マルチメディア通信と分散処理　103－12
（2001．6．8）

有効帯域幅と有効バッファを利用した複数クラスのトラヒックのための改良型コネクション受付制御

金 容珍 †　根元　義章 †
† 東北大学大学院情報科学研究科
〒980-8579　仙台市青葉区荒巻字青葉05
Tel.022-217-7140　E-mail:yjkim@nemoto.ecei.tohoku.ac.jp

あらまし　近年ネットワークの研究において、サービス品質の保証が大きな課題となっている。有効帯域幅という概念はコネクション受付制御と資源計画において、QoS を統計的に保証するための有効な方式として提案されている。コネクション受付制御は実時間トラヒック制御過程である。したがって、遅延を防止するために制御のための負荷は最小化されないといけない。同時に、できるだけ多くの利用者にサービスを提供するためにネットワーク資源を効率よく利用する必要がある。しかし、いままでは制御負荷を減らしながらネットワーク資源を効率的に利用するのは困難であった。そこで我々は、帯域幅とバッファの分離解析手法に基づいて、有効帯域幅と有効バッファという概念を新たに提案する。我々の手法を使うことによって、制御負荷を減らしながらネットワーク資源を効率的に利用することが出来ることを示す。

キーワード　サービス品質，コネクション受付制御，有効帯域幅

Improved Connection Admission Control by Effective Bandwidth and Effective Buffer with Multi-class Traffic

YongJin Kim †　Yoshiaki Nemoto †
† Graduate School of Information Sciences, Tohoku University.
Aza-Aoba05, Aramaki Aoba-ku, Sendai, 980-8579 Japan
Tel.022-217-7140
E-mail:yjkim@nemoto.ecei.tohoku.ac.jp

Abstract

Recently, the grarantee of QoS has become a hot issue in network research. Effective bandwidth has been proposed as a promising scheme for connection admission control and capacity planning to guarantee QoS statistically. Connection admission control is a real-time traffic control procedure. Hence, processing load should be minimized to prevent delay. At the same time, network resource efficiency should be enhanced to accomodate more users. However,so far, reducing process load and obtaining high network resource efficiency has been considered to be a contradictory matter. Here, we propose effective bandwidth and effective buffer concept based on bandwidth/buffer separation analysis scheme.We show that our method based on effective bandwidth and effective buffer concepts can achieve high network resource efficiency with reduced process load.

keywords　QoS, Connection Admission Control, Effective Bandwidth

1 Introduction

Major problems of network system design include how to decide whether a network accepts a new connection or not in real time and how to allocate network resources. A Connection Admission Control(CAC) function decides whether the network will accept a new connection or not. A new connection will be accepted if the network has enough resources to meet the QoS requiremets of a connection without damage to the QoS of the existing connections. A CAC is organized based on network resource requirement of connections.

To date, many CAC schemes have been proposed [1][2]. Each of the many CAC methods proposed so far has its own strengths and weaknesses. Generally, we can classsify existing methods whether the buffering effect is taken into account in evaluating performance level or not. Without considering buffering effect, CAC function could simply be made by allocating the peak rate bandwidth to the connection. A new call will be accepted if the sum of the peak rate of all existing connections and peak rate of a new connections is less than the capacity of the link. However, under that scheme, the network could not be utilized efficiently because each connection does not always transmit at its peak rate. On the other hand, in case buffering effect is considered, we need to model the queueig process. A queue will build up in the buffer as long as the total arrival rate exceeds the link capacity. When the queue length reaches buffer capacity, packet loss occurs. A strength of this method is that they can achieve high bandwidth efficiency. However, they require a lot of processing power and they are dependent on the input traffic model.

As we can see from above two CAC methods, there exists a contradictary problem between satisfying high bandwidth efficiency and keeping processing power low.In this paper, we propose bandwidth/buffer separation analysis scheme to solve such a contradictary problem. Based on bandwidth/buffer separation analysis scheme, we propose effective bandwidth and effective concepts. Through which, efficient CAC and bandwidth/buffer allocation can be performed satisfying declared QoS. We show that resource, bandwidth and buffer, can be utilized efficiently with light processing load.

