

QoS Provisioning Strategy for QoS-Unaware Applications in an IntServ/DiffServ Interoperating Network

Teerapat Sanguankotchakorn[†], Senior Member IEEE and Ruenun Urwijitaroon^{† †}

[†]School of Engineering and Technology, Asian Institute of Technology, Thailand

^{† †}Huawei Technologies (Thailand) Co., Ltd.

Summary

Due to the demands for variety of services in the Internet, the quality assurance and reliability in IP network have been increasingly important. DiffServ is a scalable QoS-enabled architecture proposed to be an Internet core network. The network has to provide both new QoS-aware and legacy QoS-unaware applications with appropriate resources provisioning. Admission control is a key component to achieve this goal. However, most of the proposed admission control algorithms support only QoS-aware applications, even though most of the current Internet applications are QoS-unaware.

This work proposes an admission control algorithm for providing QoS guaranteed services to QoS-aware applications and special marking to QoS-unaware applications when the network is in light-load situation. The proposed algorithm utilizes the analyzable curve-based calculation for both admission control and special marking decision.

The simulation results show that the proposed algorithm can provide the deterministic QoS guarantees to QoS-aware applications as well as QoS-unaware applications without any QoS degradation.

Key words: Call Admission Control, DiffServ, IntServ, QoS.

1. Introduction

1.1 Background

The current Internet architecture is best-effort service provision. However, with requirement on emerging real-time applications, such as Internet telephony and video-conferencing, the Internet QoS (Quality of Service) provisioning has become more important. To address this problem, two major mechanisms have been developed and proposed by the Internet Engineering Task Force (IETF): Integrated Services (IntServ) and Differentiated Services (DiffServ).

To date, IntServ has been developed for supporting two classes of applications [1]: Real-time applications (with strict bandwidth and latency requirements) and Elastic applications (with less significance in latency requirements). IntServ model works on per-flow mechanism and it provides QoS guarantee by reserving requested resource along the

path of the flow by means of Resource ReSerVation Protocol (RSVP).

While IntServ has attractions on providing deterministic end-to-end QoS guarantee, it obviously embraces two problems [2]. First, IntServ model is a connection-oriented QoS model which is not suitable for Internet short-lived flow. Moreover, end-to-end resource reservation causes the interdomain problem. Second, IntServ model has the scalability problem because all routers have to maintain the state of each flow traverses on them.

An alternative concept for a scalable QoS supporting Internet is DiffServ model. The basic idea of DiffServ is to aggregate the flows into limited number of service classes depending on their service requirements by using Differentiated Service Code Point (DSCP) in the Type of Service (ToS) byte of IP headers. In order to overcome scalability problem, only boundary routers process traffic on a per-flow basis and aggregate the flows with similar QoS requirement into the same Per-Hop Behavior (PHB). Core routers only forward packets based on their PHBs. Since there is no need to maintain per-flow states in the core routers, DiffServ is more scalable than IntServ. However, DiffServ cannot provide stringent end-to-end QoS guarantees to an individual flow because of best-effort treatment within a service class [3].

In this work, we propose the scheme to provide QoS guarantee in interoperation approach of IntServ/DiffServ model for QoS-aware applications and QoS-unaware applications by special marking method. The basic idea is to use DiffServ as the core network and IntServ as the access network while admission control is implemented as border router between IntServ and DiffServ domain. The proposed algorithm utilizes the curve-based calculation for both admission control and marking decision. This paper is structured as follows: Section 2 deal with the preliminary and general concept, Section 3 explains the proposed scheme; Section 4 illustrates the simulation results while Section 5 concludes the works. All detailed proof of the formula can be found at Appendix.

2. Preliminary and the Proposed General Concept

2.1 Analytical Network Model

The network model following the general reference framework proposed in RFC2998 [2], as shown in Fig. 2.1, is used.

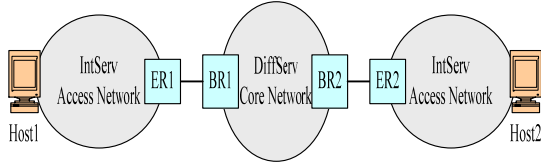


Fig. 2.1: Analytical Network Model [2]

With this architecture, translation between the two domains is performed at the DiffServ border routers. An end-to-end QoS support is feasible if an appropriate admission control and mapping policy between two domains are satisfied.

2.1.1 Services in IntServ Access Network

In this work, the IntServ access network provides two types of service: controlled-load (CL) service, and best-effort (BE) service. For controlled-load service, the flows are classified into two classes: delay-sensitive class and loss-sensitive class.

The CL service is used for QoS-aware applications that require timing property guarantee or probability of delivery guarantee. The delay-sensitive CL service provides timely QoS guarantee while loss-sensitive CL service provides loss probability QoS guarantee. Both are depending on their required TSpec parameters.

The BE service is used for QoS-unaware applications. This service treats all packets equally as much as they can without any guarantee.

To provide QoS guarantee to IntServ CL service flows, the IntServ network utilizes the generalized processor sharing (GPS) discipline [8] as its scheduler.

2.1.2 Services in DiffServ Core Network

The main consideration of this work is on the service in DiffServ core network. The DiffServ core network provides two types of services consisting of two classes of AF (Assured Forwarding) and best-effort (BE) [4].

The AF service is used for serving IntServ CL service flows. One of AF classes is used for serving IntServ delay-sensitive CL flows, while another is used for IntServ loss-sensitive CL flows. These service classes are also used for IntServ BE flows if the available resources in DiffServ domain are enough. In case of serving IntServ BE flows, the flows are mapped into any appropriate AF classes, but they have to be ensured that they will not deteriorate IntServ CL services. The BE is the service for two types of IntServ flows:

- IntServ BE flows that can be mapped to one of DiffServ AF classes, but there is not enough resource to ensure that there is no QoS deterioration.
- IntServ BE flows that cannot be mapped to any DiffServ AF class.

The services provided in both AF classes are relative to each other. The flows in delay-sensitive AF class are guaranteed to obtain the service with less maximum packet delay than the flows in loss-sensitive AF class. On the other hand, the flows in loss-sensitive AF class are guaranteed to obtain the service with less maximum packet loss probability than delay-sensitive AF class. However, in case that the relative class has very small traffic transmission that brings very small delay for loss-sensitive class or loss probability for delay-sensitive class, the flows in delay-sensitive class and loss-sensitive class will be guaranteed at a predefined threshold value.

2.1.3 IntServ/DiffServ Admission Control Policy

In order to provide QoS guarantee in IntServ/DiffServ environment, the function of admission control is the most important. Even though the IntServ network is able to provide per flow QoS guarantees, this IntServ/DiffServ network can provide only per class guarantee because of DiffServ treatment nature.

In this work, the admission control agent responses to the request of IntServ CL flows. The policies for making the decision of acceptance or rejection of the requesting flow are as follows:

- For delay-sensitive class, the effective bandwidth of delay sensitive traffic aggregation including both requesting and admitted flows must be less than the maximum capacity of delay sensitive class as shown in Eq. 2.1.

$$E_D^{ds} \{A_{req}^{ds} + \sum A_i^{ds}\} < C_{ds} \quad (2.1)$$

Where A_{req}^{ds} is the arrival curve of requesting delay-sensitive flow, A_i^{ds} is the arrival curve of the i th admitted flow of delay-sensitive class, C_{ds} is the maximum capacity of delay-sensitive class, and $E_D^{ds} \{A_{req}^{ds} + \sum A_i^{ds}\}$ is the effective bandwidth with delay constraint D of aggregated delay-sensitive traffic. The required delay constraint (D) for effective bandwidth calculation is

$$D = \max[\lambda d_{ls,max}, d_{th}] \quad (2.2)$$

Where λ is delay weight ($0 < \lambda \leq 1$), $d_{ls,max}$ is the maximum packet delay of loss-sensitive class, d_{th} is the packet delay threshold for delay-sensitive class.

