Worst-Case Delay Control in Multi-Group Overlay Networks*

Wanqing Tu
Department of Computer Science, The Hong Kong University of Science and Technology
wqtu@ust.hk

Abstract

This paper proposes a novel and simple adaptive control algorithm for the effective delay control and resource utilization of EMcast when the traffic load becomes heavy in the multi-group network with real-time flows constrained by the (σ, ρ) regulators. The control algorithm is implemented at the overlay networks, and provides more regulations through a novel (σ, ρ, λ) regulator at each group end host who suffers from heavy input traffic. To our knowledge, it is the first work to incorporate the traffic regulator into the end host multicast to control the heavy traffic output. Our further contributions include a set of theoretical analysis and results. We prove the existence and calculate the value of rate threshold ρ^* such that for a given set of K groups, when the average rate of traffic entering the group end hosts $\bar{\rho} > \rho^*$, the ratio of worst-case multicast delay bound of (σ, ρ, λ) regulator over traditional (σ, ρ) regulator is $O(\frac{1}{K^n})$ for any integer n. We also use the computer simulations to evaluate our novel algorithm and regulator that have been proved efficient in decreasing the worst-case delays.

1 INTRODUCTION

End host multicast (EMcast) has emerged as an alternative to inter-domain IP multicast. A large number of end host multicast protocols [1-17] has been proposed since NARADA [1] demonstrated the feasibility of EMcast. Few of these protocols were designed for the multi-group networks. In the multigroup network, end hosts may join in several multicast groups. When one end host belongs to more than one group, the host is in face of processing multiple simultaneously entering flows. As the group communications usually generate real-time flows and the real-time flows are characterized with the high flow rates, the end hosts that join in multiple groups are apt to suffer from bottleneck that incurs unacceptable multicast delays and compromised scalability performances.
To release bottleneck, a popular way is to design the capacity-aware end host multicast protocol [7,16-17] that assigns the direct child members for each end host based on the end host output capacity. However, such bottleneck-avoidance performance is achieved at the cost of the increased lengths of multicast paths from the source to the group receivers (as illustrated in Figure 1). Suppose each flow in the multicast network has the uniform rate ρ and each end host has the same output capacity $C = 5\rho$. Figure 1 (a) gives the capacity-aware tree when all of the end hosts only join in one group

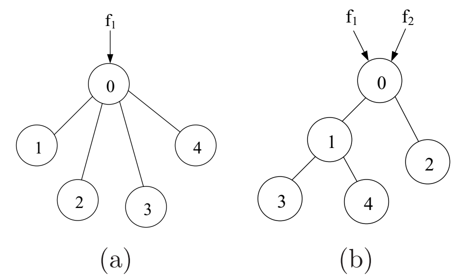


Figure 1. An example of constructing the capacity-aware multicast tree. (a) for one single-source group, (b) for two single-source groups.

G^1 in which only one transmission flow f_1 exists. In this case, each end host may have at most $\lfloor \frac{5\rho}{\rho} \rfloor = 5$ direct child members. End host 0 where the flow enters has the capacity to output packets to all of other end hosts 1, 2, 3 and 4 simultaneously. But when all of the end hosts join in two single-source groups G^1 and G^2, they may only connect to at most $\lfloor \frac{5\rho}{2\rho} \rfloor = 2$ child members directly. The reconstructed multicast tree is shown in Figure 1 (b). End host 0 will not forward packets to end hosts 3 and 4 who will receive the packets from end host 1 instead. It can be seen that the height of the multicast tree increases with the number of groups in which the end hosts join. Therefore, longer multicast delays are created. Such longer multicast delays are partially formed by the propagation delays of the underlying links that are newly added into the multicast system, and also include the delays caused by the way of packet transmission in EMcast where packets are forwarded by the end hosts and therefore experience the delays when they transmit between the IP layer and the application layer. (We analyze such delay in [18].) Moreover, the end hosts usually take longer latency to replicate and forward packets than the network routers because of the less capacities (e.g., CPU clock speed) of end hosts. Under the heavy network traffic load, the network transmission delays are already long. If the paths are increased, the unacceptable delay performances occur. Hence, instead of the capacity-aware scheme, a multicast traffic control mechanism that controls the output of simultaneously input flows at each end host is necessary for decreasing multicast delays when the network traffic becomes heavy.
Traffic control has been studied for the applications with various constraints in speed, quality and consistency of data delivery. But, not many mature researches have been done to address the multicast traffic control and all of these researches (e.g., Representative [21], RLA [22], TFMCC [23], MTCP [24] and Golestani [25]) are designed for IP multicast. Dur-

*Part of the work is finished when the author pursued her PhD degree in Department of Computer Science, City University of Hong Kong.

ing the network communication, different hosts have different instantaneous capacities. IP multicast distributes packets to the multicast address instead of each host's individual IP address. It is obvious that enabling IP multicast traffic control to recognize each individual host and its capacity incurs complexity. For example, TCP-friendly schemes excessively employ the benefit of TCP layer. EMcast holds the promise to implement traffic control in multicast communications. It is because packets are replicated and forwarded by end hosts and the control scheme can be carried out by software without other layers' support and expensive hardware.

Our motivation in this paper is to decrease the worst-case delay bound (WDB) of EMcast through traffic control when the traffic load becomes heavy in the multi-group network where the traffic is usually the high rate real-time flows constrained by (σ, ρ) regulators [19-20]. By worst-case delay, we refer to the longest packet delay at the end host who is the last one to receive the packets. There are two classical traffic control methods: the leaky-bucket mechanism [28-30] and the (σ, ρ) regulator. The leaky-bucket mechanism enforces a rigid output pattern at the average rate ρ not matter how bursty the input traffic is. For real-time applications, a more flexible mechanism is needed to process large burstiness that allows the short delay output, preferably one that does not lose data. Thus, we employ the (σ, ρ) regulator that introduces burstiness into the traffic model to analyze the worst-case delay bounds for real-time flows. In [19-20], the burstiness constraints of (σ, ρ) regulator for a given stream partially characterize the stream in the following way. Given any positive number ρ, there exists a (possible infinite) number σ such that if the stream is fed to a server that works at rate ρ while there is work to be done, the size of the backlog will never be larger than σ. Based on the regulator control method, we propose a novel and simple *adaptive control algorithm* that is implemented in the overlay network. The algorithm enables each host to adaptively employ the novel (σ, ρ, λ) regulator according to the instantaneous network situation. The communication bottleneck is released without increasing the lengths of multicast paths and requiring service reservation and control feedback. To our knowledge, it is the first work to incorporate traffic regulator into EMcast traffic control. Apart from the *adaptive control algorithm*, we present a set of theoretic analysis and results on the worst-case delay bound for the single regulated end host and the regulated EMcast respectively. More specifically, our contributions include

- The existence of the rate threshold ρ^* is proved such that $\hat{D} \geq D$ for the real-time flows with the average input rate $\bar{\rho} \leq \rho^*$ and $\hat{D} \leq D$ for $\bar{\rho} \geq \rho^*$, where $\hat{D}$ and D are the worst-case delay bounds of real-time flows constrained by the (σ, ρ, λ) and the (σ, ρ) regulators respectively;

- For a single regulated end host with K input flows, $\rho^* = 0.73C(0.79C)$ for the homogeneous (heterogeneous) flows, respectively, where C is the available output capacity of the multiplexer;

- For a single regulated end host with K input flows, the ratio of the worst-case delay bound of (σ, ρ, λ) regulator over (σ, ρ) regulator is $O(\frac{1}{K^n})$ for any positive integer n when $\bar{\rho} > \rho^*$;

- For a multicast group G with the size n, the height of DSCT tree [18] is upper bounded by $\lceil \log_k^{[k+(n-j_0)(k-1)]} \rceil$, where k (set as 3 in [11]) is a random positive integer, and $j_0 \in [0, k-1]$;

- For a multi-group network with K groups if each group $G^i (i \in [1, K])$ has n_i members that construct a DSCT tree, we derive $\rho^* = 0.73C(0.79C)$ for the homogeneous (heterogeneous) flows in the multi-group network. The ratio of the worst-case delay bound with (σ, ρ, λ) regulators over (σ, ρ) regulators is $O(\frac{1}{K^n})$ for any positive integer n when $\bar{\rho} > \rho^*$.