It should also be noted that, so far, only bandwidth allocation is studied. Many methods so far are based on infinite buffer assumption. It is clear that bandwith and buffer shows trade-off relationship. That is to say, when there is enough bandwidth, the need of larger buffer is weaken. On the contrary, when there is enough buffer space, the need of larger bandwidth is weaken. The CAC and bandwidth/buffer allocation method that takes into consideration this trade-off relation is not adequately studied so far.Using our proposed scheme,this trade-off relation can be efficiently used.

The remainder of this paper is as follows. In section 2, we propose bandwiddth/buffer separation analysis scheme. In section 3, we proposed effective bandwidth and effective buffer concepts based on the proposed analysis scheme. In section 4, we present CAC and bandwidth/buffer allocation method for a multiplexing system. In section 5, we perform numerical study with extremal ON-OFF process to verify the accuracy of the proposed scheme. Finally, we present the conclusions in section 6.

2 Bandwidth/Buffer Separation Analysis Scheme

We propose buffer/bandwidth separation analysis scheme. Through which, difficult calculaion of loss probability of buffered model is reduced to two simple model.We begin with the characterization of the ith virtual circuit of class i by stationary random arrival process a_i that reprsents instantaneous loads. To establish a framework for extracting statistical multiplexing gains, we assume the statistical independence of traffic sources.

We consider a multiplexing system as two independent resources, channel and buffer. When arrival sources exhaust both channel resource and buffer resource, packet loss occurs. That is, loss probability is calculated as the production of two probability events,channel overflow and buffer overflow. Overall concept is depicted in Fig.1, where,

$$a_i(t) = u_i(t) + v_i'(t) \tag{1}$$

Here, we shall conservatively bound the process v_i by an ON-OFF process, which takes the peak value for the on period.

By the most basic probabilty principle, we can calcuate loss probability as following.

$$P_{loss} \approx Pr\{\sum_{i=1}^{I} K_i u_i > C\} Pr\{\sum_{i=1}^{I} K_i v_i > B\} \tag{2}$$

where K_i, is number of each source class. Using

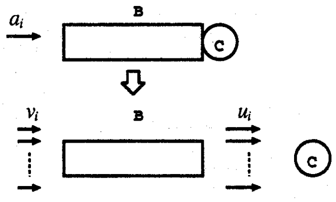

Figure 1: Separation Analysis

chernoff's bound,

$$P_{loss} \le e^{\tilde{F}_{\mathbf{K}}(s)} e^{\hat{F}_{\mathbf{K}}(s)} \tag{3}$$

where,

$$\begin{aligned} \tilde{F}_{\mathbf{K}}(s) &= \sum_{i=1}^{I} K_i \log \tilde{M}_i(s) - sC & (4) \\ \hat{F}_{\mathbf{K}}(s) &= \sum_{i=1}^{I} K_i \log \hat{M}_i(s) - sB & (5) \end{aligned}$$

where, $\tilde{M}_i(s), \hat{M}_i(s)$ is moment generating function of u_i and v_i.

By large deviation approximation, we can approximate loss probabilty as following.

$$P_{loss} \approx \inf_{s \ge 0}\{e^{\tilde{F}_{\mathbf{K}}(s)}\} \inf_{s \ge 0}\{e^{\hat{F}_{\mathbf{K}}(s)}\} \tag{6}$$

Following stability condition and loss occuring condition should be satisfied.

$$\sum_{i=1}^{I} K_i E(u_i) < C, \lim_{s \to \infty} \sum_{i=1}^{I} K_i \frac{\tilde{M}_i'(s)}{\tilde{M}_i(s)} < C \tag{7}$$

$$\sum_{i=1}^{I} K_i E(v_i) < B, \lim_{s \to \infty} \sum_{i=1}^{I} K_i \frac{\hat{M}_i'(s)}{\hat{M}_i(s)} < B \tag{8}$$

3 Effective Bandwidth and Effective Buffer Concepts

In this chapter, we propose effective bandwidth and effective buffer concepts that is based on queueing system analysis method proposed in the above section. So far, effective bandwidth concept has been proposed as a eminent CAC scheme. Under effective bandwidth approach, the amount of bandwidth required by a connection is estimated individually by the answer to the following question. What should the service rate be for a link with buffer capacity to achieve a declared QoS if the connection is input to this link?