- For loss-sensitive class, the equivalent capacity of loss sensitive traffic aggregation including both requesting and admitted flows must be less than the maximum capacity of loss-sensitive class as shown in Eq. 2.2.

$$F_L^{ls} \{A_{req}^{ls} + \sum A_i^{ls}\} < C_{ls} \quad (2.3)$$

Where A_{req}^{ls} is the arrival curve of requesting loss-sensitive flow, A_i^{ls} is the arrival curve of the i th admitted flow of loss-sensitive class, C_{ls} is the maximum capacity of loss-sensitive class, and $F_L^{ls} \{A_{req}^{ls} + \sum A_i^{ls}\}$ is the equivalent capacity with loss constraint L of aggregated loss-sensitive traffic. The required packet loss constraint (L) for equivalent capacity calculation is

$$L = \max[\gamma_{ds,max}, l_{th}] \quad (2.4)$$

Where γ is the loss weight ($0 < \gamma \leq 1$), $\gamma_{ds,max}$ is the maximum loss probability of delay-sensitive class, l_{th} is the loss probability threshold for loss-sensitive class.

For both classes, the constraints used in calculating effective bandwidth and equivalent capacity are depending on the amount of traffic accepted in another class.

These constraints are calculated for guaranteeing less delay and loss probability of delay- and loss-sensitive classes, respectively. In case that the relative class has very small traffic admitted, the threshold value will be used as the calculation constraint. The procedure of Flow Admission Control (AC) algorithm can be illustrated as shown in Fig. 2.2.

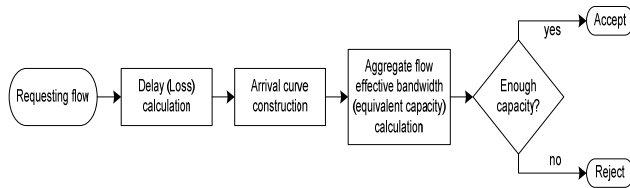


Fig.2.2 Proposed Flow Admission Control Algorithm

2.1.4 Best-effort (BE) Traffic Promotion Policy

QoS-unaware applications are normally treated as BE which will be accepted and forwarded by DiffServ network as much as possible.

Normally, in traditional DiffServ network, all BE traffics will be marked as default PHB which is the lowest priority forwarding treatment. However, based on our proposed scheme, when BE traffic enters DiffServ domain, the network will perform the AC process but with different algorithm than CL flows. The AC algorithm for BE traffic is used for marking the IntServ BE traffic to appropriate DiffServ PHB rather than the default PHB. In BE marking process, Multi-Field (MF) classifier is used in this model. The MF classifier uses five parameters in packet header, namely source and destination addresses, source and destination ports, and protocol ID and the knowledge of available resources at that time to mark the packets into the

appropriate DiffServ PHB. The IntServ BE packets can be classified to AF or BE class based on the following rules:

- The BE packets will be marked into DiffServ delay-sensitive class if and only if the source and destination ports identify that the application requires timely QoS service and available resources within delay-sensitive AF PHB is enough to ensure that the delay-sensitive CL flows using that AF PHB will not be deteriorated.
- The BE packets will be marked into DiffServ loss-sensitive class if and only if the source and destination ports identify that the application requires probability of delivery QoS service and available resources within loss-sensitive AF PHB is enough to ensure that the service to loss-sensitive CL flows will not be deteriorated.
- Otherwise, the packets will be served in BE class in DiffServ domain.

By these BE marking rules, the BE traffic can be promoted to higher priority DiffServ class and can receive better service when the load is light in the network. The BE promotion policy is summarized as shown in Fig. 2.3.

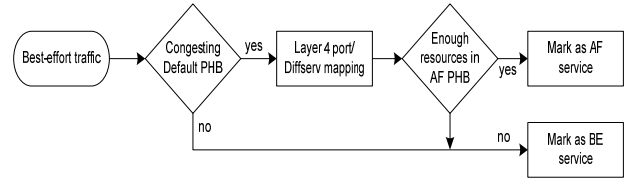


Fig.2.3 Proposed Best-effort (BE) Packet Marking Policy

2.2 Mathematical Analysis

2.2.1 IntServ Controlled-load (CL)

As mentioned in Section 2.1.1, the IntServ network is assumed to utilize the concept of Generalized Processor Sharing (GPS) as its scheduling policy.

By the concepts of GPS [7] and network calculus [5], the three fundamental bounds of particular controlled-load flow traverses the GPS which is implemented as IntServ network can be derived.

a) GPS Service Analysis

In order to analyze the output curve, delay, and backlog bound of IntServ flow traversing the GPS network, the service curve of GPS system has to be firstly analyzed.

By work conserving concept of GPS and the service characteristic, the service curve of GPS system depends on number of flows in busy and idle period. The service-curve of GPS system can be expressed as general form of piecewise linear functions as follows: (detailed proof in Appendix A)

$$S_u(t) = s_k(t - t_{k-1}) + S_u(t_{k-1}) \\ = s_k t - \sum_{j=0}^{k-1} (s_{j+1} - s_j) t_j; t_{k-1} < t \leq t_k, \forall k = 1, 2, \dots, N.$$

(2.5)

where $S_u(t)$ is the universal service curve which is identical in all flows, $S_u(t) = S_i(t)/\phi_i \forall i$, t_k is the time such that the k^{th} flow ends its busy period, s_k is the universal slope of $S_u(t)$ in the region of $t_{k-1} < t \leq t_k$, $t_0 = 0$ and $s_0 = 0$.

b) IntServ Controlled-load (CL) Analysis

In this work, the IntServ/DiffServ AC agent uses the output curve of the IntServ flow as the arrival curve to calculate the resource requirement. Combining the GPS universal service curve equation and network calculus basic, the output curve, delay, and backlog bound can be derived. The output curve of IntServ flows can be classified into three different cases: (1) The slope of the service case is always greater than traffic token rate (2) The slope of the service case is greater or equal to traffic token rate (3) The traffic token rate is always greater than the slope of service case (detailed proof in Appendix B).

Regarding delay and backlog bound calculation; delay bound can be expressed as the maximum horizontal distance between arrival and output curve, while backlog bound can be expressed as the maximum vertical deviation between arrival and output curve.

2.2.2 Admission Control Resource Calculation

The traffic arrival to DiffServ network, both delay- and loss-sensitive classes, can be constrained by the aggregation of output curve of IntServ controlled-load flows. As mentioned in Section 2.2.1, the IntServ's output flows can be classified into three different cases. However, the output curves of IntServ flow are piecewise linear functions in all cases. Hence, the output curve of IntServ flow being considered as per-flow arrival curve of DiffServ network can be rewritten as

$$A_i^*(t) = \begin{cases} 0 & \dots & t \leq \tau_0^* \\ a_1^* t & \dots & \tau_0^* < t \leq \tau_1^* \\ a_2^*(t - \tau_1^*) + A_i^*(\tau_1^*) & \dots & \tau_1^* < t \leq \tau_2^* \\ \vdots & \vdots & \\ a_N^*(t - \tau_{N-1}^*) + A_i^*(\tau_{N-1}^*) & \dots & \tau_{N-1}^* < t \leq \tau_N^* \\ a_{N+1}^*(t - \tau_N^*) + A_i^*(\tau_N^*) & \dots & t > \tau_N^* \end{cases} \quad (2.10)$$

where $a_1^*, a_2^*, \dots, a_N^*, a_{N+1}^*$ are slope of output-curves of IntServ flows, and $\tau_0^*, \tau_1^*, \tau_2^*, \dots, \tau_N^*$ are the time such that the

slope of IntServ output curve changes. The DiffServ arrival

curve of aggregate IntServ flows (A_{DS}) which is the summation of all per-flow arrival curve can be expressed as

$$A_{DS}(t) = \begin{cases} 0 & \dots & t \leq \tau_0 \\ a_1 t & \dots & \tau_0 < t \leq \tau_1 \\ a_2(t - \tau_1) + A_{DS}(\tau_1) & \dots & \tau_1 < t \leq \tau_2 \\ \vdots & \vdots & \\ a_K(t - \tau_{K-1}) + A_{DS}(\tau_{K-1}) & \dots & \tau_{K-1} < t \leq \tau_K \\ a_{K+1}(t - \tau_K) + A_{DS}(\tau_K) & \dots & t > \tau_K \end{cases} \quad (2.11)$$

where $a_1, a_2, \dots, a_K, a_{K+1}$ are slope of aggregate arrival

curve of DiffServ network, and $\tau_0, \tau_1, \tau_2, \dots, \tau_K$ are the time such that the slope of arrival curve changes.