The rest of the paper is organized as follows. Section 2 introduces the *adaptive control algorithm.* Section 3 presents the theorems on the single regulated end host's *worst-case delay bound*, *input rate threshold* and *worst-case delay improvement.* The theoretic analysis for EMcast is presented in Section 4. Section 5 observes the WDB performances for single regulated end host and different EMcast schemes through simulations. Section 6 concludes the paper.

2 ADAPTIVE CONTROL ALGORITHM

Similar to [19-20], to control the input flows, each end host in the multi-group network is served with multiplexer (MUX) that merges the flows arriving at its two or more input links into the only one output link. In our traffic service disciplines, the general MUX is considered. A general MUX is such a MUX that a packet of one flow may have priority over a packet of another flow for transmission. We define an end host consisting of a MUX regulated by a $(\sigma, \rho, \lambda)/(\sigma, \rho)$ regulator on each of its input links as the $(\sigma, \rho, \lambda)/(\sigma, \rho)$-regulated end host.

Suppose there are K groups with $n_i (i \in [1, K])$ members each. Denote these groups as $G^i = \{g_1^i, ..., g_j^i, ..., g_n^i\}$, where g_j^i refers to the j-th ($j \in [1, n_i]$) member of G^i. Without loss of generality, we assume that only one real-time flow with the rate ρ_i (under the situation of no burstiness transmit) in each group. Therefore, there are totally K flows in the multi-group network. The basic idea of *adaptive control algorithm* is that each end host adaptively employs the same traffic control model as the (σ, ρ) regulator under the normal traffic load situation, but provides more regulations by using a new (σ, ρ, λ) regulator to control the traffic output in the overlay network under the heavy traffic load situation. The (σ, ρ, λ) regulator blocks the simultaneously entering flows a short period in turn when the average flow input rate is so large (larger than the end host rate threshold) that the end host is going to suffer from bottleneck if it outputs all received flows at the same time. The operations of *adaptive control algorithm* are given in Algorithm 1.

Algorithm 1 Adaptive Control Algorithm

Input: Multicast group $G^i = \{g_1^i, ..., g_j^i, ..., g_{n_i}^i\}$, and the input rate threshold ${\rho_j}^*$ of g_j^i who joins in K groups; // n_i is the size of G^i, $i \in [1, K]$, $j \in [1, n_i]$

Output: Traffic control model;

1. For $j = 1$ to n_i do {

2. End host g_j^i calculates the average input rate $\bar{\rho}_j$ of K real-time flows that come from K groups respectively;

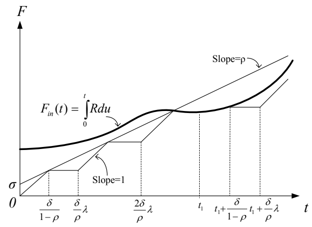


Figure 2. An example of the (σ, ρ, λ) regulator operations and $\lambda = \frac{1}{(1-\rho)}$.

3. If $(\bar{\rho_j} \in (0, \rho_j^*))$ {
   g_j^i employs the same traffic control model as the (σ_i, ρ_i) regulator;}
4. Else if $(\bar{\rho_j} \in [\rho_j^*, \frac{1}{K}))${
   g_j^i employs $(\sigma_i, \rho_i, \lambda_i)$ regulators to control the output of K input flows by two steps alternatively:
   (1) *On-state*: it works in a work-conserving way for $W_i = \frac{\sigma_i}{1-\rho_i}$ time units and then
   (2) *Off-state*: it takes a vacation of $V_i = \frac{\lambda_i \sigma_i}{\rho_i} - \frac{\sigma_i}{1-\rho_i}$ time units by turning off the input flow f_i. }}

It can be seen that the key problem of *adaptive control algorithm* is to find the input rate threshold ρ_j^* at which the algorithm should change the traffic control model. We will prove the existence and address the calculation of ρ_j^* through the theoretic analysis later. For brevity, in the theorems, we assume that each link in the network has a uniform available capacity $C = 1$. (Our theorems are also feasible if such assumption is released.) The inequality $\sum_{i=1}^{K} \rho_i \leq 1$ at each end host who joins in K groups is the *stability condition* of the multi-group network. For K homogeneous flows with the input rate ρ, the *stability condition* at each group member can be simplified as $K\rho \leq 1$.
Figure 2 shows the operations of (σ, ρ, λ) regulator. In order to smooth the simultaneous burstiness of K flows, the (σ, ρ, λ) regulator idles V time units after working each W. Once the duration of one flow's vacation expires, the regulator starts to serve the flow again. The period $(V + W)$ is called the *regulator period*. In our algorithm, we set it as $\frac{\sigma}{\rho}\lambda$. We will state the physical rationalness of the value of *regulator period* shortly later. λ is used to decide the vacation period V. We now focus on how λ is decided and its impact on the delay introduced by the regulator. Consider the i-th input flow at the end host g_j^i with the rate function $R_i \sim (\sigma_i, \rho_i)$, in order to guarantee that the total number of traffic output at g_j^i should be not greater than the total number of input traffic in the regulator during the period of m of *on* and $(m-1)$ of *off* states, we have $mW_i \leq \sigma_i + [mW_i + (m-1)V_i]\rho_i$. More specifically, we have $\frac{m\sigma_i}{1-\rho_i} \leq \sigma_i + (\frac{(m-1)\lambda_i\sigma_i}{\rho_i} + \frac{\sigma_i}{1-\rho_i})\rho_i$. That is, $\lambda_i \geq \frac{1}{1-\rho_i}$. Obviously, the smaller λ_i generates the shorter vacation period. To reduce the worst-case delay, we have

$$\lambda_i = \frac{1}{1-\rho_i} \tag{1}$$

Based on the *regulator period* $\frac{\sigma_i}{\rho_i}\lambda_i$, equation (1) infers $V_i = \frac{\sigma_i}{\rho_i}$ that shows how σ_i affects the vacation interval. Furthermore,

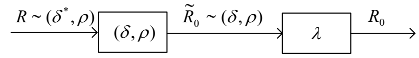


Figure 3. A concatenation of two network elements.

we consider K homogeneous flows. By the stability condition, we assume $\rho \to \frac{1}{K}$ in the worst case, then we have $V = \frac{\sigma}{\rho} \approx K\sigma = \frac{(K-1)\sigma}{(1-\frac{1}{K})} \approx (K-1)W$. It implies that when the input rate on each link is very high, the vacation interval of each regulator is nearly the same as the summation of working intervals of other $(K-1)$ regulators. Therefore, the introduction of *regulator period* and vacation has its physical rationalness.

