マルチメディア通信と分散処理　102－12
（2001．3．21）

Diffserv AF PHBのATMによるエミュレーション時のTCP性能評価

阿野 茂浩　　長谷川 亨
KDD研究所

近年、IP ネットワークにおいて、QoS (Quality of Service)の提供を可能とする、Diffserv (Differentiated Services)の構築が急務であるが、Diffserv の AF (Assured Forwarding) PHB (Per-Hop Behavior)は、現状のルータ機能としては提供されていない。これに対して、AF PHB を ATM の SBR3 (Statistical Bit Rate 3)を用いてエミュレーションする手法は現実的な解の 1 つである。しかし、独自のフロー制御を行う TCP が、SBR3 による AF PHB 上で齟齬なく動作し、ベストエフォートによる TCP トラヒックと性能面で差別化できるかについて明らかではない。そこで、筆者らは、AF PHB をエミュレーションした ATM SBR3 上に収容された TCP トラヒックを、ベストエフォートによる TCP トラヒックと中継回線上で混在させ、スループットについて比較実験を行った。本稿では、その実験結果と差別化の実現性について論じる。

Experimental TCP Performance Evaluation on Diffserv Assured Forwarding PHB Emulation by ATM

Shigehiro Ano　Toru Hasegawa
KDD R&D Laboratories Inc.

In these days, the deployment of Diffserv (Differentiated Services) that enables the QoS guarantee is urgently required by the 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 is aggregated. This paper describes the experimental results and discusses the possibility of the TCP performance differentiation between AF PHB and DF PHB over ATM.

1. Introduction

According to the growing demand for QoS guaranteed services over IP networks, the deployment of Differentiated Service (DiffServ)[1] is urgently required. Among the Diffserv PHBs (Per Hop Behaviors), EF (Expedited Forwarding) PHB has been already supported by commercial routers. However, AF (Assured Forwarding) PHB has been evaluated only by simulation studies[2]. Since AF PHB was not supported by commercial routers, it is an alternative solution to deploy AF PHB over ATM SBR3 (Statistical Bit Rate 3)[3] service with SCD (Selective Cell Discard). In order to evaluate the feasibility, we have experimentally evaluated the performances of AF PHB and DF (Default Forwarding) PHB over ATM services. 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.

2 Diffserv AF PHB over ATM

2.1 Diffserv AF PHB

AF PHB is used to build an assured bandwidth end-to-end service without jitter/latency guarantee. 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. The drop precedence values are changed by the policer that watches whether the traffic is conforming to the subscribed rate or not. Depending on the drop precedence values, the packets are scheduled to drop or queue in the congestion periods.

The traffic contract of service implemented by AF PHB consists of assured packet rate and maximum burst size.

2.2 Mapping of PHBs and ATM Services

(1)AF PHB

It is a natural way to map AF PHB to SBR3 with SCD in ATM, as discussed by the ATM Forum[4]. In SBR3, a traffic contract consists of PCR for CLP (Cell Loss Priority) = 0+1 [PCR01], SCR (Sustainable Cell Rate) and MBS (Maximum Burst Size) values for CLP=0 [SCR0, MBS0].

The traffic contracts of AF PHB and ATM SBR3 are mapped in the following way: SCR0 and MBS0 correspond to "Assured Packet Rate" and "Maximum Burst Size", respectively. The CLP of SBR3 is mapped to AF drop precedence values. CLP=0 and CLP=1 correspond to high drop precedence value and medium / low 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, violation cells are only tagged to CLP=1.

(2) DF PHB

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

We performed the experiments according to the following principles.

1. To confirm the differentiation of TCP throughput level, we compare between AF PHB and DF PHB during the congestion periods.
2. A PVP (Permanent Virtual Path) of SBR3 is used to transfer an aggregate of the traffic forwarded by AF PHB. A PVP of UBR is used to transfer an aggregate of DF PHB.
3. The above 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.
4. Each PVP accommodates multiple TCP/IP connection.
5. The ATM traffic contract values of PVP for AF PHB are determined as follows:

PVP emulating AF PHB (SBR3 with SCD) :

PCR01 = *"Trunk line speed"*
SCR0 = *"Trunk line speed"* / *"Number of PVPs"*
MBS0 = *"The sum of TCP window sizes of the aggregated TCP connections"*[5]