Resource Required for Delay-sensitive Class

For delay-sensitive AF class, the required resource to guarantee the maximum packet delay D is characterized by aggregate effective bandwidth (detailed proof in Appendix C):

$$E_D^{ds} = \max \left\{ a_{K+1}^{ds}, \frac{V_i^{ds}}{D + \tau_i^{ds}} \forall i \in 1, 2, \dots, K \right\} \quad (2.12)$$

where E_D^{ds} is the effective bandwidth of delay-sensitive class, a_{K+1}^{ds} is the long term slope of arrival curve which equals the summation of token bucket rate of all delay-sensitive flows, and V_i^{ds} is the amount of bits arrival at critical time τ_i^{ds} .

a) Resource Required for Loss-sensitive Class

For loss-sensitive AF class, the required resource to guarantee the maximum packet loss probability L is characterized by aggregate equivalent capacity (detailed proof in Appendix C):

$$F_L^{ls} = \max \left\{ a_{K+1}^{ls}, \frac{(1-L)V_i^{ls} - Q_{ls}}{\tau_i^{ls}} \forall i \in 1, 2, \dots, K \right\} \quad (2.13)$$

where F_L^{ls} is the equivalent capacity of loss-sensitive class,

a_{K+1}^{ls} is the long term slope of arrival curve which is equal to the summation of token bucket rate of all loss-sensitive flows,

V_i^{ls} is the amount of bit arrival at critical time τ_i^{ls} , and Q_{ls} is the maximum buffer size of loss-sensitive class.

2.2.3 Available Resource for Best-effort Promotion

In order to promote BE packet into appropriate higher priority class without QoS deterioration, the available

resource of higher priority class guaranteeing the required QoS service has to be calculated. The high priority available resource that can be used for serving BE traffic is again calculated based on the concept of effective bandwidth and equivalent capacity.

Regarding the BE traffic that can be promoted to delay-sensitive AF class, the available resource for promotion is calculated using effective bandwidth and can be expressed as (detailed proof in Appendix D):

$$R_D^{be} = \min \left\{ C_{ds} - a_{K+1}^{ds}, \frac{(D + \tau_i^{ds})C_{ds} - V_i^{ds}}{\tau_i^{ds}} \quad \forall i \in 1, 2, \dots, K \right\} \quad (2.14)$$

where R_D^{be} is the available bandwidth of delay-sensitive class that can serve the promoted BE packets and C_{ds} is the maximum capacity of the delay-sensitive class.

Regarding the BE traffic that can be promoted to loss-sensitive AF class, the available resource for promotion is calculated using equivalent capacity which can be expressed as (detailed proof in Appendix D):

$$R_L^{be} = \min \left\{ \frac{C_{ls}\tau_i^{ls} + Q_{ls} - (1-L)V_i^{ls}}{(1-L)\tau_i^{ds}} \quad \forall i \in 1, 2, \dots, K \right\} \quad (2.15)$$

where R_L^{be} is the available bandwidth of loss-sensitive class that can serve the promoted BE packets and C_{ls} is the maximum capacity of the loss-sensitive class.

The available bandwidth for promotion is used as the maximum rate that the BE packets can be marked as AF service.

Even though, the limit of maximum BE promotion rate can guarantee the non-degradation of the actual AF traffic's QoS, however, there is no prediction of the forthcoming AF arrivals. If the new AF flows simultaneously request for their services, the QoS may be deteriorated.

In order to avoid this situation, three additional conditions are also used as BE promotion policy:

- (1) The BE traffic cannot be marked as AF class if the effective bandwidth or equivalent capacity of traditional AF traffic is more than the half of maximum capacity.
- (2) The BE traffic cannot be marked as AF class if the queue length of AF class is more than the half of maximum buffer size.
- (3) The BE traffic cannot be marked as AF class if the cumulative average queue rate of AF class reaches the reserved effective bandwidth or equivalent capacity.

3. Proposed Scheme Simulation Model

3.1 Proposed Scheme Simulation

For validity verification, Network Simulator (NS-2) [10] is used as a tool to simulate the proposed algorithm.

3.1.1 Simulation Network Model and Parameters

The network model used in simulation follows the analytical network model as illustrated in Fig. 2.1. Since this work focuses only on the AC at DiffServ border routers, the network model can be actually simplified to the one shown in Fig. 3.1.

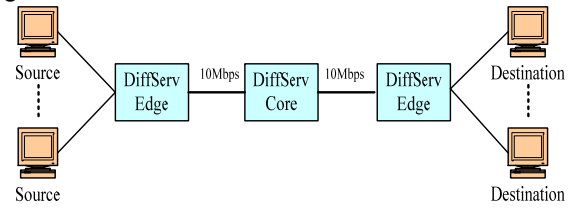


Fig.3.1 Simplified Network Model

In simulation model, the IntServ networks are assumed to be merged in the source and destination hosts. The source hosts will generate three types of traffic: delay-sensitive CL service, loss-sensitive CL service, and BE (QoS-unaware). For all traffic types, MPEG-4 and VBR source are used as traffic generator.

In DiffServ domain, only one DiffServ core node is used for representing DiffServ cloud. The AC process is enabled at DiffServ edge node for both acceptance on QoS-aware traffic and marking on QoS-unaware traffic. Within DiffServ domain, three classes of DiffServ PHBs are used: two classes of AF, and one class of BE.

3.1.2 Performance Evaluation Criteria

The performance evaluation criteria are defined as follows:

- Acceptance Probability of controlled-load
- Average Throughput (Per-class network utilization)
- Average Packet Queuing Delay
- Packet Loss Probability

The comparison is done between this proposed scheme and typical IntServ/DiffServ model without admission control algorithm.

3.1.3 Network Simulator Modifications

To simulate the proposed network model, four parts of NS-2 are modified or developed: DiffServ module, IntServ agent, VBR traffic source, and performance tracing tool.

a) DiffServ Module

The original built-in DiffServ module in NS-2 has no function of admission control. Therefore, DiffServ module is modified in two parts: Admission Control Agent (for CL requesting), and BE Marker (for BE promotion algorithm),

as shown in Fig. 3.2. The marking algorithm of BE marker follows the flow chart in Fig. 2.3.

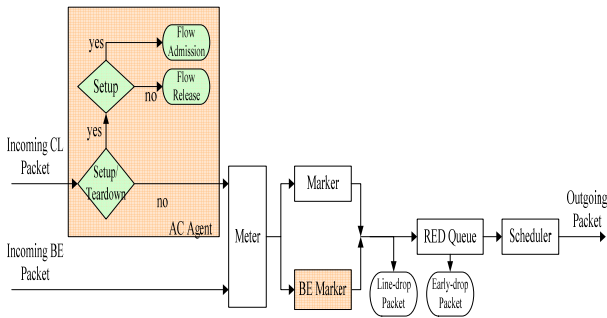


Fig.3.2 Modified NS-2 DiffServ Module Functions (modified from [10])

b) IntServ Agent

The IntServ network is simplified by merging into source and destination hosts. That is, the IntServ agent is created to function as the conformance policer of IntServ network. Any traffic source using IntServ agent can transmit only the packets conforming to the requested TSpec parameters. The IntServ agent functions are shown in Fig. 3.3.

