A Method of Admission Control Based on Both Resource Requests and Traffic Measurement and Its Dynamics Under On/Off Model Traffic

Yasusi Kanada

Central Research Laboratory, Hitachi, Ltd.
Higashi-Koigakubo 1-280 Kokubunji, Tokyo 185-8601, Japan
E-mail: Yasusi.Kanada.yq@hitachi.com

Abstract: *A method of admission control based on both resource requests by applications and class-based traffic measurement results was developed. In this method, a wide range of admission-control policy can be realized by adjusting three parameters, α, β, and γ. A policy-server prototype using this method and simulated voice traffic was used in traffic measurements. The measurements results show that the proposed method improves bandwidth usage and decreases call-blocking ratio while incurring low measurement load. Interesting but possibly harmful dynamics (i.e., system behavior) were observed by the simulations using traffic generated by an on/off model. That is, this admission-control method may cause oscillation or long-term evolution that lasts for 100 to 150 minutes, and it may also cause bandwidth "overshooting". The range of parameters with which such effects can be properly suppressed and the admission control correctly works was experimentally obtained.*

1. Introduction

Next-generation networks (NGNs), which have been standardized in the ITU-T, 3GPP, and other standardization organizations, contain a resource and admission control function (RACF). Admission control [Geo 08] refers to a judgement that either allows a communication session when the specified resources including the bandwidth are available or denies it when the resources are not available. Admission control is an established technique that has been used for the public switched telephone network (PSTN) and asynchronous transfer mode (ATM) [Lew 98] [Hab 00]. However, there are many problems concerning admission control on IP networks. Among them, there is a problem whether the admission control depends on the bandwidth requested by the application or on the bandwidth given by traffic measurement.

The conventional methods of admission control used for ATM or PSTN depend on a predefined value (i.e., request). A problem with these methods is that most applications usually use much less bandwidth than the requested value; consequently, the network cannot be used efficiently.

Many methods for measurement-based admission control (MBAC), such as, Nam, et al. [Nam 08], have been proposed. These methods allow so-called overbooking, namely, bandwidth allocation in which the sum of allocated maximum bandwidths exceeds the limit, so they reduce wasted bandwidth. However, these methods may fail and cause a prediction error and congestion.

It is almost impossible to avoid prediction error completely. However, if both the request- and measurement-based methods are used for the prediction, a more precise prediction must be possible. A method to combine these two methods was proposed by Georgoulas, et al. [Geo 08]. They used an admission-control method that uses not only mean values but also distribution of traffic amount. However, there are few studies on this issue. The purpose of the present study was to develop a method that combines requests and measurements and to evaluate the system behavior, i.e., dynamics, by simulation using an on/off model. This paper describes the method, and the prototype and the evaluation results of the method.

2. Request-and-Measurement-Based Admission-Control Method

A method of admission control called the request-and-measurement-based admission-control (RMBAC) method is proposed and the values of the parameters of this method are discussed.

2.1 Method

In communications using voice, video, and other multimedia traffic, the framework of NGNs can be used. Resources can be reserved at the beginning of the session, and resources can be released at the end. Session Description Protocol over Session Initiation Protocol (SDP/SIP) can then be used to declare the required resource. If more detailed QoS conditions are to be specified, they can be described by using the extended SDP or using a protocol for resource reservation, such as RSVP (Resource Reservation Protocol).

The traffic is assumed to pass through a backbone network in which QoS is managed using differentiated services (DiffServ). The policy server (or RACF in the terminology of NGNs) of the backbone makes admission-control decisions when reserving resources. So that each class of traffic in DiffServ does not surpass the predefined bandwidth limit, the per-class total bandwidth is managed for each network path from an ingress edge router to an egress edge router or for each path from a subnet to another subnet. A class of traffic that passes through a path is called a *macro flow* hereafter.

The feature of the proposed method is that an admission-control decision is made by using both resource-request and on-line traffic-measurement results. This means that, instead of fully reserving the requested amount of resource, it allows statistical multiplexing, or "overbooking." To avoid traffic overflow, the policy server obtains a per-class measured traffic-amount from each router by using the NetFlow protocol [Cla 04] and uses it for admission control.

Several alternative nodes exist for enforcing admission control. It was decided to put this function on ingress edge nodes and to use a measurement function that the ingress edge nodes of the backbone have. The policy server periodically collects traffic-measurement results per DiffServ class from each edge node by using NetFlow, and the server estimates the traffic once per measurement. The estimated effective bandwidth $\mathrm{Bu}(f, t)$ for macro flow f at time $t + 1$ (the next time period) is calculated by using the following equation with the measured values until time t.

$$\mathrm{Bu}(f, t) = \gamma (1 + \beta) \mathrm{Ba}(f, t) + (1 - \gamma) \mathrm{Br}(f, t)$$

Here, $\mathrm{Ba}(f, t)$ is measured bandwidth, and $\mathrm{Br}(f, t)$ is the sum of requested bandwidths of admitted flows.