3 Analysis of Worst-Case Delay Bound for The Single Regulated End Host

The following lemma characterizes the delay of any input flow with the rate function $R \sim (\sigma^*, \rho)$ at the (σ, ρ, λ)-regulated end host.
Lemma 1 *If the input flow's rate function R satisfies the burst constraint of (σ^*, ρ) regulator, i.e., $R \sim (\sigma^*, \rho)$, the delay incurred by the (σ, ρ, λ) regulator is upper bounded by*

$$D = \frac{(\sigma^* - \sigma)}{\rho} + \frac{2\lambda\sigma}{\rho}. \tag{2}$$

Proof. As shown in Figure 3, the (σ, ρ, λ) regulator is equivalent to the (σ, ρ) regulator and the λ controller.
In case $\sigma^* \leq \sigma$, there exists $\tilde{R_0} \sim (\sigma, \rho)$. Obviously, the largest backlog occurs at each end of a vacation. Without loss of generality, let $B(s)$ (s is an integer) denote the backlog of the regulator at time $\frac{s\lambda\sigma}{\rho}$ which is the end of a vacation. By the burst constraint of R, there is $B(0) \leq \sigma$. We may infer that $B(s) \leq (1+\lambda)\sigma$ for all $s \geq 0$ by induction on s. For simplicity, we denote the input flow rate bound without burstiness as ρ. At time $\frac{s\lambda\sigma}{\rho}$ the traffic arrived at the (σ, ρ, λ) regulator is $\lambda\sigma$ during the period of $\frac{\lambda\sigma}{\rho}$. On the other hand, the amount of traffic output by the regulator is at least $\lambda\sigma$. From this, and based on the induction assumption, we can infer that $B(s) \leq (1+\lambda)\sigma < 2\lambda\sigma$. Because $B(s)$ may be output by the regulator at the rate of ρ, the maximum delay could be as long as $\frac{2\lambda\sigma}{\rho}$.
In case $\sigma^* > \sigma$, because $\tilde{R_0} \sim (\sigma, \rho)$, it can be seen that the regulator may take some additional time to process the burst traffic $(\sigma^* - \sigma)$ originating from the input flow with the rate ρ. Therefore, the delay is $\frac{(\sigma^* - \sigma)}{\rho}$. Taking the two cases into consideration, we have the following delay bound for the (σ, ρ, λ) regulator $D = \frac{(\sigma^* - \sigma)^+}{\rho} + \frac{2\lambda\sigma}{\rho}$.Q.E.D.

3.1 Worst-Case Delay Bound

Theorem 1 *Let the rate function of the input flow f_i be $R_i \sim (\sigma_i, \rho_i)$, $1 \leq i \leq K$, and $\sigma_i^* = \rho_i(1-\rho_i). \min_{1 \leq j \leq K}\{\frac{\sigma_j}{\rho_j(1-\rho_j)}\}$, then the maximum delay experienced by a traffic bit in a general MUX with the $(\sigma_i^*, \rho_i, \lambda_i)$ regulator is upper bounded by $\hat{D}_g = \sum_{i=1}^{K} \frac{\sigma_i^*}{1-\rho_i} + 2\min_{1 \leq i \leq K}\{\frac{\sigma_i}{\rho_i(1-\rho_i)}\} + \max_{1 \leq i \leq K}\{\frac{\sigma_i - \sigma_i^*}{\rho_i}\}$.*

Proof. Without loss of generality, the delay experienced by any traffic bit of $f_j(j \in [1, K])$ is upper bounded by $\hat{D_g} \le D_1 + D_2$, where D_1 is the bit delay as it passes through the corresponding regulator, and D_2 is the delay bound of the multiplexer. By Lemma 1 and $\lambda_i = \frac{1}{1-\rho_i}$, there exists $D_1 \le \frac{2\lambda_i \sigma_i^*}{\rho_i} + \max_{1\le i\le K}\{\frac{\sigma_i - \sigma_i^*}{\rho_i}\} = 2\min_{1\le i\le K}\{\frac{\sigma_i}{\rho_i(1-\rho_i)}\} + \max_{1\le i\le K}\{\frac{\sigma_i-\sigma_i^*}{\rho_i}\}$. It can be seen that the amount of data bits from any flow f_i arriving at the multiplexer in any period of $\min_{1\le i\le K}\{\frac{\sigma_i}{\rho_i(1-\rho_i)}\}$ time units is upper bounded by $P^{(i)} = \frac{\sigma_i^*}{1-\rho_i}$, hence, the total amount of data bits arriving at the multiplexer in any period of $\min_{1\le i\le K}\{\frac{\sigma_i}{\rho_i(1-\rho_i)}\}$ time units is no more than $\sum_{i=1}^K P^{(i)} = \sum_{i=1}^K \frac{\sigma_i^*}{1-\rho_i}$.
Since the multiplexer is work-conserving with service rate $C = 1$, the above inequality means that each backlog at the multiplexer at any time is upper bounded by $D_2 = \sum_{i=1} K \frac{\sigma_i^*}{1-\rho_i}$. In other words, it is the upper bound on delay for any bit passing through the multiplexer. Thus, the theorem is proved. Q.E.D.
Theorem 2 gives the WDBs of K homogeneous real-time flows passing through the (σ, ρ, λ)-regulated general MUX.
Theorem 2 *For a regulated general MUX with K homogeneous input flows, let the input traffic rate functions be $R_i \sim (\sigma_0, \rho)$, $1 \le i \le K$, and $\rho \le \frac{1}{K}$. Then, the maximum delay experienced by any data bit in a (σ, ρ, λ)-regulated general MUX is upper bounded by $\hat{D}_g = \frac{K\sigma}{1-\rho} + \frac{(\sigma_0-\sigma)^+}{\rho} + \frac{2\lambda\sigma}{\rho}$.*
The proof of Theorem 2 is similar to the one of Theorem 1 and thus is omitted here.

3.2 Input Rate Threshold ρ^*

Now we are going to derive the rate threshold ρ^* for our *adaptive control algorithm* to distinguish the high rate real-time traffic from the normal rate traffic. We give the following notations,

$$\xi_{max} = \max_{1\le i\le K}\{\rho_i(1-\rho_i)\}, \xi_{min} = \min_{1\le i\le K}\{\rho_i(1-\rho_i)\},$$

$$\rho_{min} = \min_{1\le i\le K}\{\rho_i\}, \bar{\rho} = (\sum_{i=1}^K \rho_i)/K, \tag{3}$$

and introduce a condition that will be employed by the following inference

$$\frac{\xi_{max} - \xi_{min}}{\xi_{max}} \le \frac{\rho_{min}}{\bar{\rho}}. \tag{4}$$

Theorem 3 *Assume that a (σ, ρ, λ)-regulated MUX with the general service displine has K input links with the rate function for each link given by $R_i \sim (\sigma, \rho_i)$, $1 \le i \le K$, and $\sum_{i=1}^K \rho_i \le 1$. If $K \ge 2$ and condition (6) are satisfied, there exists a rate threshold $0 < \rho^* < \frac{1}{K}$ such that*
(i) if $\rho^* \le \bar{\rho} < \frac{1}{K}$, $\hat{D}_g \le D_g$, and if $0 < \bar{\rho} \le \rho^*$, $D_g \le \hat{D}_g$, where $\hat{D}_g(D_g)$ are the worst-case delay bounds of real-time flows constrained by the (σ, ρ, λ)-regulated ((σ, ρ)-regulated) general MUX respectively, and $\bar{\rho}$ is the average input rate of K flows;
(ii) when K is large enough, the ratio of the ranges $[\rho^*, \frac{1}{K})$ to the total range $(0, \frac{1}{K})$ is approximately given by $\frac{1/K-\rho^*}{1/K} \approx \frac{5-\sqrt{21}}{1} \approx 0.21$.

