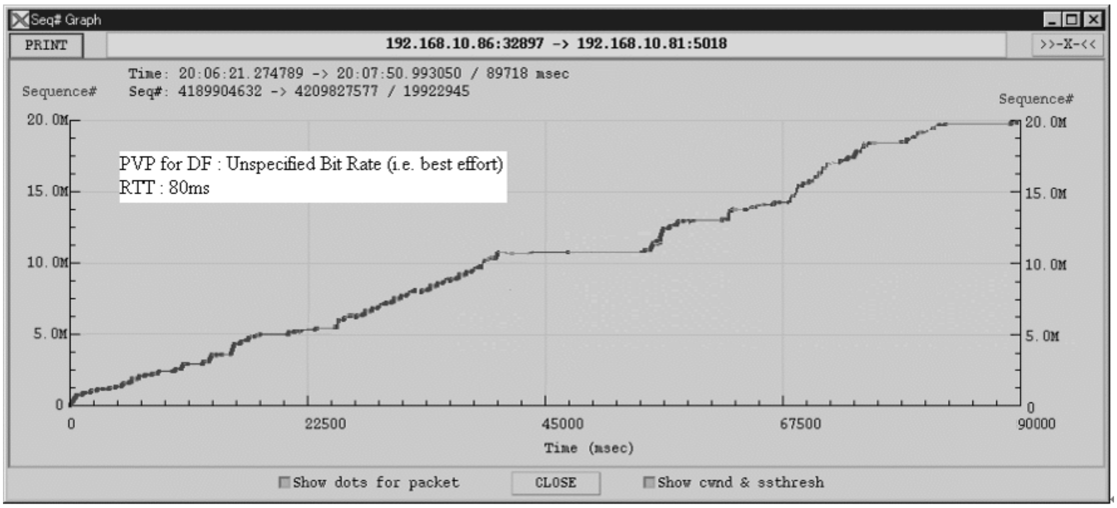

Fig. 8 TCP sequence number in PVP for DF.

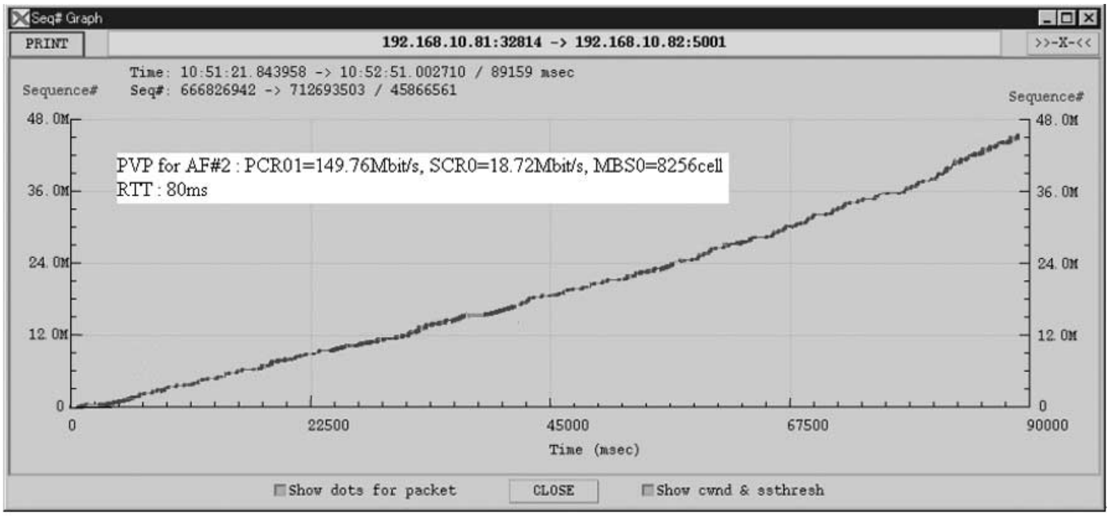

Fig. 9 TCP sequence number in PVP for AF.

using a retransmission timer. In addition, *Exponential backoff*[23] is sometimes triggered at a TCP sender due to a loss of a retransmitted segment (e.g., 40,500 ms in Fig. 8). The retransmission timer value is doubled every retransmitted segment loss, and in Fig. 8, the exponential backoff continues more than 10 seconds.

On the contrary, in the case of AF PHB, most retransmissions of TCP segments are invoked by fast recovery and fast retransmission as shown in **Fig. 9**. These retransmissions are invoked when three duplicate ACKs (acknowledgments) are received by a TCP sender. The duration of no data transfer range from 200 ms to 500 ms.

As a result, the fairness degradation for DF PHB is considered due to the frequency of timeout retransmission and exponential backoff because these generally make TCP throughput unstable.