The bandwidth limit that can be allocated for macro flow f is defined to be $\mathrm{Bm}(f)$. The difference between $\mathrm{Bm}(f)$ and $\mathrm{Bu}(f, t)$ is the bandwidth that can allocated to new traffic flows. Consequently, if the difference is equal to or less than the requested bandwidth,[1] the

[1] The requested bandwidth is used as the estimated effective bandwidth of

policy server allocates the bandwidth. The server gets the measurement result per macro flow by NetFlow and calculates an exponential moving average to smooth the measured values by using the following equation.

$$\mathrm{Ba}(f, t) = \alpha\,\mathrm{ba}(f, t) + (1 - \alpha)\,\mathrm{Ba}(f, t - 1),\ \ \mathrm{Ba}(f, 0) = \mathrm{Bm}(f)$$

Here, ba(f, t) is the raw measured data obtained by NetFlow. In this admission-control method, three policy parameters, α, β, and γ, which are explained below, are used.

- *Smoothing coefficient* α: a parameter that controls the smoothness of the moving average (i.e., smoothed measured bandwidth) ($0 \le \alpha \le 1$).
- *Bandwidth margin ratio* β: a parameter that controls the margin between the bandwidth limit and the smoothed measured bandwidth ($0 \le \beta \le 1$).
- *Measured-value contribution ratio* γ: the contribution ratio (or the relative weight) of the smoothed measured bandwidth by the estimated bandwidth limit for the admission control ($1 - \gamma$ is the contribution ratio of requested bandwidths) ($0 \le \gamma \le 1$).

2.2 Numerical example

An example of numerical values derived by the RMBAC method is shown in **Figure 2.1**. It is assumed that no bandwidth is newly allocated in the time span shown in the figure. The parameter values are $\alpha = 0.3$, $\beta = 0.5$, and $\gamma = 0.7$. In this case, the moving average is smoother than the measured values, ba(f, t).

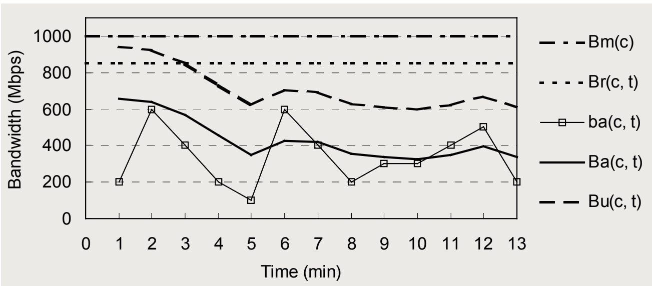


Figure 2.1: Measured and estimated values of traffic amount ($\alpha = 0.3$, $\beta = 0.5$, and $\gamma = 0.7$)

2.3 Various types of admission control

By using the proposed method and selecting values of parameters α, β, and γ, various types of admission control, as described below, can be realized; in other words, an admission-control policy can be selected from a wide range. The best selection depends on the nature of the network traffic.

- *CBR-style control that only depends on request total* Br(f, t): If $\alpha = 0$, the measured values are ignored, and the difference between bandwidth limit Bm(f) and requested bandwidth total Br(f, t) can be allocated to new traffic at time t. In Figure 2.1, the difference between Bm(f) and Bu(f, t), i.e., 150 Mbps, can be allocated for new traffic. This method handles new traffic as constant bit rate (CBR) traffic.
- *Memory-less control that only depends on last measured value* ba(f, t): If $\alpha = 1$, only the last measured value, i.e., ba(f, t), and requested total Br(f, t) are used for the control; the older values at time t, ba(f, tt) ($tt < t$), are ignored (no moving average is calculated). Examples of estimated values are show in **Figure 2.2** (where only the value of α is different from Figure 2.1).

---

the new flow, because no measurement data is available for the flow.

- *Measurement-based control that only depends on smoothed measured value* Ba(f, t): If $\gamma = 1$, the requested bandwidths are ignored, and only the measured values are used for the traffic estimation. Examples of estimated values are show in **Figure 2.3** (where only the value of γ is different from Figure 2.1).
- *Full-function control that depends on both request total and smoothed measured value*: If $0 < \alpha < 1$ and $0 < \gamma < 1$, both the requested bandwidths and measured bandwidths are taken into account. However, because an exponential moving average is used, the influence of the old measured values, ba(f, tt) ($tt < t$), decreases exponentially. (See Figure 2.1.)