Proof. (i) By condition (6), D_g in Theorem 1 can be rewritten as $\hat{D}_g \le \frac{(\sum_{i=1}^K \rho_i)\sigma}{\xi_{max}} + \frac{2\sigma}{\xi_{max}} + \frac{\xi_{max}-\xi_{min}}{\xi_{max}} \frac{\sigma}{\rho_{min}}$.[1]
Noting that $h(x) = x(1-x)$ is an increasing function in the interval $[0, 1/K)$ when $K \ge 2$. For $\rho_i \in [0, 1/K)$, we have $\xi_{max} = \max_{1\le i\le K}\{\rho_i(1-\rho_i)\} \ge \bar{\rho}(1-\bar{\rho})$.
With (5), condition (6) and inequality (7), we have

$$\hat{D}_g = \frac{K\sigma}{1-\bar{\rho}} + \frac{2\sigma}{\bar{\rho}(1-\bar{\rho})} + \frac{\sigma}{\bar{\rho}}. \tag{5}$$

On the other hand, D_g in (3) can be represented as $D_g = \frac{K\sigma}{1-K\bar{\rho}}$.
Let

$$g_1(\bar{\rho}) = \frac{K}{1-\bar{\rho}} + \frac{2}{\bar{\rho}(1-\bar{\rho})} + \frac{1}{\bar{\rho}}, g_2(\bar{\rho}) = \frac{K}{1-K\bar{\rho}}. \tag{6}$$

Considering the equation $g_1'(\bar{\rho}) = \frac{K}{(1-\bar{\rho})^2} - \frac{2(1-2\bar{\rho})}{\bar{\rho}^2(1-\bar{\rho})^2} - \frac{1}{\bar{\rho}^2} = 0$,with the positive solution by $\bar{\rho_0} = \frac{-3+\sqrt{9+3(K-1)}}{K-1}$,it is clear that $\bar{\rho_0}$ is the minimum point of the function $g_1(\bar{\rho})$. Thus, the function $g_1(\bar{\rho})$ increases in $[\bar{\rho_0}, 1)$ such that $lim_{\bar{\rho}\to 1} g_1(\bar{\rho}) = +\infty$, and decreases in $(0, \bar{\rho_0}]$ such that $\lim_{\bar{\rho}\to 0} g_1(\bar{\rho}) = +\infty$. Since $g_2'(\bar{\rho}) \ge g_1'(\bar{\rho}), 0 < \bar{\rho} < \frac{1}{K}$,it can be inferred that the equation $g_1(\bar{\rho}) = g_2(\bar{\rho})$ has an unique positive solution ρ^* such that $0 < \rho^* < 1/K$. Consequently, $g_1(\bar{\rho}) \le g_2(\bar{\rho})$ when $\bar{\rho} \in [\rho^*, 1/K)$, and $g_1(\bar{\rho}) \ge g_2(\bar{\rho})$ when $\bar{\rho} \in (0, \rho^*]$. Thus (i) is proved.
(ii) By (i), ρ^* is the unique positive solution of the equation $g_1(\bar{\rho}) = g_2(\bar{\rho})$, which can be deduced to the equation $(K^2-2K)\bar{\rho}^2 + (3K+1)\bar{\rho} - 3 = 0$.Solving this equation, we have $\rho^* = \frac{-(3K+1)+\sqrt{(3K+1)^2+12(K^2-2K)}}{2(K^2-2K)}$.It is easy to see that $\lim_{K\to\infty} \frac{1/K-\rho^*}{1/K} = \lim_{K\to\infty}(1-K\rho^*) = \frac{5-\sqrt{21}}{2}$.Since it has been assumed that $C = 1$, thus (ii) holds. Q.E.D.
Theorem 4 gives the rate threshold ρ^* for K homogeneous flows entering single regulated end host.
Theorem 4 *Assume that a (σ, ρ, λ)-regulated MUX with the general service discipline has K input links with rate function for each link given by $R_i \sim (\sigma_0, \rho), 1 \le i \le K$ and $\rho \le 1/K$. When $K \ge 2$, there exists a $0 < \rho^* < 1/K$ such that*
(i) if $\rho^* \le \rho < \frac{1}{K}$, $\hat{D}_g \le D_g$, and if $0 < \rho \le \rho^*$, $D_g \le \hat{D}_g$, where $\hat{D}_g(D_g)$ is the worst-case delay bound of real-time flows constrained by the (σ, ρ, λ)-regulated ((σ, ρ)-regulated) general MUX respectively;
(ii) when K is large enough, the ratio of the range $[\rho^*, 1/K)$ with respect to the overall range $(0, 1/K)$ is about $\frac{\frac{1}{K}-\rho^*}{\frac{1}{K}} \approx 2-\sqrt{3} \approx 0.27$.
The proof of Theorem 4 can be similarly established as the proof of Theorem 3 and thus is omitted here.

[1] For each part of the expression of D_g in Theorem 1, we have

$$\sum_{i=1}^K \frac{\sigma_i^*}{1-\rho_i} = \sum_{i=1}^K \rho_i \min_{1\le j\le K}\{\frac{\sigma_j}{\rho_j(1-\rho_j)}\} \le \frac{(\sum_{i=1}^K \rho_i)\sigma}{\xi_{max}},$$

$$\min_{1\le i\le K}\{\frac{\sigma_i}{\rho_i(1-\rho_i)}\} \le 2\frac{\sigma}{\xi_{max}},$$

$$\max_{1\le i\le K}\{\frac{\sigma_i-\sigma_i^*}{\rho_i}\} = \max_{1\le i\le K}\{\frac{\sigma - \sigma\rho_i(1-\rho_i)\frac{1}{\xi_{max}}}{\rho_i}\}$$

$$\le \frac{\sigma(1-\frac{\xi_{min}}{\xi_{max}})}{\rho_{min}} = \frac{\xi_{max}-\xi_{min}}{\xi_{max}} \frac{\sigma}{\rho_{min}}.$$

.

3.3 Improvement of Worst-Case Delay Bound

We now analyze the WDB improvement of (σ, ρ, λ) regulator over (σ, ρ) regulator for the heterogeneous (Theorem 5) and the homogeneous (Theorem 6) real-time flows respectively.

