Worst-Case Delay Control in Multi-Group Overlay Networks

Wanqing Tu, *Member, IEEE,* Cormac J. Sreenan, *Fellow, BCS,*

Department of Computer Science,

University College Cork, Ireland

Email: {wt1,cjs}@cs.ucc.ie

Weijia Jia, *Member, IEEE*

Department of Computer Science

City University of Hong Kong

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 a multi-group network with real-time flows constrained by (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 traffic regulators into the end host multicast to control heavy traffic output. Our further contributions include theoretical analysis and a set of results. We prove the existence and calculate the value of the 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 the worst-case multicast delay bound of the proposed (σ, ρ, λ) regulator over the traditional (σ, ρ) regulator is $O(\frac{1}{K^n})$ for any integer n. We also prove the efficiency of the novel algorithm and regulator in decreasing worst-case delays by conducting computer simulations.

Index Terms

Worst-case delay control, overlay multicast, multiple groups, traffic control.

I. INTRODUCTION

End host multicast (EMcast) has emerged as an alternative to inter-domain IP multicast. A large number of end host multicast protocols [1-13] have been proposed since NARADA [1] demonstrated the feasibility of EMcast. Few of these protocols were designed for multi-group networks. In a multi-group network, end hosts may join in several multicast groups. When one end host belongs to more than one group, the end host has to process multiple simultaneously entering flows. As such and because the group flows are usually high rate real-time flows, the end hosts that join in multiple groups are prone to become bottlenecks, incurring unacceptable multicast delays and compromised scalability performance.

A popular way to free bottlenecks is to design capacity-aware end host multicast protocols [5,12-13] that assign the direct child members for each end host based on the end host output capacity. Thus, the end host has enough capacity to output the received packets to all its direct child members and will not become a communication bottleneck. However, such bottleneck-avoidance is achieved at the cost of increasing the lengths of the multicast paths from the source to the group receivers. As illustrated in Fig. 1, suppose each flow in the multicast network has the uniform rate ρ, and each end host has the same output capacity $C = 5\rho$. Fig. 1 (a) gives the capacity-aware tree when all the end hosts join in exactly one group, and in which only one transmission flow exists. In this case, each end host may have at most $\lfloor \frac{5\rho}{\rho} \rfloor = 5$ direct child members. Therefore, end host 0 (where the flow enters) has the capacity to output packets to all other end hosts 1, 2, 3 and 4 simultaneously. When the end hosts subscribe to two single-source groups, however, 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 Fig. 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 end host groups. Therefore, longer multicast delays are created. Such longer multicast delays are not only caused by the propagation and transmission delays of the newly-added underlying links,

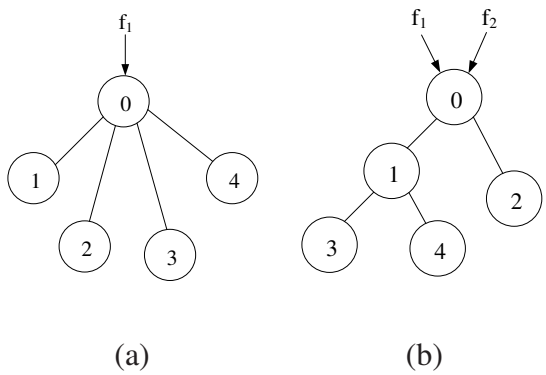


Fig. 1. An example of constructing the capacity-aware multicast tree. (a) for one single-source group, (b) for two single-source groups with the capacity-aware scheme.

but also include the delays caused by the way that packets transmit in EMcast. In EMcast, packets are forwarded by the end hosts and therefore experience delays when they transmit between the IP layer and the application layer. (We analyzed such delays in [14].) Moreover, end hosts usually take more time to replicate and forward packets than network routers because of end hosts' lesser capacities (e.g., CPU clock speed). Under heavy network traffic load, network transmission delays are already long. If path lengths are increased, unacceptable delay performance usually results. Hence, instead of the capacity-aware scheme, a multicast traffic control mechanism that does not increase the tree height is desirable.