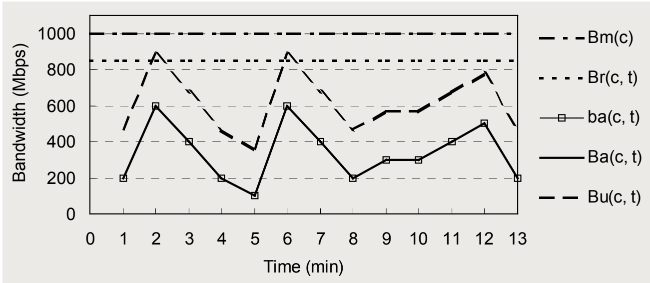


Figure 2.2: Estimated values of traffic amount when $\alpha = 1$ ($\beta = 0.5$, $\gamma = 0.7$)

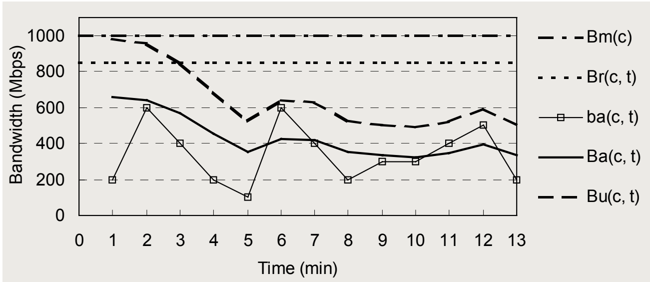


Figure 2.3: Estimated values of traffic amount when $\gamma = 1$ ($\alpha = 0.3$, $\beta = 0.5$)

3. Outline of Prototype

The RMBAC method has been built into our policy-server prototype [Kan 08]. The outline of this prototype is described below.

3.1 Traffic-measurement method

Many commercial routers have NetFlow functions. However, it is not possible to configure a Cisco's or Alaxala's router to report measurement results by NetFlow per interval less than one minute. A shorter interval is preferable for QoS guarantee, and it does not cause large overhead if DiffServ is used, because traffic is classified into a small number of classes. However, we decided to use NetFlow functions because there is currently no widely available better function.

Various formats are available for NetFlow output records. However, our prototype currently only receives destination-prefix-ToS aggregation records. **Table 3.1** lists the contents of this type of record. This format is selected by a configuration command, and 32 (bits) is specified as "mask destination minimum"; consequently, the flows that have the same ToS (including DSCP), the same output interface, and the same autonomous system (AS) are aggregated into one macro flow.

If this method is used, for example, when there are five DiffServ classes and 1000 destinations, the number of records generated by each router is 5000. If each UDP packet is 1500 bytes long and

Table 3.1: Destination-Prefix-ToS Aggregation record of NetFlow

Field name	Explanation
Flows	Number of aggregated flows
Packets	Number of packets in the aggregated flows
Bytes	Total number of bytes of the aggregated flows
First time stamp	Time when the first packet in the flow was received [s]
Last time stamp	Time when the last packet in the flow was received [s]
Destination prefix	Prefix of destination IPv4 address
Destination mask bits	Mask bit of destination IPv4 address prefix mask-bit number.
ToS	Type of service
Destination AS	AS number of destination or destination neighbor peer
Output interface	SNMP interface index of output interface

contains 25 records, a stream of 200 packets per minute (i.e., a 40-kbps stream) is generated. This stream can be sent by a network node and received by the policy server without incurring a large overhead.

3.2 Admission-control method

The RMBAC procedure is outlined below (See **Figure 3.1**).

- *Traffic measurement and estimation of used bandwidth*: Using NetFlow, the policy server periodically receives the traffic-measurement results from each edge nodes. It computes available bandwidth Ba(f, t) by using the equation given in Section 2. It obtains class-based (aggregated) information and smooths it.
- *Processing resource request* (*session information*): If the policy server receives a resource request by SIP through a SIP proxy, it performs admission control using the requested bandwidth and the estimated bandwidth. If the request does not exceed the estimated bandwidth limit, the server allows the request; if it exceeds the limit, the server denies the request. The reply is returned to the requester by means of a SIP reply. If necessary, edge-node configuration commands are sent to the edge nodes.