Theorem 5 *Let the input rate functions be $R_i \sim (\sigma, \rho_i), (1 \leq i \leq K)$ with $\sum_{1\leq i\leq K} \rho_i \leq 1$, and D_g ($\hat{D}_g$) be the worst-case delay bounds for a general MUX regulated by the (σ_i, ρ_i) ($(\sigma_i, \rho_i, \lambda_i)$) regulators respectively. When the number of input links $K \geq 2$, for any positive integer n such that $\frac{1}{K} - \frac{1}{K^{n+1}} \geq \rho^*$, we have $\frac{D_g}{\hat{D}_g} \geq O(K^n)$, whenever $\bar{\rho} \in [\frac{1}{K} - \frac{1}{K^{(n+1)}}, \frac{1}{K})$.*

Proof. By Theorem 1, we know the worst-case delay of $(\sigma_i, \rho_i, \lambda_i)$ regulator is bounded by $\hat{D}_g = \sum_{i=1}^{K} \frac{\sigma_i^*}{1-\rho_i} + 2\min_{1\leq i\leq K}\{\frac{\sigma_i}{\rho_i(1-\rho_i)}\} + \max_{1\leq i\leq K}\{\frac{\sigma_i - \sigma_i^*}{\rho_i}\}$. By [19], the worst-case delay of (σ_i, ρ_i) regulator is bounded by $D_g = \frac{\sum_{1\leq i\leq K} \sigma_i}{1-\sum_{1\leq i\leq K} \rho_i}$. When K is large enough, Theorem 3 proves that $\rho^* \approx \frac{\sqrt{21}-3}{2K}$. Thus, when n is chosen properly, the inequality $\frac{1}{K} - \frac{1}{K^{n+1}} \geq \rho^*$ holds. Then, for any $\bar{\rho} \in [\frac{1}{K} - \frac{1}{K^{n+1}}, \frac{1}{K})$, it is easy to infer that $\bar{\rho} \in [\rho^*, \frac{1}{K})$, we have $\frac{D_g}{\hat{D}_g} \geq \frac{K\bar{\rho}(1-\bar{\rho})}{(1-K\bar{\rho})[3+(K-1)\bar{\rho}]} \geq \frac{(1-\frac{1}{K^n})(1-\frac{1}{K})K^n}{4} = O(K^n)$. Q.E.D.

For K homogeneous flows, we give the worst-case delay improvement in Theorem 6.

Theorem 6 *Let the input rate function R_i of homogeneous flows be the same as the above theorem, $D_g(\hat{D}_g)$ be the worst-case delay bounds for a general MUX regulated by the $(\sigma, \rho)((\sigma, \rho, \lambda))$ regulators, respectively. When the number of input links $K \geq 2$, there exists $0 < \rho^* < 1/K$, for any n such that $\frac{1}{K} - \frac{1}{K^{n+1}} \geq \rho^*$, we have $\frac{D_g}{\hat{D}_g} = O(K^n)$, whenever $\rho \in [\frac{1}{K} - \frac{1}{K^{n+1}}, \frac{1}{K})$.*

Theorem 6 can be proved in the similar way as we prove Theorem 5 and thus we omit it here.

4 Analysis of Worst-Case Delay Bound for The End Host Multicast

In our analysis, we use DSCT tree [18] as the model of EMcast architecture. Lemma 2 gives the height bound H of DSCT tree.

Lemma 2 *For a multicast group with n members, the height of the constructed DSCT tree is upper bounded by*

$$H = \lceil \log_k^{[k+(n-j_0)(k-1)]} \rceil, \tag{7}$$

where k is a random integer that is set as 3 in [11], $j_0 (0 \leq j_0 \leq k-1)$ is the number of the last unassigned members in the lowest layer L_1 of DSCT tree.

Proof. According to (1) and (2) in [18], the n members construct the highest DSCT tree when the size of each cluster equals to k.

Suppose DSCT tree has l layers. We use i_1 to denote the number of clusters with the size k in the lowest layer L_1, and j_1 to denote the remaining members who haven't joined in any of the i_1 clusters. It can be inferred that $i_1 = \lfloor \frac{n}{k} \rfloor$ and $0 \leq j_1 \leq k-1$. The j_1 members will form a new cluster in L_1. Hence, there are at most $(i_1 + 1)$ clusters in L_1. We have $n = i_1 k + j_1$. Because the core of each cluster joins in the immediate upper layer L_2, we can infer $i_1 + 1 = i_2 k + j_2$, where $i_2 = \lfloor \frac{i_1+1}{k} \rfloor$ is the number of clusters with the size k in L_2 and $j_2 \in [0, k-1]$ is the number of members who haven't joined in any of the i_2 clusters. Similarly, in the layer L_l, we can derive

$$i_{l-1} + 1 = i_l k + j_l, \tag{8}$$

where $i_l = \lfloor \frac{i_{l-1}+1}{k} \rfloor$ is the number of clusters with the size k and $j_l \in [0, k-1]$ is the number of members who haven't joined in the i_l clusters. Based on the above equations, using the iteration, we have

$$n = j_1 + i_1 k = j_1 + (j_2 - 1)k + i_2 k^2 = ... = j_1 + (j_2 - 1)k + (j_3 - 1)k^2 + ... + (j_l - 1)k^l + i_l k^{l+1}. \tag{9}$$

Because there is only one member in the highest layer L_l, we have $i_l = 0$ and $j_l = 1$. (9) shows that the tree has the highest height when $j_2 = j_3 = ... = j_{l-1} = 2$. Thus, we can infer from (9)

$$n = j_0 + k + k^2 + ... + k^{l-1} = j_0 + \frac{k - k^l}{1 - k}. \tag{10}$$

It can be achieved from (10) that $l = \lceil \log_k^{[k+(n-j_0)(k-1)]} \rceil$. Q.E.D.

Theorem 7 *Suppose there are K groups in the regulated multi-group network and each group has $n_i (i \in [1, K])$ end hosts that construct a DSCT tree. If one group has one real-time flow and the flow is constrained by the rate function $R_i \sim (\sigma_i, \rho_i)$, with the stability condition at each end host joining in K groups $\sum_{i=1}^{K} \rho_i < 1$, let $\lambda_i = \frac{1}{1-\rho_i}$, and $\sigma_i^* = \rho_i(1-\rho_i)\min_{1\leq j\leq K}\{\frac{\sigma_j}{\rho_j(1-\rho_j)}\}$,*
(i) the maximum multicast delays experienced by any bit passing through the multi-group network regulated with $(\sigma_i^*, \rho_i, \lambda_i)$ regulators are upper bounded by

$$\hat{D_{mg}} = (\hat{H} - 1)[\sum_{i=1}^{K} \frac{\sigma_i^*}{1-\rho_i} + 2\min_{1\leq i\leq K}\{\frac{\sigma_i}{\rho_i(1-\rho_i)}\} + \max_{1\leq i\leq K}\{\frac{(\sigma_i - \sigma_i^*)}{\rho_i}\}],$$

where $\hat{H} = \max_{1\leq i\leq K}\{H_i\}$ and H_i is the height bound of G_i;
(ii) if $K \geq 2$ and condition (6) are satisfied, there exists $0 < \rho^* < \frac{1}{K}$ such that $\hat{D_{mg}} \leq D_{mg}$ if $\rho^* \leq \rho < \frac{1}{K}$, and $D_{mg} \leq \hat{D_{mg}}$ if $0 < \rho \leq \rho^*$, where D_{mg} is the worst-case delay bound of DSCT with the (σ_i, ρ_i)-regulated general MUX;
(iii) when K is large enough, the ratio of the range $[\rho^*, \frac{1}{K})$ to $(0, \frac{1}{K})$ is approximately given by $\frac{\frac{1}{K}-\rho^*}{\frac{1}{K}} \approx \frac{5-\sqrt{21}}{2} \approx 0.21$;
(iv) for any positive integer n such that $\frac{1}{K} - \frac{1}{K^{n+1}} \geq \rho^*$, we have $\frac{D_{mg}}{\hat{D_{mg}}} \geq O(K^n)$, whenever $\bar{\rho} \in [\frac{1}{K} - \frac{1}{K^{n+1}}, \frac{1}{K})$.