There are two classical traffic control methods: the leaky-bucket mechanism [20-22] and the (σ, ρ) regulator. The leaky-bucket mechanism enforces a rigid output pattern at the average rate, irrespective of the burstiness of input traffic. For real-time applications, a more flexible mechanism is needed, allowing the processing of bursty flows within short delays, and preferably with no data loss. The (σ, ρ) regulator is such a mechanism that introduces burstiness into the traffic model. The burstiness constraints that the regulator considers for a given traffic stream partially characterize the stream in the following way. Given any positive number ρ, there exists a (possibly infinite) number σ such that if the traffic 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 [15-16] (we explain the physical meaning of σ and ρ in Section III). Our motivation in this paper is to decrease the worst-case delay bound (WDB) in multi-group EMcast networks by adopting a new

algorithm to control heavy traffic. We employ the (σ, ρ) regulator as the model to analyze the worst-case delay bounds of real-time flows. By the worst-case delay bound, we refer to the longest packet delay at the end host who is the last one in the group to receive the packets. Like tree stability and link stress, the worst-case delay bound is an important metric for EMcast. The WDB indicates whether all of the communication groups can achieve acceptable delay performance (i.e., the performance that meets the end-to-end delay bound requirements) or not. The decision to allow a new group to join the network is therefore based on the WDB. A shorter WDB improves the network's ability to host more groups.

We propose a novel and simple *adaptive control algorithm* that is implemented in the overlay network. Unlike capacity-aware EMcast protocols, our algorithm adaptively employs the novel (σ, ρ, λ) regulators to free bottlenecks without increasing the lengths of multicast paths (λ is a control parameter that will be introduced in Section III). With the proposed regulator, when the network traffic becomes heavy, the forwarding of flows at each end host is controlled in turn based on the current network state. To our knowledge, it is the first work to incorporate traffic regulators into EMcast. As well as the *adaptive control algorithm*, we present theoretical analysis and a set of results on the worst-case delay bound for a single regulated end host and a regulated EMcast network, respectively. Denote the worst-case delay bounds of the real-time flows constrained by the (σ, ρ, λ) and the (σ, ρ) regulators as $\hat{D}$ and D respectively, the average input rate of real-time flows as $\bar{\rho}$, and the end host's available output capacity as C. To be specific, our contributions include

- The existence of the rate threshold ρ^* is proved such that $\hat{D} \geq D$ when $\bar{\rho} \leq \rho^*$ and $\hat{D} \leq D$ when $\bar{\rho} \geq \rho^*$;
- For a single regulated end host with K input flows, $\rho^* = 0.73C$ ($\rho^* = 0.79C$) for the homogeneous (heterogeneous) flows, and the ratio of $\hat{D}$ over D when $\bar{\rho} > \rho^*$ is $O(\frac{1}{K^n})$ for both homogeneous and heterogeneous flows where n is any positive integer;
- For a multicast group G with the size n, the height of DSCT EMcast tree [14] is upper bounded by $\lceil \log_k^{[k+(n-j_1)(k-1)]} \rceil$, where k (set as 3 in [11]) is a random positive integer decided by the group size

and the application requirements, and $j_1 \in [0, k-1]$;

- For a multi-group network with K groups that are denoted as $G^I (I \in [1, K])$, if each group G^I has n_I members that construct a DSCT tree, we may derive $\rho^* = 0.73C$ ($\rho^* = 0.79C$) for the homogeneous (heterogeneous) flows in the multi-group network. The ratio of $\hat{D}$ over D when $\bar{\rho} > \rho^*$ is $O(\frac{1}{K^n})$ for both homogeneous and heterogeneous flows, where n is any positive integer.

The rest of the paper is organized as follows. Section II introduces the related work. Section III gives the *adaptive control algorithm* and describes the (σ, ρ, λ) regulator. Section IV presents the theorems for the *worst-case delay bound*, the *input rate threshold*, and the *worst-case delay improvement* of the single regulated end host. The theoretical analysis for EMcast is presented in Section V. Section VI uses the simulations to observe the WDB performance for a regulated end host and for different EMcast schemes. Section VII concludes the paper.