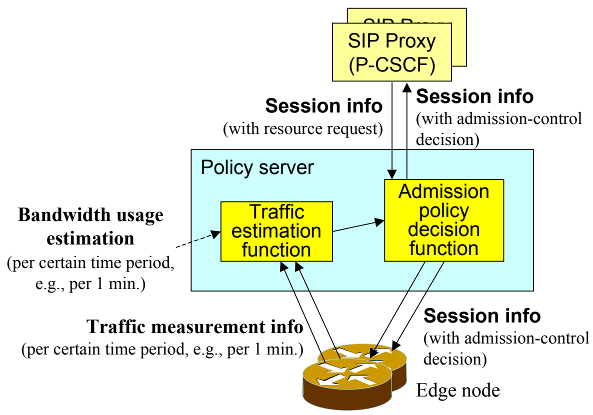


Figure 3.1: Admission control and related functions

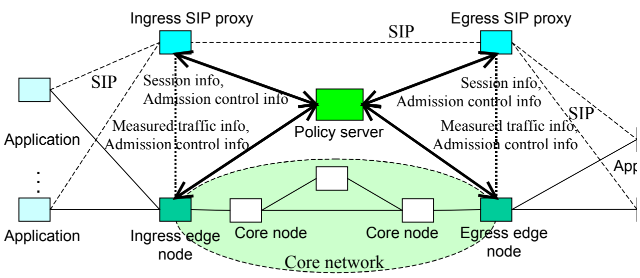


Figure 3.2: Assumed network structure

Figure 3.2 shows an example of a network structure. As shown in this figure, it is assumed that the correspondence between SIP proxies and edge nodes are one-to-one. Each SIP message therefore passes two SIP proxies, i.e., ingress and egress proxies. The policy server receives messages from these proxies by using a protocol called the resource request transport protocol (RRTP), which is a proprietary protocol.[1] The policy server can thus recognize the ingress and egress routers of the flow.

3.3 PS sequencer

A program called a policy-server (PS) sequencer, which simulates part of the network described above, was developed. The PS sequencer can be used with the policy server. It has the function to generate the following information using the built-in traffic model described in the next section and to send it to the policy server instead of using the real network shown in Figure 3.2 (See **Figure 3.3**).

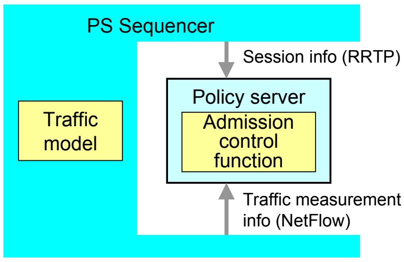


Figure 3.3: PS Sequencer

- *Session information*: In the assumed network, the SIP proxy must look at a SIP message generated by an application, and the session information included in the SIP message should be reported to the policy server. The PS sequencer autonomously generates virtual session information and sends it to the policy server by RRTP.
- *NetFlow information*: The policy server receives simulated traffic-measurement results from the PS sequencer by NetFlow protocol using destination-prefix-ToS aggregation records.

The policy server can receive NetFlow information in various formats and timings according to router configuration. However, the PS sequencer generated this information once per minute (which is a virtual time, and "real time" may be much shorter).

4. Evaluation of Admission-Control Method

The traffic-measurement and admission-control methods in the prototype are evaluated by using the PS sequencer. The evaluation method and the results are explained in this section.

4.1 Evaluation criteria and methods

Two criteria are used for the evaluation.

1. *Call-blocking ratio*: The ratio of call requests denied by the admission control.
2. *Used-bandwidth ratio*: The ratio of allocated bandwidth used. It is computed using measured values. The maximum value is usually 1.0, but it may larger than 1.0 when the traffic exceeds the bandwidth limit.

These criteria are evaluated by varying the parameters shown in Section 3 and by varying the traffic conditions as described below.

The traffic conditions are as follows (see **Figure 4.1**). The PS sequencer simulates voice flows. Arrivals of these voice flows follow a Poisson process (i.e., they are independent of each other) whose mean value of arrival interval is fixed, and the requested bandwidth of the flows is 80 kbps (64 kbps without packet headers). This means

[1] If a standardized protocol should be used, a Diameter [Cal 03] based protocol was standardized in 3GPP.

the arrival intervals are computed using exponential-distribution random numbers.[1] The mean value of the interval was 0.5 to 1.0 seconds. An exponential distribution with 120-second mean time was used for the continuation time (holding time) of the flows.[2] Each voice flow uses 60% (48 kbps) of the requested bandwidth. This means that gaps occupy 40% of each flow in time, and no packets are sent during a gap; that is, silence detection is used.

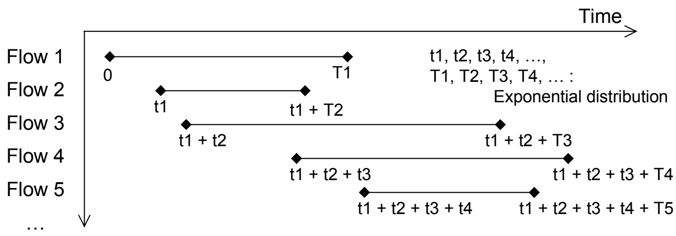


Figure 4.1: Flows virtually generated by the PS sequencer

To generate traffic with a gap (sound-less period), a method of per-flow traffic-generation using the following *on/off model* is used. (Each flow is assumed to start immediately after the session is allowed.) A human voice consists of two parts. One is talkspurt, i.e., a loud part, and the other is a gap, i.e., a silent part. In an on/off model, these parts are simulated by a Markov process, Markov chain, or a hidden Markov process with two states, i.e., on and off. A famous example is Brady's model [Bra 65] with a six-state hidden Markov model. However, a simpler model with a two-state Markov process shown in **Figure 4.2** was used. This figure also shows the transition probabilities used in the evaluation. In this model, the mean continuation time of talkspurt is 1 second (i.e., 50 times), and that of the gap is 667 ms (i.e., 33.3 times). Both of them follow exponential distributions.[3]