Proof. (i) Suppose the longest multicast path (denoted as $< s^i \rightarrow r^i >$) in G^i is the one connecting the source s^i and the receiver r^i, where $s^i, r_i \in G^i$, $s^i \neq r^i$. Assume that the path contains F forwarders that are denoted as the set $\{\gamma_1^i, ..., \gamma_m^i, ..., \gamma_F^i\} (m \in [1, F])$ and $\gamma_m^i \in G^i$. The worst-case multicast delay in G^i with the $(\sigma_i^*, \rho_i, \lambda_i)$-regulated general MUX is the worst-case delay of any bit passing through $< s^i \rightarrow r^i >$ when s^i and all γ_m^i join in all the K groups. Then, the worst-case multicast delay bound $\hat{D^i_{mg}}$ in G^i is calculated by $\hat{D^i_{mg}} = \hat{D^i_g}(< s^i \rightarrow \gamma_1^i >) + \hat{D^i_g}(< \gamma_F^i \rightarrow r^i >) +$

$\sum_{m=1}^{F-1}\hat{D}_g^i(<\gamma_m^i\to\gamma_{m+1}^i>)$,where $\hat{D}_g^i(<s^i\to\gamma_1^i>)$, $\hat{D}_g^i(<\gamma_F^i\to r^i>)$ and $\sum_{m=1}^{F-1}\hat{D}_g^i(<\gamma_m^i\to\gamma_{m+1}^i>)$ refer to the worst-case delay bounds between s^i and γ_1^i, γ_F^i and r^i, and γ_m^i and γ_{m+1}^i respectively. According to Theorem 1, they equal to $\sum_{i=1}^{K}\frac{\sigma_i^*}{1-\rho_i}+2\min_{1\leq i\leq K}\{\frac{\sigma_i}{\rho_i(1-\rho_i)}\}+\max_{1\leq i\leq K}\{\frac{(\sigma_i-\sigma_i^*)}{\rho_i}\}$. Hence, the worst-case delay $\hat{D^i_{mg}}$ of any bit passing through the DSCT tree in G^i is $\hat{D^i_{mg}}=(H_i-1)[\sum_{i=1}K\frac{\sigma_i^*}{1-\rho_i}+2\min_{1\leq i\leq K}\{\frac{\sigma_i}{\rho_i(1-\rho_i)}\}+\max_{1\leq i\leq K}\{\frac{(\sigma_i-\sigma_i^*)}{\rho_i}\}]$,where H_i is the height bound of DSCT tree in G^i.

Considering the whole multi-group network, the worst-case multicast delay occurs in the group with the highest DSCT tree. We have $\hat{D_{mg}}=\max_{1\leq i\leq K}\{\hat{D^i_{mg}}\}=(\hat{H}-1)[\sum_{i=1}^{K}\frac{\sigma^*{}_i}{1-\rho_i}+2\min_{1\leq i\leq K}\{\frac{\sigma_i}{\rho_i(1-\rho_i)}\}+\max_{1\leq i\leq K}\{\frac{(\sigma-\sigma_i^*)}{\rho_i}\}]$,where $\hat{H}=\max_{1\leq i\leq K}\{H_i\}$. Q.E.D.

The proof of (ii), (iii) and (iv) can be similarly established as the proof of Theorems 3 and 5. Theorem 8 considers the homogeneous flows.

Theorem 8 *Suppose there are K groups denoted as $G^i(i\in[1,K])$ in the regulated multi-group network and each group has n_i end hosts that construct a DSCT tree. If each group has one real-time flow that is constrained by the rate function $R_i\sim(\sigma_i,\rho_i)$, with the stability condition $\rho\leq\frac{1}{K}$ at each end host joining in K groups,*

(i) the maximum worst-case delay experienced by any bit passing through the DSCT tree with the (σ,ρ,λ)-regulated general MUX is upper bounded by $\hat{D_{mg}}=\frac{(\hat{H}-1)K\sigma}{1-\rho}+\frac{(\hat{H}-1)(\sigma_0-\sigma)^+}{\rho}+\frac{2(\hat{H}-1)\lambda\sigma}{\rho}$,where $\hat{H}=\max_{1\leq i\leq K}\{H_i\}$ and H_i is the height bound of DSCT tree in G^i that can be derived by Lemma 2;

(ii) if $K\geq 2$ is satisfied, there exists a rate threshold $0<\rho^*<\frac{1}{K}$ such that $\hat{D_{mg}}\leq D_{mg}$ if $\rho^*\leq\bar{\rho}<\frac{1}{K}$, and $D_{mg}\leq\hat{D_{mg}}$ if $0<\bar{\rho}\leq\rho^*$, where D_{mg} is the worst-case delay bound of any bit passing through the DSCT tree with the (σ,ρ) regulator;

(iii) when K is large enough, the ratio of the range $[\rho^*,\frac{1}{K})$ to the total range $(0,\frac{1}{K})$ is approximately given by $\frac{\frac{1}{K}-\rho^*}{\frac{1}{K}}\approx 2-\sqrt{3}\approx 0.27$;

(iv) for any positive integer n such that $\frac{1}{K}-\frac{1}{K^{n+1}}\geq\rho^*$, we have $\frac{D_{mg}}{\hat{D_{mg}}}\geq O(K^n)$,whenever $\bar{\rho}\in[\frac{1}{K}-\frac{1}{K^{n+1}},\frac{1}{K})$.

5 Simulation Evaluation

We have done two groups of simulations in *ns*-2 [27] on a group of SUN SOLARIS workstations.

5.1 Simulation I

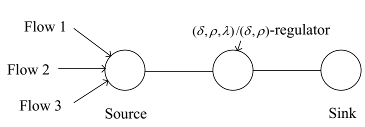
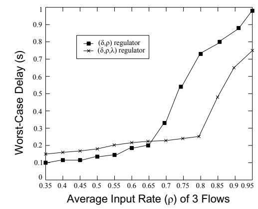
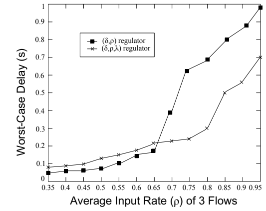


Figure 4. The simulation topology with only one $(\sigma,\rho,\lambda)/(\sigma,\rho)$-regulated end host.

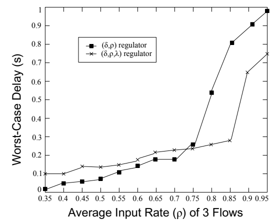
In the first group of simulations, we observe the WDB performances of single $(\sigma,\rho,\lambda)/(\sigma,\rho)$-regulated end host. Figure 4 shows the simulation topology. The source is fed with three real-time flows that are going to transmit to the sink. The intermediate node is equipped with the $(\sigma,\rho,\lambda)/(\sigma,\rho)$-regulated general MUXs respectively. Two types of real-time streams are employed: 64Kbps audio streams and 1.5Mbps MPEG-1 video streams. The regulator parameter σ is decided by the employed streams themselves. And, We compare the WDB performances of (σ,ρ,λ) regulator and (σ,ρ) regulator with 3 video streams, 3 audio streams and heterogeneous streams (one video and two audio streams) respectively when the average input rate ρ increases from 0.35 to 0.95.