II. RELATED WORK

Traffic control has been studied for the applications with various constraints in speed, quality and consistency of data delivery, not many mature researches have been done to address the multicast traffic control and most of these researches are designed for the IP multicast.

a) Traffic control in IP multicast.

IP multicast usually employs open-loop or feedback traffic control systems. In an open-loop control system [31], a pre-determined control strategy that a session makes resource reservations ahead of time is fixed. The senders control their sending rate within the reservation, and do not response to changing network conditions. Open-loop control is difficult to implement because the Internet provides best effort services without service reservation. Most multicast traffic control schemes (e.g., Representative [17], RLA [18], MTCP [20] and Golestani [21]) are based on feedback control. In a feedback control system, the control parameter is adjusted on the fly. The control result reflecting the instantaneous network situations is measured and sent back to the associate node (e.g., sender) who will then adjust the transmission

accordingly. TCP-Friendly Rate Control (TFRC) [36] is a feedback control mechanism designed to compete with TCP traffic for bandwidth in unicast Internet environment. The TFRC receiver calculates the congestion control information (i.e., the loss rate), and feedbacks the information to the sender who then measures the round-trip time and gives the acceptable transmit rate. Such receiver-based mechanism enables TFRC to be easier extended to a multicast traffic control (TCP-Friendly Multicast Congestion Control (TFMCC) [19]). In TFMCC, each receiver measures the loss rate and its RTT to the sender and then decides a sending rate based on the equation for TCP throughput. The sender selects the receiver who reports the lowest rate as the current limiting receiver.

IP multicast distributes packets to the group's multicast address instead of each group member's individual IP address. During network communications, different hosts have different instantaneous capacities. Traffic control in IP multicast is usually complex in order to guarantee the communication qualities of end hosts with different capacities. For example, TFMCC excessively employs the benefit of TCP layer. Furthermore, to detect the current network state, considerable feedback overheads are introduced into the network.

b) Traffic control in EMcast.

Several end-to-end TCP-friendly multicast traffic control algorithms [28-29] have been studied for EMcast. J. Lu [31] proposed OverlayTFMRC focusing on QoS transport support for multimedia streaming and scalable TCP-friendly multicast congestion control. Although the *clustered-receiver-based* random delay strategy is employed, the scheme cannot remove the control overheads yet. Actually, many large scale practical systems (e.g., Overcast [10] and ALMI [4]) implement congestion control in the overlay path between each pair of nodes. These systems are actually implicity TCP-friendly where overlay paths are constructed through using TCP. Generally, because the sending sources employ the lowest sending rate calculated by the receivers [30], the throughput in the system with the TCP-friendly control scheme decreases with the increasing number of receivers. Therefore, in [31], the hop-by-hop control rather than the end-to-end technique is suggested to be implemented because of the end hosts' storing and forwarding functions in EMcast.

G. U. Keller [32] proposed a fixed-size window-based traffic control protocol for adjacent nodes. Each end host maintains a buffer for each of its outgoing interface. A fixed window control is applied to prevent the host from forwarding any packets to another host with the full-sized window. Other typical EMcast traffic control schemes include Cost-Benefit [33] and ROMA [34]. Basically, these schemes implement the quasi-hop-by-hop control and lack of the efficiency to support real-time streaming media. A single slow receiver can degrade the performance of the entire system. When it comes to the multiple group environment, the situation becomes worse because there is no coupled process for different group streams. Our proposed algorithm implements traffic control on a hop-by-hop basis. Without introducing feedback overheads, each host adaptively decides its control models based on the traffic input rates. Furthermore, each receiver implements traffic control based on its own capacity and therefore the slow receiver problem is solved.

III. ADAPTIVE CONTROL ALGORITHM

In this section, we introduce how the *adaptive control algorithm* works for traffic control in overlay networks. The model of the novel (σ, ρ, λ) regulator is presented. Also, the values of some parameters (e.g., λ and *regulator period*) are defined and calculated.