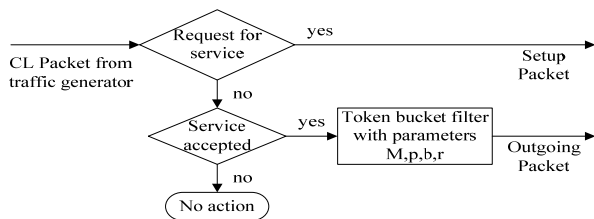


Fig.3.3 IntServ Agent Functions

c) Traffic Source

The traffic sources used in this work are MPEG-4 and VBR. The MPEG-4 traffic is created by traffic generator [6]. This traffic source is used in general case scenario simulation. In the worst case simulation where all traffic sources have to burst their traffic simultaneously, VBR traffic is used instead due to the MPEG-4 traffic statistical characteristic limitation.

d) Performance Tracing Tool

The original NS-2 software has already supported probability of controlled-load acceptance and probability of packet loss tracing function. Only packet delay and per-class network utilization tracing tools are newly developed for this work.

The packet delay tracing tool uses the timestamp in packet header for providing the time arrival of packet at the network and the queuing delay is collected when the packet is transmitted from that network element. At DiffServ egress node, the cumulative delay is collected as Edge-to-Edge Packet Queuing Delay as shown in Fig.3.4.

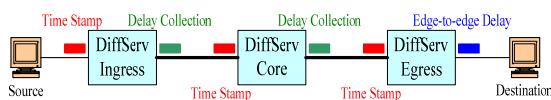


Fig.3.4 The Edge-to-Edge Queuing Delay Tracing

Regarding the per-class network utilization tracing tool, it uses the byte-count value stored in queue module to calculate average packet transmission rate. This value is kept with per-class basis. Since this tracing tool is implemented in queue module, this function can be enabled at any point of the network.

3.1.4 Simulation Scenarios

We perform three scenarios: (1) Admission control of bursting traffic (2) Best effort promotion of bursting traffic (3) Best-effort promotion of normal traffic. The fixed parameters used in all scenarios are shown in Table 3.1 and 3.2.

Table 3.1: Network Model Parameters

Network Model Parameters	Parameter Value
Per-class maximum capacity	10 Mbps
Per-class buffer size	100 Packets
Link delay	5 ms

Table 3.2: MPEG-4 and VBR Traffic Parameters

Traffic Type	Peak Rate (bps)	Bucket Rate (bps)	MBS* (bytes)	MTU** (bytes)
MPEG-4	1,325,877	210,994	5,524	500
VBR	1,325,877	210,994	5,524	500

Note: * = Maximum Burst Size, ** = Maximum Transmission Unit.

The rest of parameters are varying depending on the scenario.

Scenario 1: Admission Control of Bursting Traffic

In this scenario, we study the effect of AC parameters on network performance. In each case, the effect of delay threshold and loss probability threshold are studied. The parameters used in this simulation are illustrated in Table 3.3 and 3.4.

Table 3.3: Parameter of Traffic Sources in Scenario 1

Traffic Source	Number of Traffic Sources	Inter Arrival Time (sec)	Holding Time (sec)
VBR	50	0.0	1.0

Table 3.4: Parameters Setting in Scenario 1

Parameters	Control Variables	
	Delay Threshold	Loss Probability Threshold
Delay Threshold	Vary	0.0001
Loss Probability	10	Vary

Scenario 2: Best-effort Promotion of Bursting Traffic

In this scenario, the effect of the proposed BE promotion algorithm on network performance in burst situation where all traffic sources start bursting simultaneously is studied.

We consider the effect of amount of CL traffic in the network instead of AC parameters as in previous scenario. The parameters used in this scenario are illustrated in Table 3.5.

Table 3.5: Parameters of Traffic Sources in Scenario 2

Traffic	Number of Traffic Sources	Inter Arrival Time (sec)	Holding Time (sec)
VBR	200	0.0	1.0

For all cases, the AC parameters: delay threshold and loss probability threshold are set to 5 ms and 0.001, respectively.

Scenario 3: Best-effort Promotion of Normal Traffic

In this scenario, the effect of proposed BE promotion scheme in normal traffic situation is studied. The AC parameters are the same as in scenario 2. The traffic parameters are shown in Table 3.6.

Table 3.6: Parameters of Traffic Sources in Scenario 3 [6]

Traffic	Number of Traffic Sources	Inter Arrival Time (sec)	Holding Time (sec)
MPEG-4	200	0.05	2.0

4. Simulation Results and Discussions

4.1 Scenario 1: Admission Control of Bursting Traffic

4.1.1 Effect of Delay Threshold on Network Performance

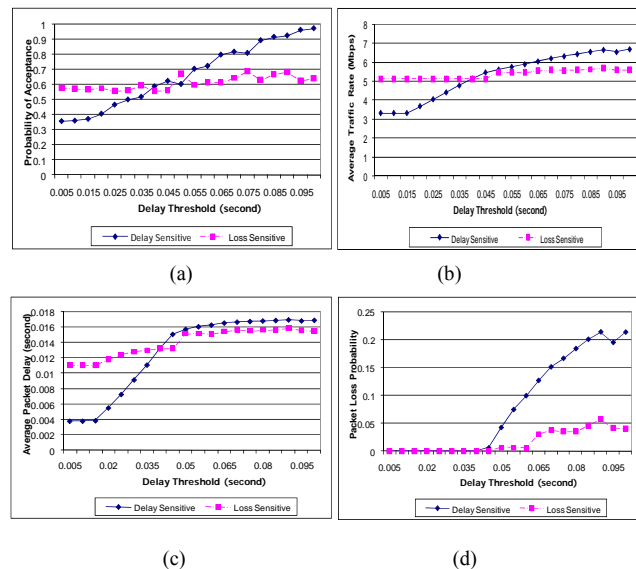


Fig.4.1 Effect of Delay Threshold on (a) Acceptance Probability (b) Average Throughput (c) Average Delay (d) Packet Loss Probability

a) Discussion of Simulation Results

Fig. 4.1 (a) and (b) illustrate the same trend of results. This is because when delay threshold is very small, many effective bandwidths are required to accept one more flow. When delay threshold is large, the network cannot accommodate any more traffic.

In Fig.4.1 (c), in case of Delay-sensitive class, the packet delay proportional to the number of accepted flows while in case of Loss-sensitive class, the increment of delay threshold affects packet delay. This is because of round robin scheduling manner. When delay threshold is small, a small number of Delay-sensitive packets are transmitted in the network. Therefore, the capacity in the network will be available to serve Loss-sensitive packets which results in less delay service.

In Fig. 4.1 (d), the results show that if delay threshold is less than the maximum queuing delay, the network will always be able to serve all packets, then causing no queuing loss. However, if delay threshold is more than the maximum queuing delay, the Delay-sensitive traffic will be over-admitted which leads to increment of packet loss of both traffics.

4.1.2 Effect of Loss Probability Threshold on Network Performance

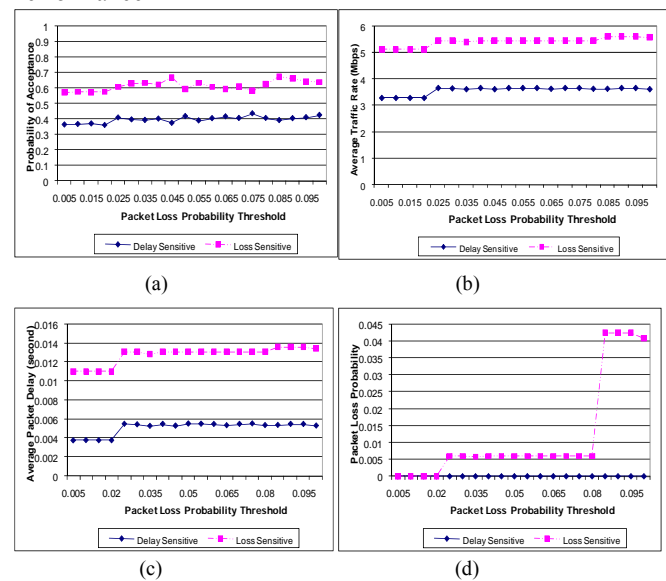


Fig.4.2 Effect of Loss Probability Threshold on (a) Acceptance Probability (b) Average Throughput (c) Average Packet Delay (d) Packet loss Probability

a) Discussion of Simulation Results

The effect of loss probability threshold on Acceptance Probability, Average throughput and Average delay (Fig. 4.2 (a)-(d)) is not so much perceptible as the effect of delay threshold. Since the requirement of loss probability

threshold of 10-6 or 10-5 does not much influence admission control result. However, as illustrated in Fig. 4.2 (d), only the packet loss probability of Loss-sensitive class is affected by loss probability threshold.