(a)



(b)



(c)

Figure 5. The worst-case delay performances. (a) for three audio streams, (b) for three video streams and (c) for one video stream and two audio streams.

Figure 5 (a) illustrates the worst-case delay performances when three 64Kbps audio steams pass through the network in Figure 4. The cross point of the two curves is 0.66, i.e., the input rate threshold in this simulation is 0.66. When $\rho<0.66$, the worst-case delays with the (σ,ρ,λ) regulator are longer than the ones with the (σ,ρ) regulator. Otherwise, the worst-case delays with the (σ,ρ,λ) regulator are shorter than the ones with the (σ,ρ) regulator. The rate threshold difference between the simulation result and the theoretic analysis is because our theoretic analysis does not take into account of fluctuation of network throughput in the practical network. Also, it can be seen from the figure that the maximum worst-case delay improvement of (σ,ρ,λ) regulator over (σ,ρ) regulator is $\frac{0.72}{0.26}\approx 2.8$ when $\rho=0.8$. According to Theorem 6 and $K=3$, we can derive $n\approx 1$. Figure 5 (b) illustrates the WDB performances of 3 homogeneous video streams. The rate threshold is 0.67 that is a little less than the theoretic result 0.73 for the fluctuation of network throughput. The maximum improvement of worst-case delays of the (σ,ρ,λ) regulator over the (σ,ρ) regulator is $\frac{0.72}{0.26}\approx 2.82$ when $\rho=0.8$. With Theorem 6 and $K=3$, we can also derive $n\approx 1$. Figure 5 (c) gives the comparison of worst-case delay performance of heterogeneous real-time streams in the network. It can be seen that the input rate threshold is 0.74 that is a little less than the theoretic value 0.79 in Theorem 3. When $\rho\geq 0.74$, the worst-case delays with the (σ,ρ,λ) regulator are much shorter than the ones with the (σ,ρ) reg-

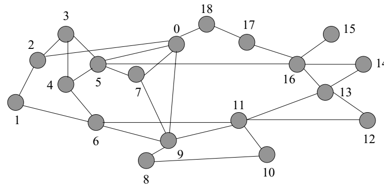


Figure 6. The backbone network topology in the simulations.

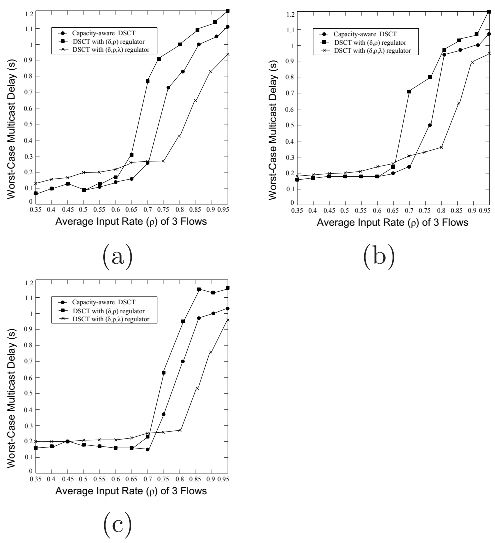


Figure 7. The worst-case delay performances when each of the three groups is fed with (a) the audio stream, (b) the video streams and (c) the video stream or the audio streams.

ulator. The maximum improvement of the worst-case delay is $\frac{0.85}{0.27} \approx 3.15$ when $\rho = 0.85$. Such improvement meets the theoretic results in Theorem 5 when $n = 1$.

5.2 Simulation II

In the second group of simulations, we observe the worst-case delay performances of real-time streams in the multi-group network. There are 665 end hosts in the network who join in 3 groups. Figure 6 shows the backbone network topology. The 665 group members are with the $(\sigma, \rho, \lambda)/(\sigma, \rho)$-regulated general MUXs and attach to the routers in the backbone network directly or indirectly through some intermediate network components (e.g., the hubs). Each group has one real-time flow. Namely, each host needs to serve 3 real-time flows. Also, there are two types of simulation streams: 64Mbps audio streams and 1.5Mbps MPEG-1 video streams in the multi-group network. And,we compare the WDB performances of three EMcast schemes: the *capacity-aware DSCT tree*, *DSCT tree with (σ, ρ) regulator* and *DSCT tree with (σ, ρ, λ) regulator*. The traffic pattern is the same as the first group of simulations.

Figure 7 (a) illustrates the worst-case delay performances of the *capacity-aware DSCT*, *DSCT with (σ, ρ) regulator* and *DSCT with (σ, ρ, λ) regulator* when each of the three groups is fed with the same 64Kbps audio stream. From the figure, we can see that the *capacity-aware DSCT* achieves shorter delay performances than *DSCT with (σ, ρ) regulator*. And, when $\rho \geq 0.7$, *DSCT with (σ, ρ, λ) regulator* achieves the best delay performances in the three multicast schemes. Compared to *DSCT with (σ, ρ) regulator*, the rate threshold in the simulation is 0.65 that is a little less than the theoretic value 0.73 in Theorem 8. And, the maximum improvement of the worst-case delays of *DSCT with (σ, ρ, λ) regulator* over *DSCT with (σ, ρ) regulator* is $\frac{0.95}{0.27} \approx 3.52$ when $\rho = 0.75$. It meets the theoretic results in Theorem 8 when $n = 1$. Figure 7 (b) shows the worst-case multicast delay performances of video streams. The *capacity-aware DSCT* achieves shorter delay performances than *DSCT with (σ, ρ) regulator*, and when $\rho \geq 0.7$, *DSCT with (σ, ρ, λ) regulator* achieves the shortest delay performances in the three multicast schemes. As for the comparison of *DSCT with (σ, ρ, λ) regulator* to *DSCT with (σ, ρ) regulator*, the simulation rate threshold of 3 flows is 0.65, and the maximum worst-case multicast delay improvement of *DSCT with (σ, ρ, λ) regulator* over *DSCT with (σ, ρ) regulator* is $\frac{1.18}{0.32} = 3.69$ when $\rho = 0.8$. Figure 7 (c) gives the worst-case delay performance comparison when one group is fed with the video stream and each of other two groups is fed with the audio stream. The simulation results also tell us that the rate threshold is 0.735 that is a little less than the theoretic result 0.79 in Theorem 7 because of the network throughput fluctuation in the practical network. And the maximum worst-case delay improvement of *DSCT with (σ, ρ, λ) regulator* over *DSCT with (σ, ρ) regulator* is $\frac{1.15}{0.27} \approx 4.26$ when $\rho = 0.8$.

6 Conclusion

In this paper, we addressed the problem of decreasing the worst-case delay bound for EMcast when the group members are in face of having no enough capacities to output the simultaneous input traffic. We presented a novel *adaptive control algorithm*. Based on the instantaneous network situations, the algorithm adaptively employs the (σ, ρ) regulator under the normal traffic load situation and the (σ, ρ, λ) regulator under the heavy traffic load situation to control the traffic output at each end host. The (σ, ρ, λ) regulator adopts two states: *on* and *off* to assign the output of the simultaneous heavy input flows in turn. By our new algorithm and regulator, proved by our theoretic analysis and simulation evaluation, the possible bottleneck can be avoided without increasing the lengths of multicast paths.