We first introduce the communication environment of the algorithm. Similar to [15-16], end hosts in the multi-group network are equipped with multiplexers (MUX) to control the input flows. The function of each MUX is to merge the flows arriving at its two or more input links into its single output link. In our traffic service algorithm, the general MUX is considered. A general MUX is a MUX such that a packet of one flow may have priority over a packet of another flow for transmission. For brevity, we define an end host consisting of a MUX that is regulated by a (σ, ρ, λ)/(σ, ρ) regulator on each of its input links as a (σ, ρ, λ)/(σ, ρ)-regulated end host. Suppose there are K groups with $n_I (I \in [1, K])$ end hosts each. Denote each group as G^I $(I \in [1, K])$. Without loss of generality, a member g_j^i in the multi-group network may join in $\hat{K}$ groups ($\hat{K} \in [1, K]$, $i \in [1, \hat{K}]$, $j \in [1, n_i]$, and n_i is the size of the i-th group that is

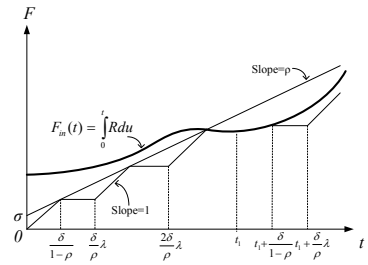


Fig. 2. An example of the operations of a (σ, ρ, λ) regulator. The x-axis represents the communication time and the y-axis represents the amount of data from the stream since the beginning of the communication. R is the traffic instantaneous input rate and $\lambda = \frac{1}{(1-\rho)}$ is the traffic control factor of the (σ, ρ, λ) regulator.

denoted as G^i). Assume that there is only one real-time flow in each group. We denote the i-th flow's burst data amount as σ_i and long term average input rate as ρ_i. Therefore, there are totally K flows in the multi-group network. For simpleness, in this paper, we assume that each link in the network has a uniform available capacity $C = 1$. When the assumption is released, the theorems and their proofs can be similarly developed by multiplying σ_i and ρ_i by C.) The inequality $\sum_{i=1}^{\hat{K}} \rho_i \leq 1$ at each end host g_j^i is regarded as the *stability condition* of the multi-group network. For $\hat{K}$ homogeneous flows with the input rate bound ρ, the *stability condition* at each group member can be simplified as $\hat{K}\rho \leq 1$.

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 new (σ, ρ, λ) regulators in the overlay network to control the traffic output under the heavy traffic load situation. Fig. 2 gives the operations of (σ, ρ, λ) regulator serving for one of the K flows. As illustrated by the zig-zag curve, the (σ, ρ, λ) regulator blocks the flow's output for V time units after outputting the flow for W time units. We call the V time units (i.e., the horizontal parts of the zig-zag curve) as the flow's vacation period, and the W time units (i.e., the sloping parts of the zig-zag curve) as the flow's working period. Other flows will experience the similar operations through their (σ, ρ, λ) regulators. In order to smooth the simultaneous burstiness of $\hat{K}$ flows, the *adaptive control algorithm* at each end host enables one regulator to work for its flow at each time in turn while other regulators block

their flows at the same time. The period of $(V+W)$ time units is defined as one *regulator period* of the flow and equals to $\frac{\sigma}{\rho}\lambda$. We will explain the physical rationalness of $\frac{\sigma}{\rho}\lambda$ later in this section. λ is a new parameter employed by our regulator. It decides the flow's vacation period and working period. We now see how to decide W_i, V_i and λ_i for the flow from the i-th group.