3.2 Experimental configuration

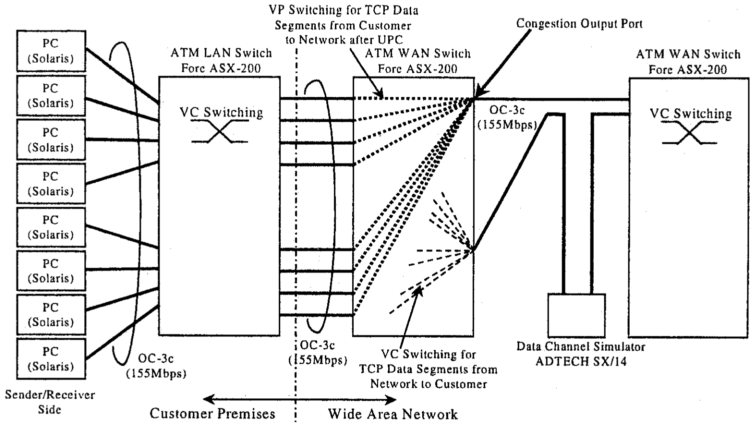

Figure 1 Experimental Configuration.

Figure 1 shows the configuration of the experiments. Eight PCs (Personal Computers: Pentium III 500MHz 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 OC3-c 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 OC3-c 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 OC3-c 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 OC3-c output line handled as the ATM WAN trunk line. The output buffer size for each PHB in the ATM WAN switch is set to 10000 cells. The data channel simulator (ADTECH SX/14) for insertion of propagation delay is connected. 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 8192 bytes. Based on the principle described in section 2.1, ATM traffic contract values are set as follows:

PCR01 = 149.76Mbit/s
SCR0 = 18.72Mbit/s
MBS0 = 8256cell

SCD threshold is fixed to 90% through the testing. During the *ttcp* execution, we also measure the packet queuing delay using 2048byte ICMP (Internet Control Message 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 and the RTT (Round Trip Time) value is set to 20ms, 80ms or 160ms.

4. Results of TCP Performance Measurement

4.1 Differentiation between AF PHB and DF PHB

Under the configuration of Fig.1, 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. Figure 2 shows each TCP throughput in the case of RTT = 80ms. 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.

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:

Assumption :

(a) The residual bandwidth with the exception of the subscribed traffic rate for PVPs for AF PHB is equally shared by the all PVPs for AF and DF PHBs.

(b) According to TCP behavior and *ttcp* software, the size of all transmitted TCP segments is MSS (=8192byte).

(c) A 8192byte TCP segment is transferred by 172 ATM cells. It is because the TCP segment is appended to TCP/IP header (=40byte), LLC/SNAP header (=8byte) and AAL type 5 trailer and padding field.

(d) Total bandwidth in ATM level is 149.76Mbit/s and traffic contract value of AF in ATM level is 18.72Mbit/s as described in section 3.2.

Calculation :

(DF PHB)

"*Estimated PVP bandwidth*" = (149.76Mbit/s − 18.72Mbit/s * 2) / 8 = 14.04Mbit/s

"*Estimated TCP throughput*" = 14.04Mbit/s / (172cell * 53byte) * 8192byte / 6connection = 2.10Mbit/s

(AF PHB)

"*Estimated PVP bandwidth*" = 14.04Mbit/s + 18.72Mbit/s = 32.76 Mbit/s

"*Estimated TCP throughput*" = 32.76Mbit/s / (172cell * 53byte) * 8192byte / 6connection = 4.91Mbit/s

The average TCP throughput values of Fig.2 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 0ms and the maximum queuing delay is 7ms. These results are largely smaller than 8ms average delay and 48ms maximum delay in PVPs for DF PHB.

In the case of 20ms RTT (heavy congestion), the similar results of throughput and queuing delay could be observed, but, in the case of 160ms RTT (no congestion), the no differentiation could be observed.

4.2 Different Assured Packet Rates of AF PHB

The experiments changing SCR and MBS values were performed to study the differentiation between the traffic contracts of different assured packet rates. The following PVPs are adopted for testing:

(a) 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.

(b) 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.

(c) The other six PVPs are used for DF PHB. UBR is applied to those PVPs.