4.2 Scenario 2: BE Promotion of Bursting Traffic

4.2.1 Effect of BE promotion of Bursting Traffic on Network performance

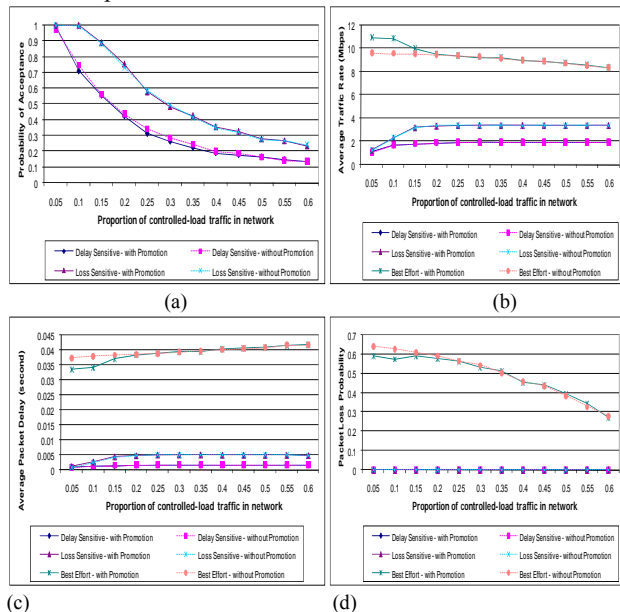


Fig.4.3 Effect of BE Promotion on (a) Acceptance Probability (b) Average Throughput (c) Average Packet Delay (d) Packet Loss Probability

a) Discussion of Simulation Results

The results in scenario 2 simulation can be considered in two topics: the effect of BE promotion of bursting traffic on CL and BE traffic.

First, consider the effect of BE promotion scheme on CL traffic. From Fig. 4.3 (a)-(d), the results show that the BE promotion scheme has no effect on QoS guarantees of CL service. The CL flows still receive the same service guarantees as if there is no BE promotion; although some BE traffics are marked as CL AF class.

Next we consider the effect of BE promotion on BE traffic. Fig. 4.3 (b) shows that only in light CL situation, the BE traffic can achieve the better throughput. This is because the BE marking policy requires reserving a half of maximum capacity for guaranteeing QoS to forthcoming CL flows. In case of light CL traffic, there is available bandwidth for promoting BE traffic to AF class. Hence, some BE traffics share the resource with CL traffic and make higher overall BE traffic throughput. However, in heavy CL traffic, only small or none of BE traffic will be marked as AF class. The BE traffic receives the same treatment as no BE promotion scheme.

In Fig. 4.3 (c), since BE traffic shares some resource of CL traffic, the promoted packets will receive the same service as CL flows. This reduces the overall average BE packet delay comparing to typical marking scheme. Again, in heavy CL situation, there is not enough resource for promoting BE packets. The performance in terms of average packet delay is not improved.

In Fig. 4.3 (d), It is obvious that the less packet loss probability is achieved when the BE promotion scheme is enabled. The difference on packet loss probability between proposed and typical scheme depends on how many packets can be promoted. Because the promoting policy limits the promotion only if the queue of BE class is full, the packet loss probability is proportional to the number of packets promoted.

4.3 Scenario 3: BE Promotion of Normal Traffic

4.3.1 Effect of BE Promotion of Normal Traffic on Network Performance

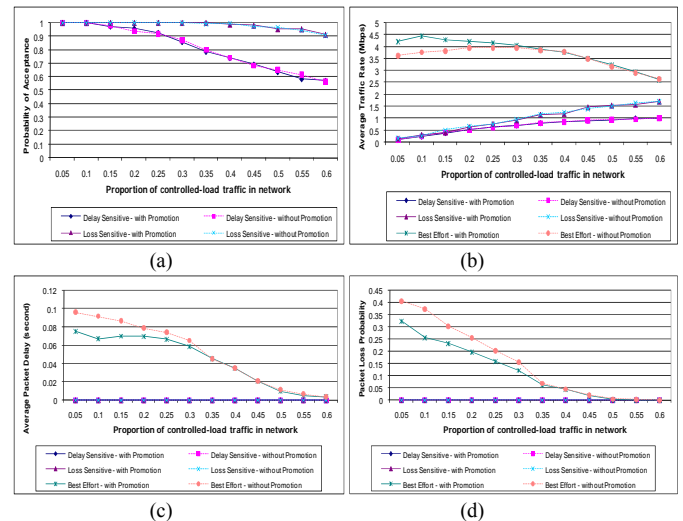


Fig.4.4 Effect of BE Promotion on (a) Acceptance Probability (b) Average Throughput (c) Average Packet Delay (d) Packet Loss Probability

a) Discussion of Simulation Results

Fig. 4.4 (a)-(d) show that the BE promotion algorithm does not affect the QoS guarantees of CL traffic, while improves the performance of BE traffic treatment. The transmission rate of BE traffic is greater than the typical marking policy. The average packet queuing delay and packet loss probability are reduced.

Comparing to the results in scenario 2, this scenario illustrates better performance when the CL traffic is increased, since the flows do not concurrently send their packets. At any instantaneous time, there is always available resource to promote BE traffic. Hence, more packets are promoted and the network performance is better improved.

5. Conclusions

Admission control agent is the most important network element in QoS provisioning in IntServ/DiffServ interoperating network to provide QoS guarantee. In this research, the QoS provisioning algorithms used in IntServ/DiffServ environment are studied and analyzed.

The class-relative QoS guarantee policy is proposed. This policy provides QoS guarantee using class service status as QoS parameter. The Delay-sensitive applications will receive less-delay service guarantee, while Loss-sensitive applications will receive less-loss service guarantee. The curve-based admission control policy used to provide these QoS guarantee is proposed and analyzed. This admission control policy utilizes the concept of effective bandwidth and equivalent capacity in resource calculation. However, with curve-based admission control, the low network utilization cannot be avoided. In this work, special marking policy for BE packets is proposed to improve the network utilization as well as the service to QoS-unaware applications. This marking policy will mark the BE packets as appropriate AF traffic if there is enough available resource on AF class to guarantee no QoS degradation.

To verify the proposed schemes: class-relative admission control and BE promotion algorithm, three simulation scenarios are established. Two types of traffic source are used: MPEG-4 and VBR traffic.

From simulation results, it can be concluded that the proposed admission control algorithm can provide class-relative QoS guarantees in both worst-case and normal-case situations. Regarding the BE promotion algorithm, the simulation results show that there is no effect on the service of CL traffic in both burst and normal situations. The QoS provided to Delay-sensitive and Loss-sensitive applications is not deteriorated. In light CL situation, the service provided to BE traffic is improved in terms of average packet delay and packet loss probability as well as the network utilization.

Appendix

A. Derivation of GPS Service Curve

To analyze the service curve of GPS offered to IntServ CL flows. The worst case situation called "all-greedy regime" [7] is considered.