Denote the (σ,ρ)/(σ,ρ,λ) regulator of the i-th flow at g_j^i as the (σ_i,ρ_i)/$(\sigma_i,\rho_i,\lambda_i)$ regulator respectively, and the instantaneous input rate of the i-th flow at g_j^i as R_i. Similar to the (σ_i,ρ_i) regulator in [15-16], our $(\sigma_i,\rho_i,\lambda_i)$ regulator considers that $R_i \sim (\sigma_i,\rho_i)$. The term $R_i \sim (\sigma_i,\rho_i)$ holds when $\int_{t_1}^{t_2} R_i dt \leq \sigma_i + \rho_i(t_2-t_1)$ satisfies, where σ_i and ρ_i are the i-th flow's burst data amount and long term average input rate respectively, and t_1 (t_2) is the communication time with $t_2 \geq t_1$. $R_i \sim (\sigma_i,\rho_i)$ shows the physical meaning that the input amount of the i-th flow in any interval is upper bounded by the burst data amount plus the product of the long term average input rate and the length of the interval.

In order to guarantee that the total amount of traffic output at the end host g_j^i should be not greater than the total number of input traffic in the regulator during the period of m of W_i and $(m-1)$ of V_i, g_j^i's output should satisfy $mW_i \leq \sigma_i + [mW_i + (m-1)V_i]\rho_i$. According to Fig. 2, the cross points of the zig-zag curve and the trend line indicate the time that all of the blocked data from the flow are output by the (σ,ρ,λ) regulator. Furthermore, because all of the output capacity $C=1$ is occupied by the flow, the value of the slope of the (σ,ρ,λ) regulator curve is 1. Based on these analysis, we achieve that $W_i = \frac{\sigma_i}{(1-\rho_i)}$. It infers that $\frac{m\sigma_i}{1-\rho_i} \leq \sigma_i + [\frac{(m-1)\lambda_i\sigma_i}{\rho_i} + \frac{\sigma_i}{1-\rho_i}]\rho_i$. That is, $\lambda_i \geq \frac{1}{1-\rho_i}$. Since $V_i = \frac{\sigma_i}{\rho_i}\lambda_i - W_i$, a smaller λ_i generates a shorter vacation period. Therefore, considering of reducing the worst-case delay, we have

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

Equation (1) infers $V_i = \frac{\sigma_i}{\rho_i}$.

We now analyze the physical meaning of *regulator period*. For brevity, we suppose there are $\hat{K}$ homogeneous flows (i.e., $\rho_i = \rho$, $i \in [1,\hat{K}]$, $j \in [1,n_i]$). By the stability condition, we assume $\rho \to \frac{1}{\hat{K}}$ in the

worst case, then we have $V = \frac{\sigma}{\rho} \approx \hat{K}\sigma = \frac{(\hat{K}-1)\sigma}{(1-\frac{1}{\hat{K}})} \approx (\hat{K}-1)W$. 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 $(\hat{K}-1)$ regulators. Therefore, the introduction of *regulator period* and vacation has the physical rationalness. The detailed operations of *adaptive control algorithm* are given below.

Adaptive Control Algorithm

Input: The input rate threshold ρ_j^* of the member g_j^i who joins in $\hat{K}$ groups $G^i = \{g_1^i, ..., g_j^i, ..., g_{n_i}^i\}$;

// $i \in [1, \hat{K}]$, $j \in [1, n_i]$, n_i is the size of

and G^i

Output: Traffic control model;

1. End host g_j^i calculates the average input rate $\bar{\rho}$ of $\hat{K}$ real -time flows that belong to $\hat{K}$ groups respectively;

2. If $(\bar{\rho_j} \in (0, \rho_j^*))$ {

    g_j^i employs the same traffic control model as the (σ_i, ρ_i) regulator;}

3. Else if ($\bar{\rho_j} \in [\rho_j^*, \frac{1}{\hat{K}}))$\{

    g_j^i employs $(\sigma_i, \rho_i, \lambda_i)$ regulators to control the output of $\hat{K}$ input flows by the following 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 $V_i = \frac{\lambda_i \sigma_i}{\rho_i} - \frac{\lambda_i}{1-\rho_i}$ time units by turning off the input of the i-th flow at the end host g_j^i.}