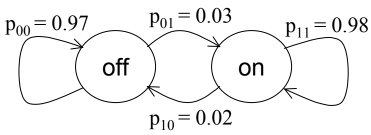


Figure 4.2: Simple on/off model expressed by a two-state Markov process

The reason only one type of the traffic (i.e., voice) was used is that with this admission-control method, the bandwidth is managed class-by-class, so the influence of other types of traffic on the voice traffic is small. Especially, voice traffic is usually given highest precedence (or expedited-forwarding per-hop behavior (EF PHB) is used in DiffServ), the influence is very small.

4.2 Evaluation results

4.2.1 Comparison with a non-measurement method

The RMBAC method was first compared with a conventional admission-control method, in which measured bandwidths are not taken into account. **Figure 4.3** shows the time evolution of call-blocking ratio (CBLR) and used-bandwidth ratio (UBWR) when bandwidth limit Bm(f) is 10 Mbps and the mean flow arrival interval was 0.79 seconds. The parameter values were $\alpha = 1.0$, $\beta = 0.2$, and $\gamma = 0.8$. NetFlow information came from the PS sequencer every minute, at which time the CBLR and UBWR were computed. They were computed twice for each set of parameter values using different random number seeds. The average UBWR was 0.70 and the average CBLR was 0.04 (7 min $\leq t \leq$ 33 min).

Figure 4.4 shows CBLR and UBWR when $\gamma = 0$. Other parameter values are the same as in Figure 4.3. Note that $\gamma = 0$ means the measured values are not taken into account. The admission control was therefore performed only using the requested bandwidth. The average UBWR was 0.57, which is much lower than that in Figure 4.3. This UBWR is a result of the bandwidth usage of this voice traffic; namely, it used 57% of the requested bandwidth on average. As a result, the average CBLR was low, i.e., 0.18. This result is close to the CBLR, 0.2, calculated by using Erlang B formula [Bro 48] with 125 (= 10 M / 80 k) telephone lines.

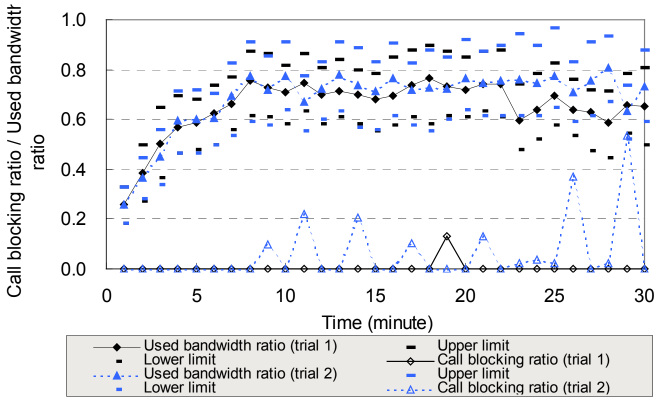


Figure 4.3: CBLR and UBWR when $\alpha = 1.0$, $\beta = 0.2$, and $\gamma = 0.8$

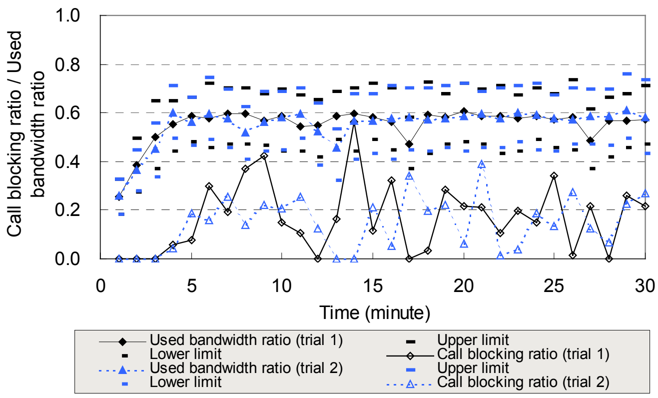


Figure 4.4: CBLR and UBWR when $\alpha = 1.0$, $\beta = 0.2$, and $\gamma = 0$

These graphs also show the upper and lower bounds of UBWR. The meaning of these plots is as follows. UBWR varies within a one-minute measurement interval. The number of packets virtually generated from the on/off model was taken into account, and UBWR was calculated at 20-ms intervals (assumed voice-packet interval). The maximum value (upper bound) and minimum value (lower bound) of UBWRs within a minute are shown in the graphs. The maximum value may exceed 1.0, but it did not in these measurements. The key results on policy parameters α, β, and γ are summarized as follows.

1. The value of α was 1.0. If the traffic was steady, the result did not change much even if α was varied from 0 to 1, except the first 10 minutes.