4.2.1 Effect of SCR0

Figure 3 and Figure 4 show the results for change of SCR0 when RTT values are set to 80ms and 160ms, respectively. As shown in Fig.3, if the SCR0 is more than 8.5Mbit/s, the TCP aggregate of AF #2 gets more bandwidth than SCR0 plus the fair share of the residual bandwidth. However, if the SCR0 is less than equal to 8.5Mbit/s, the TCP aggregate of AF #2 gets the bandwidth equal to the fair share. When the SCR0 is 8.5Mbit/s, the estimated TCP aggregate throughput is 23.82Mbit/s which is calculated in the same way as section 4.1. The actual TCP aggregate bandwidth is 23.48Mbps, and is almost the same the estimated bandwidth. It is considered that as the difference between SCR0 values of PVPs becomes larger, the bandwidth of PVP whose SCR0 is smaller gradually becomes the estimated bandwidth.

In the case of RTT 160ms, the total bandwidth delay product of TCP connections is not always enough large to fulfill TCP data in the network. In other words, the link is not always utilized due to the small window sizes of TCP. In this case, the bandwidth of PVP whose SCR0 is smaller cannot get the residual bandwidth.

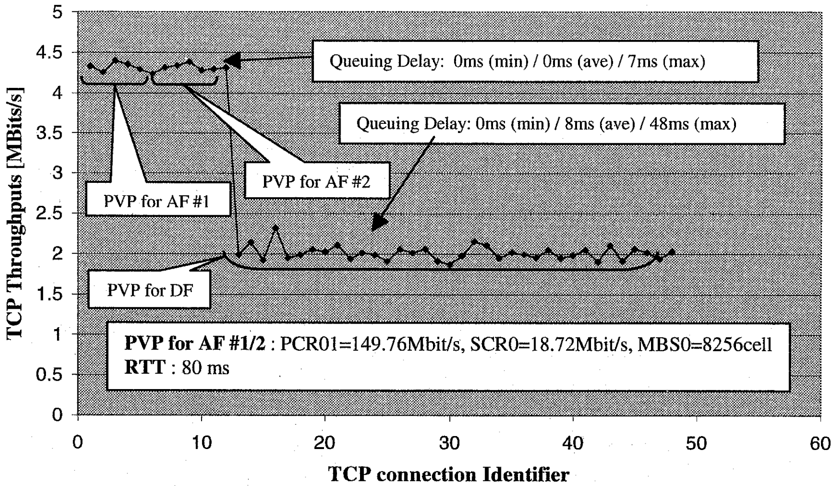

Figure 2 Each TCP throughput under PVPs for AF and DF PHBs.

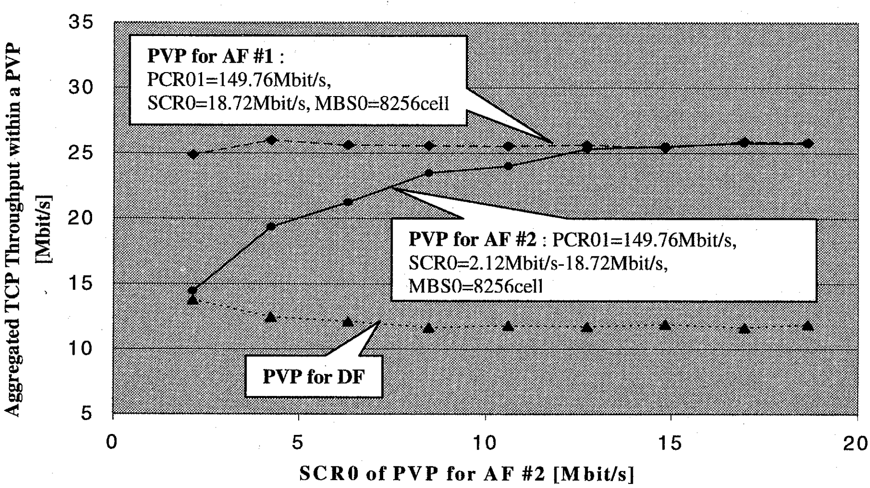

Figure 3 Effect of SCR0 with RTT = 80 ms.

4.2.2 Effect of MBS0

As for influence of MBS, no influences can be observed independent of the congestion level (RTT=20ms, 80ms, 160ms). Figure 5 shows the result. We can say that changing MBS0 values have no effect on the differentiation between AF PHB classes.