Let $i = 1, 2, \dots, N$ be the i^{th} CL admitted flow. Each flow must conform to the token bucket parameter (b_i, r_i) as defined in RFC2212 [8], where b_i and r_i are the bucket size and bucket rate of the i^{th} CL admitted flow, respectively. Hence, the i^{th} flow traffic transmitted to the network is

$$A_i(t) \leq \min[M_i + p_i t, b_i + r_i t] \quad \text{for } t \geq 0 \quad (\text{A.1})$$

where M_i and p_i are the maximum datagram size and peak rate of the i^{th} CL admitted flow. By the definition of CL service, the peak rate (p_i) is not defined and it can be assumed to be the line rate of incoming interface [12]. For simplification, without loss of generality, we assume that the peak rate of flow is infinite. Hence,

$$A_i(t) \leq b_i + r_i t \quad \text{for } t \geq 0 \quad (\text{A.2})$$

Let $S_i(t)$ denotes the service curve of the i^{th} flow, ϕ_i denotes the GPS assignment (weight) of i^{th} flow and C denotes the GPS service capacity. The worst case scenario occurs when all sources start sending packets at their maximum rate at the same time $t = 0$ [7]. In this case, the arrival curve of i^{th} flow is

$$A_i(t) = b_i + r_i t. \quad (\text{A.3})$$

Let $S_u(t)$ denotes the universal service curve which is

$$S_u(t) = \frac{S_i(t)}{\phi_i} \quad \forall i$$

identical for all flows, t_k denotes the time such that the k^{th} flow ends its busy period ($0 < t_1 \leq t_2 \leq \dots \leq t_N$), S_k denotes the universal slope which is the slope of $S_u(t)$ in the region of $t_{k-1} < t \leq t_k$, B_k denotes the set of flows which are still in busy period when $t > t_{k-1}$.

Consider the time $0 < t \leq t_1$; all flows are in busy period. So, the service curve of each flow is

$$S_i(t) = \left(\phi_i / \sum_{j=1}^N \phi_j \right) C t \quad ; 0 < t \leq t_1$$

(A.4)

$$S_u(t) = \left(C t / \sum_{j=1}^N \phi_j \right) \quad ; 0 < t \leq t_1$$

(A.5)

and

$$t_1 = \inf_{i \in B_1} \left\{ t > 0 : \frac{A_i(t)}{\phi_i} = S_u(t) \right\} \quad (\text{A.6})$$

$$= \inf_{i \in B_1} \left\{ t > 0 : \frac{b_i + r_i t}{\phi_i} = \frac{C t}{\sum_{j=1}^N \phi_j} \right\}$$

$$s_1 = C / \sum_{j=1}^N \phi_j \quad (\text{A.7})$$

Next, consider the time $t_1 < t \leq t_2$, the $i=1^{st}$ flow already ends its busy period but still transmits packets at its token

bucket rate r_1 .

$$S_i(t) = \frac{(C-r_1)\phi_i}{\sum_{j \in B_2} \phi_j} (t-t_1) + S_i(t_1) \quad ; t_1 < t \leq t_2$$

(A.8)

$$S_u(t) = \frac{(C-r_1)}{\sum_{j \in B_2} \phi_j} (t-t_1) + S_u(t_1) \quad ; t_1 < t \leq t_2$$

and

$$t_2 = \inf_{i \in B_2} \left\{ t > 0 : \frac{A_i(t)}{\phi_i} = S_u(t) \right\}$$

$$= \inf_{i \in B_2} \left\{ t > 0 : \frac{b_i + r_1 t}{\phi_i} = \frac{(C-r_1)}{\sum_{j \in B_2} \phi_j} (t-t_1) + S_u(t_1) \right\}$$

$$s_2 = (C-r_1) / \sum_{j \in B_2} \phi_j$$

Then, consider the time $t_{k-1} < t \leq t_k$, the $i < k^{th}$ flows already end their busy period, but still transmit packet at rate r_i .

$$S_i(t) = \frac{(C - \sum_{j \in B_k} r_j) \phi_i}{\sum_{j \in B_k} \phi_j} (t-t_{k-1}) + S_i(t_{k-1}) \quad ; t_{k-1} < t \leq t_k$$

(A.12)

$$S_u(t) = \frac{(C - \sum_{j \in B_k} r_j)}{\sum_{j \in B_k} \phi_j} (t-t_{k-1}) + S_u(t_{k-1}) \quad ; t_{k-1} < t \leq t_k$$

(A.13)

and

$$t_k = \inf_{i \in B_k} \left\{ t > 0 : \frac{A_i(t)}{\phi_i} = S_u(t) \right\}$$

$$= \inf_{i \in B_k} \left\{ t > 0 : \frac{b_i + r_i t}{\phi_i} = \frac{(C - \sum_{j \in B_k} r_j)}{\sum_{j \in B_k} \phi_j} (t-t_{k-1}) + S_u(t_{k-1}) \right\}$$

$$s_k = (C - \sum_{j \in B_k} r_j) / \sum_{j \in B_k} \phi_j$$

Hence, the service curve of GPS server is

$$S_u(t) = \begin{cases} s_1 t; & 0 < t \leq t_1 \\ s_2(t-t_1) + S_u(t_1); & t_1 < t \leq t_2 \\ \vdots & \\ s_k(t-t_{k-1}) + S_u(t_{k-1}); & t_{k-1} < t \leq t_k \\ \vdots & \\ s_N(t-t_{N-1}) + S_u(t_{N-1}); & t_{N-1} < t \leq t_N \end{cases}$$

(A.16)

(A.9) And let $t_0 = 0$ and $s_0 = 0$, then (A.16) can be rewritten in simpler form as

$$S_u(t) = s_k(t-t_{k-1}) + S_u(t_{k-1})$$

$$= s_k t - \sum_{j=0}^{k-1} (s_{j+1} - s_j) t_j \quad ; t_{k-1} < t \leq t_k, \forall k = 1, 2, \dots, N$$

(A.10) 17)

(A.11)

B. Derivation of IntServ Controlled-Load Output Flows

To analyze the output curve of IntServ flows, the definition of arrival curve, service curve, and network calculus [5] are used. The arrival curve of IntServ controlled-load flow is $A_i(t) = b_i + r_i t$ and the service curve of GPS server is in (A.17). The output curve is the min-plus deconvolution between arrival and service curve:

$$A_i^*(t) = (A_i \oslash S_i)(t) = \sup_{u \geq 0} \{A_i(t+u) - S_i(u)\} \quad (B.1)$$

Let $s'_k = \phi_i s_k$ be the k^{th} slope of i^{th} flow. Then,

$$S_i(t) = \phi_i S_u(t) = \phi_i s_k t - \phi_i \sum_{j=0}^{k-1} (s_{j+1} - s_j) t_j$$

(B.2)

$$= s'_k t - \sum_{j=0}^{k-1} (s'_{j+1} - s'_j) t_j; t_{k-1} < t \leq t_k, k = 1, 2, 3, \dots, i$$

$$A_i^*(t) = \sup_{u \geq 0, 1 \leq k \leq i} \left\{ b_i + r_i(t+u) - s'_k u + \sum_{j=0}^{k-1} (s'_{j+1} - s'_j) t_j \mid t_{k-1} < u \leq t_k \right\}$$

$$= \sup_{u \geq 0, 1 \leq k \leq i} \left\{ b_i + (r_i - s'_k) u + \sum_{j=0}^{k-1} (s'_{j+1} - s'_j) t_j \mid t_{k-1} < u \leq t_k \right\}$$

(B.3)

(A.14)

Define t_m be the time such that $S_i(t_{m-1}) \leq b_i < S_i(t_m)$, t_n be the time such that $s'_n > r_i$ and let $t_0 = 0$ and $s'_0 = 0$. There are three different cases of controlled-load

(A.15) output curve.

Case I: $m > (n = 0)$

In this case, the slope of service curve is always greater than traffic token rate.