2. The value of β was 0.2. This means the margin over the measured value was 20%. This is the reason UBWR was bounded to approximately 0.8.

[1] If each call occurs independently, the arrival time follow a Poisson distribution, and the interval of arrivals follow exponential distribution. However, traffic flows in the Internet are not necessarily independent and they do not follow exponential distribution completely.

[2] Bolotin wrote that the distribution of telephone connections is more heavy-tail than exponential distribution [Bol 94].

[3] It is reported that the distribution of talkspurt and gap does not follow an exponential distribution completely [Wen 00] [Dan 04].

3. The value of γ was 0.8. This means the measured bandwidth was much more weighted than the requested bandwidth. It was assumed that there was no traffic when starting the measurements (i.e., when the time was zero). It took 10 minutes to reach near-steady state.

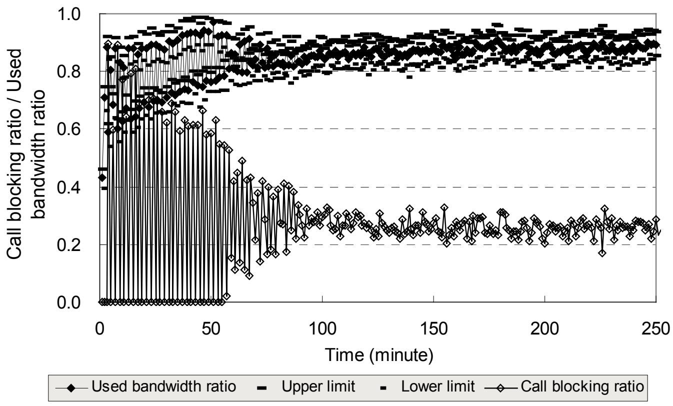
If the policy server knows that each flow uses only 60% of the requested bandwidth before starting the flow, it can increase UBWR without having to use traffic-measurement results. However, the policy server cannot usually see the real bandwidth that a flow uses, and UBWR varies while time passes. As a result, such predefined overbooking is not possible. The measurement enables effective use of bandwidth, as shown in Figure 4.3.

4.2.2 Oscillation of bandwidth usage

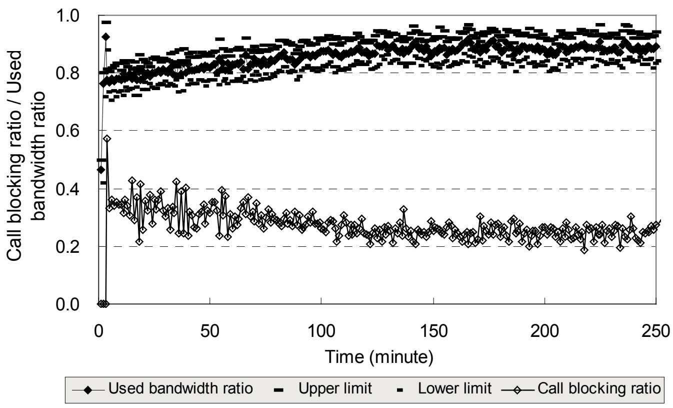
The proposed admission-control method was analyzed by changing the values of parameters and measuring CBLR and UBWR. This subsection describes the *oscillation* of UBWR, which is a type of evolution of this dynamical system.

Figure 4.5 shows the oscillation when the bandwidth limit was 100 Mbps and the value of α was changed. Figure 4.5(a) shows an example of development of CBLR and UBWR when α = 1.0, β = 0.2, and γ = 0.8. Figure 4.5(b) shows one when only the value of α was changed to 0.5; that is, other parameter values were the same. The oscillation converged in 100 minutes when α = 1.0, but it converged in only 10 minutes when α = 0.5. If the bandwidth limit is smaller, it takes more time when α = 1.0. This result is considered to mean that the oscillation makes a response to a steady-state change slower.

Figure 4.5 also shows that there is a much longer-term evolution that could not be observed in previous figures. Both when α = 1.0



(a) When α = 1.0



(b) When α = 0.5

Figure 4.5: Oscillation and convergence of CBLR and UBWR

and when α = 0.5, the average of the CBLR slowly decreased, and the average of UBWR slowly increased up till 100 or 150 minutes. This means UBWR was suppressed in the first 100 or 150 minutes. However, it is not certain whether this effect was simulation-specific (PS sequencer specific).

Xu, et al. [Xu 04] reported a similar oscillation of an admission-control system with feedback of measurement results. If positive feedback is avoided, oscillation can also be avoided. However, because this system is nonlinear, it cannot easily be analyzed, and oscillation is hard to be avoided completely. To suppress the oscillation, Xu, et al., thus proposed a stochastic method of admission control for Web servers. However, this method cannot suppress oscillation completely.