With existing effective bandwidth concept, buffer capacity was fixed or ignored for calculation simplicity, which casues network resource inefficiency. Here, we propose effective bandwidth and effective buffer concepts. It works by answering following question. What should the service rate and buffer capacity be for a link to achieve a declared QoS if the connection is input to this link. By answing this question, the amount of bandwidth and buffer required by a connection is estimated individually.In addition, using effective bandwidth and effective buffer concept, we can find required amount of bandwidth and buffer considering the trade-off relation of bandwidth and buffer.

At first, the constraint on tail behavior $P\{Q \ge B\} \le L$, where Q is the queue length and B is the buffer size, will certainly be satisfied if,

$$P_{loss} \approx \inf_{s \ge 0}\{e^{\tilde{F}_{\mathbf{K}}(s)}\} \inf_{s \ge 0}\{e^{\hat{F}_{\mathbf{K}}(s)}\} \le L \tag{9}$$

In other word, it can be expressed as following.

$$\inf_{s \ge 0}\{\tilde{F}_{\mathbf{K}}(s)\} \inf_{s \ge 0}\{\hat{F}_{\mathbf{K}}(s)\} \le \log L \tag{10}$$

Here we can consider various combinations. For the best resource efficiency, we select constraints as following.

$$\begin{aligned} \inf_{s \ge 0}\{\tilde{F}_{\mathbf{K}}(s)\} &= L_1 & (11) \\ \inf_{s \ge 0}\{\hat{F}_{\mathbf{K}}(s)\} &= L_2 & (12) \end{aligned}$$

where,

$$L_1 + L_2 = \log L \tag{13}$$

Now, We use the value L_1 and L_2 to define effective bandwidth and effective buffer. That is to say, effective bandwidth is defined as the amount of bandwidth that satisfies the constraint of L_1 and effective buffer is defined as the amount of buffer that satisfies the constraint of L_2.

4 CAC and Resource Allocation for a Multiplexing System

We apply our system analysis method and consequent effective bandwidth and effective buffer concept to CAC and bandwidth/buffer allocation to a system. Our scheme can be applied to a variety of system models. Here, we apply it to a multiplexing system and use Chernoff bound.

Now, we are interested in the admissible set of each class. So far, admission control is organized through effective bandwidth concept. However, calculating effective bandwidth of buffered system has been very complicated work. Here, we propose connection admission control scheme using both effective bandwidth and effective buffer concepts that is proposed in the above section. Through which, both goal of bandwidth efficiency and reducing processing load can be achieved. As we explained above, when QoS requirement is given as L, it is satisfied if,

$$\inf_{s\geq 0}\{\tilde{F}_{\mathbf{K}}(s)\} + \inf_{s\geq 0}\{\hat{F}_{\mathbf{K}}(s)\} \leq \log L \qquad (14)$$

With the constraint of $\tilde{F}_{\mathbf{K}}(s) = L_1$, and $\hat{F}_{\mathbf{K}}(s) = L_2$, above equation is expressed as following.

$$\inf_{s\geq 0}\{\sum_{i=1}^{I} K_i\tilde{M}_i(s) - sC\} +$$

$$\inf_{s\geq 0}\{\sum_{i=1}^{I} K_i\hat{M}_i(s) - sB\} \leq L_1 + L_2 \qquad (15)$$

The acceptance region A, consisting of values $K = \{K_1, K_2, ..., K_I\} \in R^I_+$ satisfyig above condition, has a convex complement in R^I_+, since this complement is defined as the intersection of R^I_+ with a family of half space.The tangent plane at a point $\dot{K}$ on the boundary of the region A is,

$$\sum_{i=1}^{I} K_i\tilde{M}_i(\dot{s}) - \dot{s}C + \sum_{i=1}^{I} K_i\hat{M}_i(\ddot{s}) - \ddot{s}B = L_1 + L_2 \qquad (16)$$

where $\dot{s}$ and $\ddot{s}$ attains the infimum in (15) with K replaced by $\dot{K}$. One of the choice for the above condition is,

$$\sum_{i=1}^{I} K_i\tilde{M}_i(\dot{s}) - \dot{s}C = L_1 \qquad (17)$$

$$\sum_{i=1}^{I} K_i\hat{M}_i(\ddot{s}) - \ddot{s}B = L_2 \qquad (18)$$

It can also be expressed as following.