The difference of timeout retransmission frequency depends on how ATM cells are dropped by ATM switches. The UPC mechanism for SBR3 equally discards the cells from each TCP connection of the PVP because it discards nonconforming cells in advance of the buffer overflow of the ATM switch. In addition, since cells are not consecutively discarded, consecutive TCP segment losses which cause a retransmission timeout do not occur so often. On

the other hand, cells of UBR is not controlled by ATM switches, and those are discarded together at the time of buffer overflow. This causes the devastating loss of cells (i.e., TCP segments) which is the origin of a timeout retransmission.
(4) The size of MBS, which corresponds to committed burst size of AF PHB, does not affect TCP throughput over ATM SBR3. In other words, even when a smaller MBS is used, TCP throughput does not change. This is because the aggregated TCP traffic is not so bursty as the traffic of a single TCP connection [24]. At the experiments of Section 4, six TCP connections are aggregated to a single PVP for AF PHB; therefore; the traffic is not considered bursty.
(5) The traffic control methods of ATM switches are effective to the differentiation between AF PHB and DF PHB. The ATM switch used at the experiments have independent queues (output buffers) for SBR and UBR services, and they contribute to the differentiation between AF PHB and DF PHB. However, ATM SBR3 can provide only two drop precedence values since two levels of CLP are provided by ATM switches.

In spite of the only two levels, the support of independent queues let this approach become one of the most realistic candidates for the initial deployment of Diffserv AF/DF PHB services in the wide area IP networks.

Consequently, ATM SBR3 with SCD is a practical way to realize the Diffserv Assured Forwarding using the commercial products.

6. Conclusions

The emulation of Diffserv AF PHB over ATM SBR3 service is considered a realistic solution to the early deployment of IP networks supporting AF PHB. In the solution, AF PHB and DF PHB are mapped to SBR3 and UBR services of ATM networks. We have experimentally evaluated the differentiation of AF PHB and DF PHB in TCP level in order to assess the feasibility. The following results are obtained.
(1) AF PHB and DF PHB over the ATM networks realize TCP throughput differentiation during the link congestion. ATM SBR3 with SCD can assign each TCP connection the estimated TCP throughput supposing the fair share of the residual bandwidth. On the other hand, the average queuing delay of PVP for AF PHB is largely smaller than that for DF PHB.
(2) The differentiation to assured packet rates is achieved in TCP level. In other words, AF PHBs of different assured packet rates equally share the residual bandwidth which is not allocated to AF PHB.
(3) The TCP throughput fairness to each TCP connection is achieved in AF PHB. In TCP connections of AF PHB, timeout retransmission of TCP segments do not occur frequently. This makes TCP throughput in AF PHB stable.

The results have proved the effectiveness of the emulation of Diffserv AF PHB over ATM SBR3 service. The emulation is expected to be used for the early deployment of IP networks supporting AF PHB.

Acknowledgments The authors would like to express their gratitude to Mr. T. Asami, President & CEO of KDDI R&D Laboratories Inc., and Dr. Y. Matsushima, Executive Vice President, for their continuous support of this work.

References

1) Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z. and Weiss, W.: *An Architecture for Differentiated Services*, RFC2475, The Internet Engineering Task Force (1998).
2) Jacobson, V., Nichols, K. and Poduri, K.: An Expedited Forwarding PHB, RFC2598, The Internet Engineering Task Force (1999).
3) Heinanen, J., Baker, F., Weiss, W. and Wroclawski, J.: *Assured Forwarding PHB Group*, RFC2597, The Internet Engineering Task Force (1999).
4) Clark, D. and Fang, W.: Explicit Allocation of Best Effort Packet Delivery Service, *IEEE/ACM Trans. Networking*, Vol.1, No.4, pp.397–413 (1998).
5) Ibanez, J. and Nichols, K.: *Preliminary Simulation Evaluation of an Assured Service*, Internet Draft ⟨draft-ibanez-diffserv-assured-eval-00.txt⟩, The Internet Engineering Task Force (1998).
6) Kawahara, R. and Saito, H.: Performance of TCP/IP over ATM over an ADSL, *IEICE Trans. Comm.*, Vol.E83B, No.2, pp.140–154 (2000).
7) Universal ADSL Working Group (UAWG), http://www.uawg.org/.
8) The ADSL Forum, http://www.adsl.com/.
9) Takigawa, T., Aoyagi, S. and Maekawa, E.: ATM Based Passive Double Star System Offering B-ISDN, N-ISDN, and POTS, *GLOBECOM'93*, Vol.1, pp.14–18 (1993).
10) Bessho, Y., Kozaki, S., Mukai, H. and Ichibangase, H.: Performance Evaluation of