The oscillation in this study is strong when CBLR in Earlang B equation is 0.3 or more. However, if the CBLR is around 0.2, it is hard to be observed because call blocking seldom occur. However, in other experiments, oscillation was observed if the traffic is in a steady state; that is, call blocking sometimes occurs, and UBWR drops sharply immediately after the call blocking.

4.2.3 Overshooting of bandwidth usage

UBWR sometimes overshoots. Namely, this dynamical system may evolve in an interesting but harmful way.

Figure 4.6 shows an example; CBLR and UBWR in two trials, A and B, when the arrival interval of flows is 0.6 second and α = 0.3, β = 0.2, and γ = 0.8. In trial B, UBWR overshot five minutes from the start of the measurement. Although such effect was not observed in trial A, such overshooting was often observed in other experiments. Overshooting may cause pseudo bandwidth-shortage and packet drops, so it should be avoided. A similar effect was also observed when β was varied. However, it was found that large overshooting often occurred if α was small. There thus seems to be a trade-off between overshooting and oscillation[1].

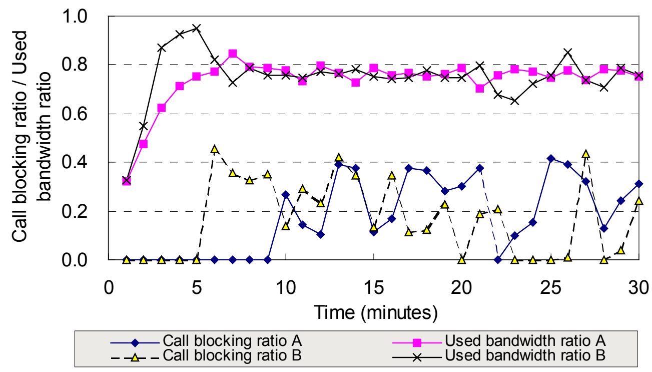


Figure 4.6: CBLR and UBWR when α = 0.3, β = 0.2, and γ = 0.8

4.2.4 Optimum bandwidth margin

If β becomes closer to zero, in addition to increasing UBWR, packet drop and delay may occur more easily. Thus, α was fixed to 0.5, γ was fixed to 0.8, β was varied, and UBWR and possible packet-drop ratio (or, more exactly, the ratio of exceeding the bandwidth limit) were computed. The packet-arrival ratio was selected so that the CBLR calculated by the Erlang B equation became 0.5. CBLR and UBWR were measured for 150 minutes for each set of parameters, and average UBWR and packet-drop ratio (except the first 20 minutes) were calculated.

[1] In linear systems, if a register is added to suppress oscillation, an overshoot (a transitive effect) is also suppressed. However, probably because the admission control system is nonlinear, it behaves differently. However, it is similar to linear systems in generating oscillations and overshoots.

The measurement results are shown in **Figure 4.7**. Three results with different bandwidth limits, i.e., 10 Mbps (corresponds to 125 telephone lines), 30 Mbps (375 lines), and 100 Mbps (1250 lines), are shown. UBWR is almost always lower than $1-\beta$, and it was 0.92 or less if $\beta = 0$.

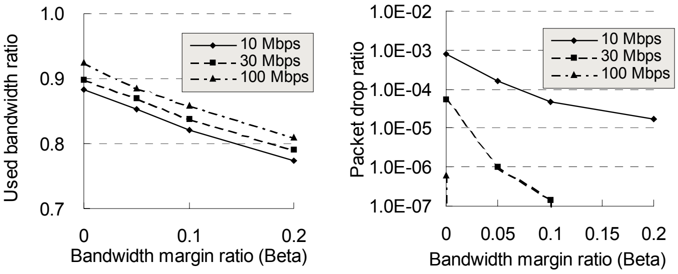


Figure 4.7: UBWR and packet-drop ratio when $\alpha = 0.15$ and $\gamma = 0.8$

To calculate the packet drop ratio, all the packets that exceed the bandwidth limit were assumed to be dropped, although the surplus packets may be forwarded normally or delayed without loss in certain situations. When the bandwidth limit is 10 Mbps, the requested drop ratio is always satisfied if it is 10^{-3}, but it is satisfied only when $\beta \geq 0.1$ if it is 10^{-4}. If the drop ratio is 10^{-5} or less, the measured values are not reliable because of large sampling errors.

5. Conclusion

By measuring traffic at edge routers using NetFlow and by using the RMBAC method, used-bandwidth ratio (UBWR) can be increased, and call-blocking ratio (CBLR) can be decreased. According to the measurement results, CBLR was 0.18 when using an admission-control method similar to conventional methods without measurement. However, CBLR was much improved by this method; namely, it became 0.04. With this method, the overhead caused by NetFlow measurement and communication is very small.

To make the proposed method practical, the values of policy parameters, α, β, and γ, must be optimized. We used a simulated VoIP traffic with 40% gap for this measurement and obtained the following three ranges of the optimal parameter values.