$$\sum_{i=1}^{I} K_i e_i = C \qquad (19)$$

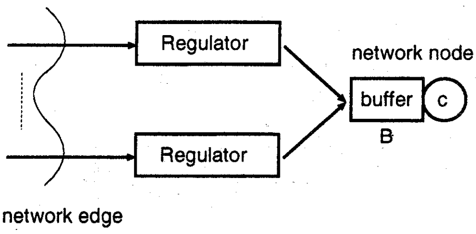

Figure 2: A multiplexing system model

$$\sum_{i=1}^{I} K_i b_i = B \qquad (20)$$

where e_j is effective bandwidth and b_j is effective buffer.

$$e_i = \frac{\tilde{M}_i(\dot{s})}{\dot{s} + L_1/C} \qquad (21)$$

$$b_i = \frac{\hat{M}_i(\ddot{s})}{\ddot{s} + L_2/B} \qquad (22)$$

Here, we can approximate,

$$\sum_{i=1}^{I} K_i e_i + \sum_{i=1}^{I} K_i b_i \approx B + C \qquad (23)$$

Accordingly, acceptance region can be expressed as following.

$$A_{L_1+L_2} = \{\mathbf{K} : \sum_{i=1}^{I} K_i\{e_i + b_i\} \leq B + C\} \qquad (24)$$

5 Numerical Study

We present numerical results to verify the accuracy of the proposed method. We consider a network node with multiple connections. Each connection is regulated by a leaky bucket regulator. We prepare a node that has a total amount of bandwidth C and buffer space B, shared by the sources. It is illustracted in figure 2.

To exploit the bursty nature of traffic sources regulated by leaky buckets, we assume that the regulated sources are extremal ON-OFF periodic processes after passing through leaky bucket regulators as illustrated in figure 3.

Here, an extremal ON-OFF periodic source is one which, when active, generates data at the peak rate P until the depletion of its token bucket; it then stays inactive until the token bucket is completely filled again. Such processes account for the

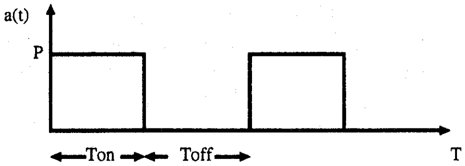

Figure 3: Traffic model

Table 1: Comparison with simulation

Loss Prob.	2.2×10^{-8}	1.7×10^{-7}	1.3×10^{-6}	6×10^{-6}
Simulation	150	160	170	180
Proposal's	148	157	165	172
Anwar's	110	119	129	137
Francesco's	116	123	134	142

worst-case statistical behavior in the sense that they maximize the average loss rate and the loss probability estimated by the Chernoff bound. T_{on} is the maximal time that a regulated source can send traffic at the peak rate P. This happens when the leaky bucket is full with σ amount of tokens.T_{off} is the time to fill an empty buffer with the token rate ρ. In order to model the statistical independence of traffic sources, we introduce indeterminate "phases" to sources as in [5]. Assume that traffic sources are grouped into I classes and for $1 \leq i \leq I$, there are K_i sources in class i. Each source of class i, is an extremal ON-OFF periodic process with leaky bucket parameters (ρ_i, σ_i, P_i) where ρ is token generation rate, σ is token bucket size and P_i is peak rate.In addition, each source of class i has an associate phase θ_i, that is,

$$a_i(t) = a_i(t + \theta_i) \tag{25}$$

As explained above section, loss probability can be estimated using the Chernoff bound. Then, with the equation mentioned so far, we can define following moment generating function for each source to channel.

$$\tilde{M}_i(s) = 1 - \omega_j + \omega_j e^{su_j} \tag{26}$$

Where ω is the probability for each state. The moment generating function for each source to buffer is expressed as a same way.

$$\hat{M}_i(s) = 1 - \omega_j + \omega_j e^{sv_j} \tag{27}$$

Using above equation, upper bound of loss probability can be obtained. In addition, effective bandwidth and effective buffer is obtained easily using above equation and proposed method in previous section.