The selection of the traffic control model is based on the flows' instantaneous input rates. The input rates indicate the instantaneous situations of underlying links. More specifically, the input rate $\rho_i (i \in [1, \hat{K}])$ at

an end host indicates the minimum instantaneous capacity amongest all underlying links that are covered by the overlay paths connecting the flow sender and the end host. The calculations of W_i and V_i are based on ρ_i and therefore based on the instantaneous capacity of underlying network links. The algorithm "intelligently" judges whether the end host is in face of congestion or not according to the average rate of all input flows. When the average rate is larger than the rate threshold ρ^* (i.e., the end host has no enough output capacity to work for all received flows at the same time), the algorithm "intelligently" blocks the simultaneous entering flows a short specific period in turn. It can be seen that the key problem of *adaptive control algorithm* is to find the input rate threshold ρ^* at which the algorithm should change the traffic control model. We will prove the existence and address the calculation of ρ^* through the theoretical analysis later.

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

We analyze the worst-case delay bound for the single regulated end host in this section. The results obtained will serve as an important basis of worst-case delay bound analysis as the packets pass through the EMcast tree.

The following lemma characterizes the delay of any input flow with the rate function of $R \sim (\sigma^*, \rho)$ at the (σ, ρ, λ)-regulated end host.

Lemma 1 *If the rate function R of input flow satisfies the burst constraint of (σ^*, ρ) regulator, i.e., $R \sim (\sigma^*, \rho)$, then 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. To prove the lemma, it is assumed that there exists $\tilde{R}_0$ that satisfies the traffic constraint of (σ, ρ) regulator, i.e., $\tilde{R}_0 \sim (\sigma, \rho)$. We now consider the following two cases.

In case $\sigma^* \leq \sigma$, 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 the

following 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 arriving at the (σ, ρ, λ) regulator is $\lambda\sigma$ during the period of $\frac{\lambda\sigma}{\rho}$. On the other hand, since $\sigma^* \leq \sigma$, at time $\frac{s\sigma}{1-\rho}$, the regulator can output the amount of traffic that equals to the amount of input traffic during the period $[\frac{(s-1)\sigma}{1-\rho}, \frac{(s-1)\sigma}{\rho}\lambda]$. That is to say, from the beginning of the communication to the time $\frac{s\lambda\sigma}{\rho}$, the maximum total backlog is $\lambda\sigma$ (i.e., the amount of traffic entered during the period $[\frac{s\sigma}{1-\rho}, \frac{s\lambda\sigma}{\rho}]$). Based on this, and considering of the induction assumption $B(0) \leq \sigma$, 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 $(\sigma^* - \sigma)/\rho$. 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.

A. Worst-Case Delay Bound

In the subsection, we present two theorems for the WDBs with K heterogeneous (Theorem 1) and homogeneous (Theorem 2) real-time flows respectively by applying Lemma 1 in the $(\sigma_i, \rho_i, \lambda_i)$-regulated general MUX, $1 \leq i \leq K$.

Theorem 1 *Let the rate function of the input flow f_i be given by R_i such that $R_i \sim (\sigma_i, \rho_i)$, $1 \leq i \leq K$, 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 from the flow $f_j (j \in [1, K])$ is upper bounded by $\hat{D}_g \leq D_1 + D_2$, where D_1 is the delay experienced by the bit passing 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 $D_1 \leq \frac{2\lambda_i\sigma_i^*}{\rho_i} + \max_{1\leq i\leq K}\{\frac{\sigma_i-\sigma_i^*}{\rho_i}\} = 2\min_{1\leq i\leq K}\{\frac{\sigma_i}{\rho_i(1-\rho_i)}\} + \max_{1\leq i\leq K}\{\frac{\sigma_i-\sigma_i^*}{\rho_i}\}$. It can be seen that the amount of data bits from any flow f_i arriving at the multiplexer in any period of $\min_{1\leq i\leq K}\{\frac{\sigma_i}{\rho_i(1-\rho_i)}\}$ 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 $\min_{1\leq i\leq K}\{\frac{\sigma_i}{\rho_i(1-\rho_i)}\}$ 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 such that $R_i \sim (\sigma_0, \rho)$, $1 \leq i \leq K$, and $\rho \leq \frac{1}{K}$. Then, the maximum delay experienced by any data bit in a (σ, ρ, λ)-regulated general MUX is upper bounded by*

$$\hat{D}_g = \frac{K\sigma}{1-\rho} + \frac{(\sigma_0 - \sigma)^+}{\rho} + \frac{2\lambda\sigma}{\rho}. \tag{3}$$

The proof of Theorem 2 is similar to that of Theorem 1 and thus is omitted here.