5. Conclusions

The goal of the above experiments was to study the differentiation between AF PHB and DF PHB over the ATM PVP services. We mapped AF PHB to SBR3 with SCD and DF PHB to UBR respectively. Using this mapping strategy, many experiments by altering the traffic contract values in the following results are obtained:

(1) The differentiation between AF PHB and DF PHB over the ATM networks is realized during the congestion. ATM SBR3 with SCD can assign each TCP connection belonging to the PVPs for AF PHB the estimated TCP throughput.

(2) The differentiation to assured packet rates is realized under some constraints. First, the difference between assured packet rates must be large. However, the required difference is not analyzed quantitatively in this extended abstract. Second, TCP must use window sizes large

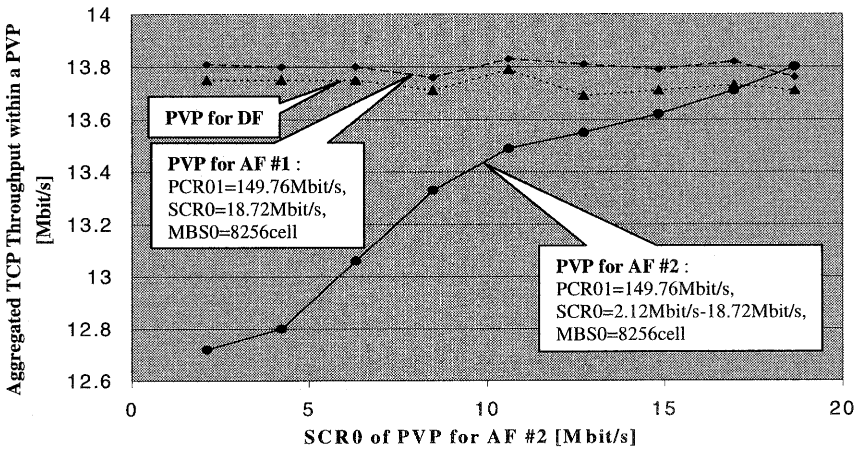

Figure 4 Effect of SCR0 with RTT = 160 ms.

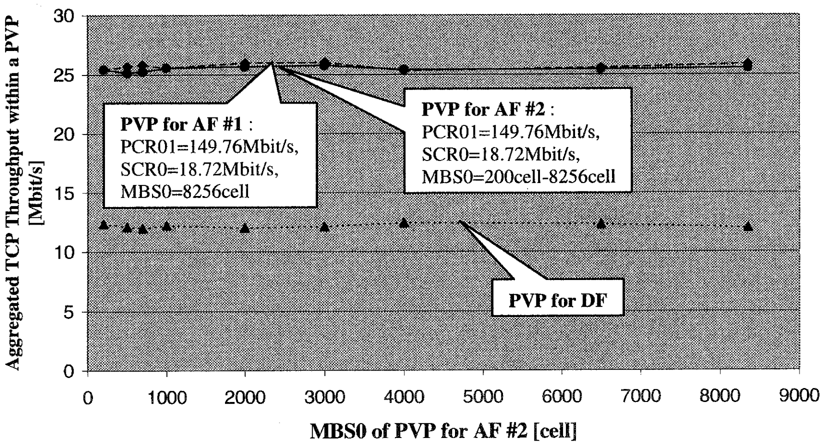

Figure 5 Effect of MBS0 with RTT = 80 ms.

enough for the bandwidth delay product of the network.

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

References

[1] S. Blake et al., "An Architecture for Differentiated Services," IETF RFC2475, Dec. 1998.

[2] D. Clark and W. Fang, "Explicit Allocation of Best Effort Packet Delivery Service," IEEE/ACM Transactions on Networking, Volume 1, Number 4, August 1993, pp. 397-413.

[3] ITU-T recommendation I.371, "Traffic control and congestion control in B-ISDN," Geneva 1996.

[4] O. Aboul-Magd et al., "Mapping of Diff-Serv to ATM Categories," The ATM Forum / 99-0093, 1999

[5] S. Ano, T. Hasegawa, T. Kato, K. Narita, and K. Hokamura, "Performance evaluation of TCP traffic over wide area ATM network," IEEE ATM'97 Workshop, May 1997.