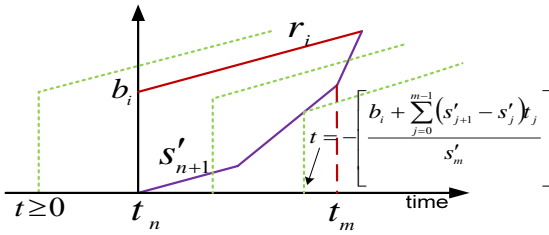


Fig B.1 Graphical computation of case I IntServ output curve

As shown in Fig. B.1, we define

$$T = \left[\left(b_i + \sum_{j=0}^{m-1} (s'_{j+1} - s'_j) t_j \right) / s'_m \right]$$

$$\text{For } t \geq 0, \quad A_i^*(t) = A_i(t+u) - S_i(u)$$

$$\text{At } u=0, \quad A_i^*(t) = A_i(t) = b_i + r_i t$$

(B.4)

$$\text{For } -T \leq t < 0, \quad A_i^*(t) = A_i(t+u) - S_i(u)$$

$$\begin{aligned} \text{At } u = -t, \quad A_i^*(t) &= A_i(0) - S_i(-t) \\ &= \left\{ b_i + \sum_{j=0}^{k-1} (s'_{j+1} - s'_j) t_j + s'_k t \mid -t_k \leq t < -t_{k-1} \right\} \end{aligned}$$

(B.5)

$$\text{For } t < -T, \quad A_i^*(t) = 0 \quad (\text{B.6})$$

Hence,

$$A_i^*(t) = \begin{cases} 0; & t < -T \\ b_i + \sum_{j=0}^{k-1} (s'_{j+1} - s'_j) t_j + s'_k t; & -T \leq t < 0 \text{ and } -t_k \leq t < -t_{k-1} \\ b_i + r_i t; & t \geq 0 \end{cases}$$

(B.7)

Case II: $m \geq n > 0$

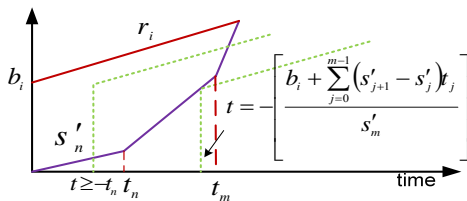


Fig. B.2 Graphical computation of case II IntServ output curve

As shown in Fig.B.2, for $t \geq -t_n$, $A_i^*(t) = A_i(t+u) - S_i(u)$

At $u = t_n$,

$$\begin{aligned} A_i^*(t) &= b_i + r_i(t + t_n) - s'_n t_n + \sum_{j=0}^{n-1} (s'_{j+1} - s'_j) t_j \\ &= b_i + (r_i - s'_n) t_n + \sum_{j=0}^{n-1} (s'_{j+1} - s'_j) t_j + r_i t \end{aligned}$$

(B.8)

$$\text{For } -T \leq t < -t_n, \quad A_i^*(t) = A_i(t+u) - S_i(u)$$

At $u = -t$,

$$A_i^*(t) = \left\{ b_i + \sum_{j=0}^{k-1} (s'_{j+1} - s'_j) t_j + s'_k t \mid -t_k \leq t < -t_{k-1} \right\}$$

(B.9)

$$\text{For } t < -T, \quad A_i^*(t) = 0 \quad (\text{B.10})$$

Hence,

$$A_i^*(t) = \begin{cases} 0; & t < -T \\ b_i + \sum_{j=0}^{k-1} (s'_{j+1} - s'_j) t_j + s'_k t; & -T \leq t < 0 \text{ and } -t_k \leq t < -t_{k-1} \\ b_i + (r_i - s'_n) t_n + \sum_{j=0}^{n-1} (s'_{j+1} - s'_j) t_j + r_i t; & t \geq -t_n \end{cases}$$

(B.11)

Case III: $n > m > 0$

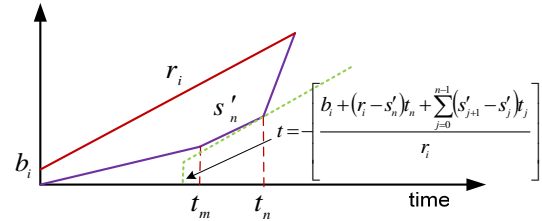


Fig. B.3 Graphical computation of case III IntServ output curve As shown in Fig. B.3, we define

$$T' = \left[\left\{ b_i + (r_i - s'_n) t_n + \sum_{j=0}^{n-1} (s'_{j+1} - s'_j) t_j \right\} / r_i \right]$$

$$\text{For } t \geq -T, \quad A_i^*(t) = A_i(t+u) - S_i(u)$$

$$\text{At } u = t_n, \quad A_i^*(t) = b_i + (r_i - s'_n) t_n + \sum_{j=0}^{k-1} (s'_{j+1} - s'_j) t_j + r_i t$$

(B.12)

$$\text{For } t < -T, \quad A_i^*(t) = 0 \quad (\text{B.13})$$

Hence,

$$A_i^*(t) = \begin{cases} 0; & t < -T' \\ b_i + (r_i - s'_n) t_n + \sum_{j=0}^{n-1} (s'_{j+1} - s'_j) t_j + r_i t; & t \geq -T' \end{cases}$$

(B.14)

C. Derivation of Effective Bandwidth and Equivalent Capacity

Delay-sensitive Class Effective Bandwidth Calculation

For simplification, the representation of arrival curve of DiffServ network in Eq. 2.11 is modified to be

$$A_{DS}(t) = \begin{cases} 0 & \dots & t \leq \tau_0 \\ a_1 t & \dots & \tau_0 < t \leq \tau_1 \\ a_2(t - \tau_1) + V_1 & \dots & \tau_1 < t \leq \tau_2 \\ \vdots & \dots & \vdots \\ a_K(t - \tau_{K-1}) + V_{K-1} & \dots & \tau_{K-1} < t \leq \tau_K \\ a_{K+1}(t - \tau_K) + V_K & \dots & t > \tau_K \end{cases} \quad (C.1)$$

where $V_i = A_{DS}(\tau_i)$; $i = 1, 2, 3, \dots, K$.

The effective bandwidth is the minimum capacity that is required for serving the arrival traffic with a constraint on delay [6]. The effective bandwidth E_D^{ds} for delay-sensitive arrival curve $A_{ds}(t)$ and delay constraint D is

$$E_D^{ds} = \sup_{s \geq 0} \left\{ \frac{A_{ds}(s)}{s + D} \right\} \quad (C.2)$$

As shown in Eq. C.1, the arrival curve of delay-sensitive aggregate input traffic is piecewise linear function. With this kind of arrival curve, the maximum value of Eq. C.2 can

occur only at the critical point $\tau_1, \tau_2, \tau_3, \dots, \tau_K$ and can be represented by the maximum slope of straight line that passes the critical point and the point $-D$ as shown in Fig. C.1.

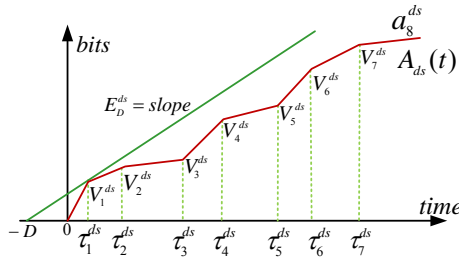


Fig. C.1 Effective bandwidth of DiffServ delay-sensitive class

In the case that the maximum slope of the line is less than the aggregate token bucket rate, the effective bandwidth will be limited to the aggregate token bucket rate. Hence, the effective bandwidth of delay-sensitive class can be expressed as

$$E_D^{ds} = \max \left\{ a_{K+1}^{ds}, \frac{V_i^{ds}}{D + \tau_i^{ds}} \forall i \in 1, 2, \dots, K \right\} \quad (C.3)$$

Loss-sensitive Class Equivalent Capacity Calculation

The equivalent capacity is the minimum capacity required for serving the arrival traffic with a constraint on backlog [6].

The equivalent capacity F_L^{ls} for loss-sensitive arrival curve $A_{ls}(t)$ and loss constraint L is

$$F_L^{ls} = \sup_{s \geq 0} \left\{ \frac{A_{ls}(s) - Q_{ls}(1 + L)}{s} \right\} \quad (C.4)$$

where Q_{ls} is the maximum buffer size of loss-sensitive class.

With loss constraint L , the loss probability relies on the amounts of bits arrival at considered critical point. The equivalent capacity can be derived as follows:

Let R_i^{ls} be the minimum service rate satisfying the maximum loss constraint L at the time τ_i^{ls} .