First, the optimum value of smoothing coefficient α is probably between 0.5 to 1, but more study is required to confirm it. The oscillation and overshooting are suppressed by using an appropriate value of α.

Second, the optimum value of bandwidth margin ratio β is probably between 0.1 and 0.2. The bandwidth can be more effectively used if β is close to zero. However, to avoid packet loss caused by sudden increase of UBWR, it should be around 0.1 to 0.2. Moreover, the optimum value depends on requested loss ratio, delay, aggregation ratio, burstiness, and so on. It is not easy to find the optimum value if the traffic model is complex.

Third, the optimum value of measured-value contribution ratio γ is probably between 0.6 and 0.8. No measurement results are used if $\gamma = 0$, and the performance is worse if $\gamma > 0$. Oscillation, a type of harmful behavior of the admission-control system, may occur if $\gamma = 1$. If oscillation occurs, the CBLR periodically becomes close to 1 and possibly exceeds 1, so it should be avoided. More experiments and/or analyses are required to confirm these results.

It is noteworthy that not only the optimum value of α but also those of β and γ may be influenced by the nature of the traffic. Therefore, non-steady state traffic and types of traffic other than voice, e.g., video, should be tested.

In addition to the oscillation, another type of interesting but harmful behavior was also observed: UBWR may overshoot. However, more experiments are required to analyze overshooting.

Future work should solve the above problems, i.e., oscillation and overshooting, of RMBAC systems. However, the system should be kept simple because a more complicated system may cause more unexpected behavior.

Acknowledgments

The author thanks Daisuke Matsubara from the Central Research Laboratory, Hitachi, Ltd., and Masami Oshita, Ken'ichiro Hatae, Hajime Umemoto, and Fuminori Kimura from the Network Systems Solutions Division, Hitachi, Ltd., for discussing the method of admission control.

References

[Bol 94] Bolotin, V. A., "Telephone Circuit Holding Time Distributions", *14th Int'l Teletraffic Congress*, Antibes, France, North-Holland, 1994.

[Bra 65] Brady, P. T., "A Technique for Investigating On-off Patterns of Speech", *Bell System Technical Journal*, Vol. 44, pp. 1–22, January 1965.

[Bro 48] Brockmeyer, E., Halstrøm, H. L., and Jensen, A., "The Life and Works of A. K. Erlang", Danish Acad. Techn. Sci., No. 2, Köbenhavn (Denmark), 1948.

[Cal 03] Calhoun, P., Loughney, J., Guttman, E., Zorn, G., and Arkko, J., "Diameter Base Protocol", RFC 3588, IETF, September 2003.

[Cla 04] Claise, B., Ed., "Cisco Systems NetFlow Services Export Version 9", RFC 3954, IETF, October 2004.

[Dan 04] Trang Dinh Dang, Sonkoly, B., and Molnar, S., "Fractal Analysis and Modeling of VoIP Traffic", *11th Int'l Telecommunications Network Strategy and Planning Symposium* (*NETWORKS 2004*), pp. 123–130, June 2004.

[Geo 08] Georgoulas, S., Trimintzios, P., Pavlou, G., and Ho, K., "An Integrated Bandwidth Aloocation and Admission Control Framework for the Support of Heterogeneous Real-time Traffic in Class-based IP Networks", *Computer Communications*, Vol. 31, pp. 129–152, 2008.

[Hab 00] Habibi, D. and Mao, G., "A Hybrid ATM Connection Admission Control Scheme Based on On-Line Measurements and User Traffic Descriptors", *IEEE Int'l Conference on Networks* (*ICON 2000*), pp. 248–254, 2000.

[Kan 08] Kanada, Y., "Policy-based End-to-End QoS Guarantee Using On-Path Signaling for Both QoS Request and Feedback", *Int'l Conference on Information Networking 2008* (*ICOIN 2008*), I-1, January 2008.

[Lew 98] Lewis, J. T., Russell, R., Toomey, F., McGurk, B., Crosby, S., and Leslie, I. M., "Practical Connection Admission Control for ATM Networks Based on On-line Measurements", *Computer Communications*, Vol. 21, pp. 1585–1596, 1998.

[Nam 08] Seung Yeob Nam, Sunggon Kim, and Dan Keun Sung, "Measurement-Based Admission Control at Edge Routers", *IEEE/ACM Transactions on Networking*, Vol.16, No. 2, pp. 410–423, April 2008.

[Wen 00] Wenyu Jiang and Schulzrinne, H., "Analysis of On-off Patterns in VoIP and Their Effect on Voice Traffic Aggregation", *9th Int'l Conference on Computer Communications and Networks*, pp. 82–87, October 2000.

[Xu 04] Xu, Zhengdao, and Bochmann, G. v., "A Probabilistic Approach for Admission Control to Web Servers", *2004 Int'l Symposium on Performance Evaluation of Computer and Telecommunication Systems* (*SPECTS'04*), pp. 787–794, 2004.