References

[1] Y. H. Chu, S. Rao, S. Seshan, and H. Zhang. A Case for End System Multicast. In Proc. of ACM SIGMETRICS 2000, page 1-12, June 17-21, 2000, Santa Clara, California, USA.

[2] P. FRANCIS. Yoid: Extending the Internet Multicast Architecture. Available at http://www.aciri.org/yoid/docs/index.html, April 2000.

[3] Y. Chawathe. Scattercast: An Architecture for Internet Broadcast Distribution as An Infrastructure Service. PhD thesis, University of California, Berkeley, August 2000.

[4] Y. H. Chu, S. G. Rao, S. Seshan, and H. Zhang. Enabling Conferencing Applications on The Internet Using An Overlay

Multicast Architecture. In Proc. of ACM SIGCOMM 2001, page 55-67, August 27-31, 2001, San Diego, CA.

[5] D. Pendarakis, S. Shi, D. Verma, and M.Waldvogel. ALMI: An Application Level Multicast Infrastructure. In Proc. of The 3rd Usenix Symposium on Internet Technologies and Systems (USITS01), page 49-60, March 2001, San Francisco, California.

[6] S. Zhuang, B. Zhao, A. D. Joseph, R. H. Katz, and J. Kubiatowicz. Bayeux: An Architecture for Wide-Area Fault-Tolerant Data Dissemination. In Proc. of NOSSDAV01, page 11-20, June 2001, Port Jefferson, New York, USA.

[7] S. Shi, and J. Turner. Routing in Overlay Multicast Networks. In Proc. of IEEE INFOCOM 2002, page 1200-1208, June 23- 27, 2002, New York, USA.

[8] S. Ratnasamy, M. Handley, R. Karp, and S. Shenker. Application-Level Multicast Using Content-Addressable Networks. In Proc. of Of The 3rd International Workshop on Network Group Communication, page 14-29, November 7-9, 2001, London, UK.

[9] S. Ratnasamy, P. Francis, M. Handley, and R. Karp. A Scalable Content- Addressable Network. In Proc. of ACM SIGCOMM 2001, page 161-172, August 27-31, 2001, San Diego, California, USA.

[10] B. Zhang, S. Jamin, and L. Zhang. Host Multicast: A Framework for Delivering Multicast to End Users. In Proc. of IEEE INFOCOM 2002, page 1366-1375, June 23-27, 2002, New York, USA.

[11] S. Banerjee, B. Bhattacharjee, and C. Kommareddy. Scalable Application Layer Multicast. In Proc. of ACM SIGCOMM, page 205-217, August 19-23, 2002, Pittsburgh, Pennsylvania, USA.

[12] M. Castro, P. Druschel, A. M. Kermarrec, and A. Rowstron. SCRIBE: A Large- Scale and Decentralized Application-Level Multicast Infrastructure. IEEE Journal on Selected Areas in Communications (JSAC), vol. 20, no. 8, page 1489-1499, October, 2002.

[13] A. Rowstron, and P. Druschel. Pastry: Scalable, Distributed Object Location and Routing for Large-Scale Peer-to-Peer Systems. Available at http://research.mecrosoft.com/antr/PAST/, 2001.

[14] J. Jannotti, D. K. Gifford, K. L. Johnson, M. Frans Kaashoek, and J. W. OToole Jr. Overcast: Reliable Multicasting with An Overlay Network. In Proc. of The 4th Usenix Symposium on Operating Systems Design and Implementation, page 194-210, October 22-25, 2000, Paradise Point Resort, San Diego, California, USA.

[15] H. Deshpande, M. Bawa, and H. Garcia-Molina. Streaming Live Media over Peers. In Proc. of First Workshop on Hot Topics in Networks (HotNets-I 2002), page 1-14, October 28-29, 2002, Princeton, New Jersey, USA.

[16] S. Banerjee, C. Kommareddy, B. B. K. Kar, and S. K huller. Construction of An Efficient Overlay Multicast Infrastructure for Real-Time Applications. In Proc. of IEEE INFOCOM 2003, page 1521-1531, March 30-April 3, 2003, San Francisco, USA.

[17] A. Riabov and L. Z. Zhen Liu. Overlay Multicast Trees of Minimal Delay. In Proc. of The 24th International Conference on Distributed Computing Systems (ICDCS 2004), page 654-661, March 23-26, 2004, Tokyo, Japan.

[18] W. Tu, and W. Jia. A Scalable and Efficient End Host Multicast for Peer-to- Peer Systems - DSCT. In Proc. of GlobeCom 2004, page 967-971, November 29 - December 3, 2004, Dallas, Texas, USA.

[19] R. Cruz. A Calculus for Network Delay, Part I: Network Elements in Isolation. IEEE Transaction on Information Theory, vol. 37, no. 1, page 114-131, January 1991.

[20] -. A Calculus for Network Delay, Part II: Network Analysis. IEEE Transaction on Information Theory, vol. 37, no. 1, page 131-141, January 1991.

[21] D. DeLucia, and K. Obraczka. Multicast Feedback Suppressionusing Representatives. In Proc. of IEEE INFOCOM 1997, page 463-470, April 7-12, 1997, Kobe, Japan.

[22] H. A. Wang, and M. Schwartz. Achieving Bounded Fairness for Multicast Traffic and TCP Traffic in The Internet. In Proc. of ACM SIGCOMM 1998, page 81-92, August 31 - September 4, 1998, Vancouver, British Columbia, Canada.

[23] J. Widmer, and M. Handley. Extending Equation-Based Congestion Control to Multicast Applications. In Proc. of ACM SIGCOMM 2001, August 27-31, 2001, San Diego, California, USA.

[24] I. Rhee, N. Ballaguru, and G. N. Rouskas. MTCP: Scalable TCP-like Congestion Control for Reliable Multicast. In Proc. of IEEE INFOCOM 1999, page 1265- 1273, March 21-25, 1999, New York, NY, USA.

[25] J. Golestani, and K. Sabnani. Fundamental Observations on Multicast Congestion Control in The Internet. In Proc. of IEEE INFOCOM 1999, page 990-1000, March 21-25, 1999, New York, USA.

[26] UC Berkeley, LBL, USC/ISI, and Xerox PARC. Ns Notes and Documentation. October 20, 1999.

[27] J. Turner. New Directions in Communications. IEEE Communication Magazine, vol. 24, no. 10, page 8-15, October 1986.

[28] M. Sidi, W. Z. Liu, I. Cidon, and I. Gopal. Congestion Control through Input Rate Regulation. In Proc. of IEEE Globecom 1989, page 1764-1768, November 27-30, 1989.

[29] A. E. Echberg, D. T. Luan, and D. M. Lucantoni. An Approach to Controlling Congestion in ATM Networks. International Journal of Digital Analog Communication System, vol. 3, page 199-209, 1990.

[30] W. Jia, H. Wang, M. Tang, and W. Zhao. Effective Delay Control for High Rate Heterogeneous Real-Time Flows. In Proc. of the 23rd IEEE International Conference on Distributed Computing Systems, page 367-375, May 19-22, 2003, Providence, Rhode Island, USA.