ATM-PON Interface for OLT, *The 2000 Society Conference of IEICE*, Vol.2, B-8-14, pp.196 (2000).
11) The International Telecommunication Union — Telecommunication Standardization Sector (ITU-T), *Broadband Optical Access Systems Based on Passive Optical Networks (PON)*, Recommendation G.983.1, Geneva (1998).
12) Ishihara, T., Tanaka, J., Goto, M. and Oda, S.: Diffserv-Based QoS over ATM Access Networks, *IEICE Trans. Comm.*, Vol.E84B, No.6, pp.1498–1503 (2001).
13) The International Telecommunication Union — Telecommunication Standardization Sector (ITU-T), *Traffic Control and Congestion Control in B-ISDN*, Recommendation I.371, Geneva (1996).
14) Ano, S., Hasegawa, T. and Kato, T.: A Study on Accommodation of TCP/IP Best Effort Traffic to Wide Area ATM Network with VBR Service Category Using Selective Cell Discard, *IEEE ATM'99 Workshop*, pp.535–540 (1999).
15) Ano, S., Hasegawa, T. and Kato, T.: An Experimental Study on Performance during Congestion for TCP/IP Traffic over Wide Area ATM Network Using VBR with Selective Cell Discard, *IEICE Trans. Comm.*, Vol.E83-B, No.2, pp.155–164 (2000).
16) Aboul-Magd, O., et al.: Mapping of Diff-Serv to ATM Categories, 99-0093, *The ATM Forum* (1999).
17) Geib, R.: *Differential Services for the Internet and ATM*, ⟨http://www.internet2.edu/qos/wg/papers/I2QoS-geib-difs-atm-02.html⟩, The Internet2 QoS Working Group (1999).
18) Ano, S. Decre, N. and Hasegawa, T.: Experimental TCP Performance Evaluation on Diffserv Assured Forwarding over ATM SBR services, *IEEE International Conference on Networking (ICN'01)*, pp.825-835 [Part II] (2001).
19) Nichols, K., Blake, S., Baker, F. and Black, D.: *Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers*, RFC2474, The Internet Engineering Task Force (1998).
20) Ano, S., Hasegawa, T., Kato, T., Narita, K. and Hokamura, K.: Performance Evaluation of TCP Traffic over Wide Area ATM Network, *IEEE ATM'97 Workshop*, pp.73–82 (1997).
21) Heinanen, J.: *Multiprotocol Encapsulation over ATM Adaptation Layer 5*, RFC1483, The Internet Engineering Task Force (1993).
22) Kato, T., Ogishi, T., Idoue, A. and Suzuki, K.: Intelligent Protocol Analyzer with TCP Behavior Emulation for Interoperability Testing of TCP/IP Protocols, *Formal Description Techniques for Distributed Systems and Communication Protocols (FORTE X) and Protocol Specification*, Testing and Verification (PSTV XVII), pp.449–464 (1997).
23) Wright, G.A. and Stevens, W.R.: *TCP/IP Illustrated — Volume 2: The Implementation*, Addison-Wesley, ISBN 0-201-63354-X (1995).
24) Ano, S., Hasegawa, T., Kato, T., Narita, K. and Hokamura, K.: Performance of TCP Traffic over VBR service in ATM Network, *IEICE Trans.*, Vol.J80-B-I, No.6, pp.375–386 (1997).

(Received June 11, 2001)
(Accepted December 18, 2001)

Shigehiro Ano received the B.E. and the M.E. degrees in electronics and communication engineering from Waseda University, Japan in 1987 and 1989, respectively. Since joining KDD in 1989, he has been working in the field of ATM switching system and ATM networking. His current interests are QoS routing and multicast protocols realized over the next generation Internet. He is also interested in the field of the today's Internet performance monitoring. He is currently a Senior Research Engineer of Network Management System Lab. in KDDI R&D Laboratories Inc. He received IPSJ Convention Award in 1995.

Nicolas Decre was born in 1978 at Nantes, France. He devoted the preparation for the competitive entrance examinations to French Engineering Schools (Lycée Hoche, Versailles) from 1996 to 1998. In 2001, he graduated from the ENST (Ecole Nationale Supérieure des Télécommunications) of Paris, three year program leading to the State Engineering Degree. Now he is working for Net2S as a consultant in networks. His measures are computer science, network protocols and network management. He has engaged in this work through the six month internship at KDDI R&D Laboratories Inc. from July 2000 to December 2000.

Toru Hasegawa received the B.E. and the M.E. degrees in information engineering from Kyoto University, Japan, in 1982 and 1984, respectively. He received Dr. Informatics Degree from the graduate school of informatics, Kyoto University in 2000. Since joining KDD in 1984, he has been working in the field of formal description technique (FDT) of communication protocols. From 1990 to 1991, he was a visiting researcher at Columbia University. His current interests are high speed protocol and ATM. He is currently the Senior Manager of Mobile IP Network Lab. in KDDI R&D Laboratories Inc. Since 2002, he has also been a Guest Professor of National Institute of Informatics. He received IPSJ Convention Award in 1987.