Based on above equation, we obtained admission set for a given QoS criteria(loss probability) and it is shown in table 1.We also compared it with simulation result that is taken from[5] . In addition, we compared our result with results of others. The meaning of above result is that it is pretty close to simulation results compared with other existing methods such as the method proposed by Anwar[5] and the method proposed by Francesco[6] with light processing load. In addition, it shows conservative nature that is desirable for better QoS guarantee.

6 Conclusion

For network design, CAC and network resource allocation are the most important problems. Moreover, to guarantee required QoS statistically, sophisticated scheme is required. We proposed connection admission control and network resource allocation scheme using effective bandwidth and effective buffer concepts. Our scheme is based on bandwidth/buffer separation analysis scheme.

Through our scheme, connection admission control and network resource allocation can be organized with light processing load satisfying high bandwidth efficiency. In addition, our scheme give very clear concept on bandwidth and buffer tradeoff relation. In addition, Our schems gives simple method how bandwidth and buffer resource should be designed to satify required QoS. One more thing we have to mention is that our scheme does not constrict to regulated traffic. That is, our scheme can be applied for general traffic source model.

References

[1] Harry, G.P. and Khaled, M.E.,"Call Admission Control Schemes:A Review",IEEE Communication Magazine., November 1996, pp.82-91.

[2] Shiomoto, K., Tamanaka, N. and Takahashi, T.,"Overivew of Measurement-Based Connection Admission Control Methods in ATM Networks",IEEE communications surveys., First quarter 1999.

[3] Anwar, I.E. and Debasis, M.,"Effective Bandwidth of General Markovian Traffic Sources

and Admission Control of High Speed Networks",IEEE/ACM Trans. on Networking., vol.1,NO 3, June 1993, pp. 329-343.

[4] Arthur, W.B. and Ward, W.,"Effective Bandwidth with Priorities",IEEE/ACM Trans. on Networking., vol.6, NO 4, August 1998,pp. 447-460.

[5] Anwar, E., Debasis, M. and Robert, H.,"A New Approach for Allocating Buffers and Bandwidth to Heterogeneous, Regulated Traffic in an ATM Node", IEEE Trans. on Selected Areas in Communications., August 1995, pp. 1115-1127.

[6] Francesco,L.P.,Zhi-Li, Z. and Jim, K.,"Source Time Scale and Optimal Buffer/Bandwidth Tradeoff for Heterogeneous Regulated Traffic in a Network Node", IEEE/ACM Trans. on Networking., vol. 7,NO 4, August 1999,pp. 490-501

[7] Hluchyj, M.G. and Bhargava, A.,"Queueing Disciplines for Integrated Fast Packet Networks", Proc. IEEE ICC '92,. pp.335A.2.7, 1992.

[8] Shiomoto, K., and Yamanaka, N.,"A Simple Multi-QoS ATM Buffer Management Scheme Based on Apaptive Admission Control, Proc. IEEE GLOBECOMM '96, pp. 447-451, London, 1996.

[9] Oki, E. et al.,"A New Multiple QoS Control Scheme with Equivalent-window CAC in ATM Networks",IEICE Trans. Commun., vol.E81-B, no.7,pp.1462-1474,July 1998.

[10] Murase, T. et al.,"A Call Admission Control Scheme for ATM Networks Using a Simple Quality Estimate",IEEE J. Select. Areas Commun., vol.9,no 9,pp.1461-1470,Dec. 1991.

[11] Sohraby, K.,"On the Asymptotic Behavior of Heterogeneous Statistical Multiplexer with Applications, Proc. IEEE INFOCOM '92, pp.839-847,1992.

[12] Sohraby, K.,"On the Theory of General on-off Sources with Applications in High-speed Networks, Proc. IEEE INFOCOM '93, pp.401-410,1993.

[13] Kelly, F.P.,"Effective Bandwidths at Multi-type queues",Queueing Syst, vol.9,pp.5-15,1991.

[14] Guerin, R., Ahmadi,H.and Naghshineh, M., "Equivalent Capacity and its Application to Bandwidth Allocation in High-speed Networks",IEEE J. Select. Areas Commun.,vol.9,no.7,pp.968-981,Sept.1991.