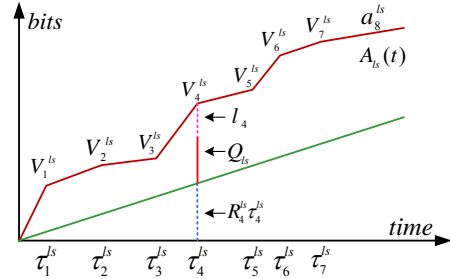


Fig. C.2 Equivalent capacity calculation example at point τ_4^{ls}

$$l_i = \frac{V_i^{ls} - Q_{ls} - R_i^{ls} \tau_i^{ls}}{V_i^{ls}} = L \quad (C.5)$$

$$R_i^{ls} = \frac{(1-L)V_i^{ls} - Q_{ls}}{\tau_i^{ls}} \quad (C.6)$$

The equivalent capacity is the maximum value of R_i^{ls} and aggregate token bucket rate. Hence, the equivalent capacity of loss-sensitive class is

$$F_L^{ls} = \max \left\{ a_{K+1}^{ls}, \frac{(1-L)V_i^{ls} - Q_{ls}}{\tau_i^{ls}} \forall i \in 1, 2, \dots, K \right\} \quad (C.7)$$

D. Derivation of Available Promotional Bandwidth

The Available Promotional Bandwidth for Delay-sensitive Class

Consider the promoted best-effort traffic as a flow with

constant traffic rate R_{be} . By utilizing the concept of effective bandwidth, the required resource for providing service with delay constraint D to the aggregated traffic between promoted best-effort and delay-sensitive controlled-load is

$$E_D^{ds} = \max \left\{ a_{K+1}^{ds} + R_{be}, \frac{V_i^{ds} + R_{be} \tau_i}{D + \tau_i} \forall i \in 1, 2, \dots, K \right\} \quad (D.1)$$

The capacity for promoting best-effort traffic is the maximum rate of promoted best-effort traffic that ensures the guarantee of delay constraint D . Therefore, the available capacity for promotion is the rate R_{be} that makes the effective bandwidth equals the maximum capacity of delay-sensitive class.

Let R_D^{be} be the available bandwidth of promotion to delay-sensitive class. We get the maximum capacity of delay-sensitive class as

$$C_{ds} = \max \left\{ a_{K+1}^{ds} + R_D^{be}, \frac{V_i^{ds} + R_D^{be} \tau_i^{ds}}{D + \tau_i^{ds}} \forall i \in 1, 2, \dots, K \right\} \quad (D.2)$$

Hence, the available bandwidth of promotion is

$$R_D^{be} = \min \left\{ C_{ds} - a_{K+1}^{ds}, \frac{(D + \tau_i^{ds})C_{ds} - V_i^{ds}}{\tau_i^{ds}} \forall i \in 1, 2, \dots, K \right\} \quad (D.3)$$

The Available Promotional Bandwidth for Loss-sensitive Class

Similarly, consider the promoted best-effort traffic as a flow with constant traffic rate R_{be} . By concept of equivalent capacity, the required resource for providing service with loss constraint L to the aggregated traffic between promoted best-effort and loss-sensitive controlled-load is

$$F_L^{ls} = \max \left\{ a_{K+1}^{ls} + R_{be}, \frac{(1-L)(V_i^{ls} + R_{be} \tau_i^{ls}) - Q_{ls}}{\tau_i^{ls}} \forall i \in 1, 2, \dots, K \right\} \quad (D.4)$$

The capacity for promoting best-effort traffic is the maximum rate of promoted best-effort traffic that ensures the guarantee of loss constraint L . Therefore, the available capacity for promotion is the rate R_{be} that makes the effective bandwidth equals the maximum capacity of loss-sensitive class.

Let R_L^{be} be the available bandwidth of promotion to loss-sensitive class.

$$C_{ls} = \max \left\{ a_{K+1}^{ls} + R_L^{be}, \frac{(1-L)(V_i^{ls} + R_L^{be} \tau_i^{ls}) - Q_{ls}}{\tau_i^{ls}} \forall i \in 1, 2, \dots, K \right\} \quad (D.5)$$

Hence, the available bandwidth of promotion is

$$R_L^{be} = \min \left\{ C_{ls} - a_{K+1}^{ls}, \frac{C_{ls} \tau_i^{ls} + Q_{ls} - (1-L)V_i^{ls}}{(1-L)\tau_i^{ls}} \forall i \in 1, 2, \dots, K \right\} \quad (D.6)$$

Acknowledgements

The authors wish to thank Prof. Akinori Nishihara of The Center for Research and Development of Educational Technology, Tokyo Institute of Technology for his support and valuable comments. This work was supported partially by a grant from National Institute of Information and Communication Technology (NICT), Japan.

References

- [1] Armitage, G., "Quality of Service in IP Networks: Foundations for a Multi-Service Internet," Macmillan Technical Publishing, Indianapolis, USA, 2000.
- [2] Bernet, Y., Ford, P., Yavatkar, R., Baker, F., Zhang, L., Speer, M., Braden, R., Davie, B., Wroclawski, J. and Felstaine, E., "A Framework for Integrated Services Operation over DiffServ Networks", RFC 2998, IETF, 2000
- [3] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z. and Weiss, W., "An Architecture for Differentiated Services," RFC 2475, IETF, 1998.
- [4] Carpenter, B. E. and Nichols, K., (2002), "Differentiated Services in the Internet," Proceedings of the IEEE, Vol. 90, No. 9, September, 2002, pp. 1479-1494.

- [5] Le Boudec, J. Y. and Thiran, P., "Network Calculus: A Theory of Deterministic Queuing Systems for the Internet," URL: http://icalwww.epfl.ch/PS_files/NetCal.htm, 2002.
- [6] Le Boudec, J.Y., "Application of Network calculus to Guaranteed service Networks", IEEE Transactions on Information Theory, Vol.11, No.3, May1998, pp.1087-1096.
- [7] Matrawy, A., Lambadaris I. and Huang C., "MPEG4 Traffic Modeling using The Transform Expand Sample Methodology," Proceedings of the 4th IEEE International Workshop on Networked Appliances (IWNA4), Gaithersburg, USA, January, 2002.
- [8] Parekh, A. K. and Gallager, R. G., "A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Single-Node Case," IEEE/ACM Transactions on Networking, Vol. 1, No. 3, June, 1993, pp. 344-357.
- [9] Shenker, S., Partridge, C. and Guerin, R., "Specification of Guaranteed Quality of Service", RFC 2212 IETF.
- [10] Urwijitarn R., "QoS Provisioning Strategy for QoS-Unaware Applications in an IntServ/DiffServ Interoperating Network" M.Eng. Thesis, Telecommunications Program, School of Advanced Technologies, Asian Institute of Technology, Pathumthani, Thailand, 2003.
- [11] Thomas, A., "Supplying Legacy Applications with QoS: A Description Syntax at Application, End-user and Network Level," IASTED International Conference on Software Engineering and Applications (SEA2002), Cambridge, USA, November, 2002.
- [12] Wroclawski, J., "Specification of the Controlled-Load Network Element Service", RFC 2211, IETF.
- [13] NS-2 <http://www.isi.edu/nsnam/ns>



Teerapat Sanguankotchakorn was born in Bangkok Thailand on 08th December 1965. He received the B.Eng in Electrical Engineering from Chulalongkorn University, Thailand in 1987, M.Eng and D.Eng in Information Processing from Tokyo Institute of Technology, Japan in 1990 and 1993, respectively under Japanese Government Scholarship. In 1993, he joined

Telecommunication and Information Systems Research Laboratory at Sony Corporation, Japan, where he holds two patents on Signal Compression. He joined Asian Institute of Technology in October 1998 as an Assistant Professor. Currently, he is an Associate Professor at Telecommunications Field of Study, School of Engineering and Technology, Asian Institute of Technology. His research interests include Digital Signal Processing, High Speed Network, IP-based multimedia applications and QoS in IP and Mobile Ad Hoc Network. He is the senior member of IEEE and member of IEICE, Japan



Ruenun Urwijitarn was born in Bangkok Thailand. He received the B.Eng in Electrical Engineering from Chulalongkorn University, Thailand in 2001 and M.Eng in Telecommunication from Telecommunications field of Study, School of Engineering and Technology, Asian Institute of Technology, Thailand in 2003. Currently, he is working at Huawei

Technologies (Thailand) Co., Ltd. in Thailand. His major fields of research are High Speed Network and Quality of Service in IP Network.