Remark 1 *By (13) in [15], if a general MUX has K heterogeneous (homogeneous) input flows and the rate function for each flow is given by $R_i \sim (\sigma_i, \rho_i)$ and $\sum_{1\leq i\leq K} \rho_i \leq 1$ (R_i such that $R_i \sim (\sigma_0, \rho)$ and $\rho \leq \frac{1}{K}$), the maximum delay in the general MUX is upper bounded by*

$$D_g = \frac{\sum_{1\leq i\leq K} \sigma_i}{1 - \sum_{1\leq i\leq K} \rho_i} (D_g = \frac{K\sigma_0}{1-K\rho}). \tag{4}$$

B. Input Rate Threshold ρ^*

Now we are going to derive the control 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.

$$\xi_{max} = \max_{1\le i\le K}\{\rho_i(1-\rho_i)\}, \xi_{min} = \min_{1\le i\le K}\{\rho_i(1-\rho_i)\}, \rho_{min} = \min_{1\le i\le K}\{\rho_i\}, \bar{\rho} = (\sum_{i=1}^{K}\rho_i)/K. \tag{5}$$

We then introduce a condition that will be employed by the following inference

$$\frac{\xi_{max}-\xi_{min}}{\xi_{max}} \le \frac{\rho_{min}}{\bar{\rho}}. \tag{6}$$

Theorem 3 *Assume that a $(\sigma_i^*, \rho_i, \lambda_i)$-regulated MUX with the general service discipline has K input links with the rate function for each link given by R_i such that $R_i \sim (\sigma_i, \rho_i)$, $1 \le i \le K$, and $\sum_{i=1}^{K}\rho_i \le 1$. If $K \ge 2$ and condition (6) are satisfied, then 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 $\hat{D}_g$ and D_g are the worst-case delay bounds of real-time flows constrained by the $(\sigma_i^*, \rho_i, \lambda_i)$-regulated general MUX and the (σ_i^*, ρ_i)-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 range (called the control range) $[\rho^*, \frac{1}{K})$ 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), for each part of the expression of D_g in Theorem 1, assume $\sigma = \min_{1\le i\le K}\{\sigma_i\}$, we have

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

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

$$\max_{1\le i\le K}\{\frac{\sigma_i-\sigma_i^*}{\rho_i}\} = \max_{1\le i\le K}\{\frac{\sigma_i-\sigma\rho_i(1-\rho_i)\frac{1}{\xi_{max}}}{\rho_i}\} \le \frac{\sigma[\frac{\sigma_i}{\sigma}-\frac{\xi_{min}}{\xi_{max}}]}{\rho_{min}} = \frac{\xi_{max}-\xi_{min}}{\xi_{max}}\frac{\sigma}{\rho_{min}} + \frac{\sigma_i-\sigma}{\rho_{min}}.$$

Then, D_g in Theorem 1 can be rewritten as

$$\hat{D}_g \le \frac{(\sum_{i=1}^{K}\rho_i)\sigma}{\xi_{max}} + \frac{2\sigma}{\xi_{max}} + \frac{\xi_{max}-\xi_{min}}{\xi_{max}}.\frac{\sigma}{\rho_{min}} + \frac{\sigma_i-\sigma}{\rho_{min}}. \tag{7}$$

Noting that $h(x) = x(1-x)$ is an increasing function in the interval $[0, 1/K)$ when $K \geq 2$, thus for $\rho_i \in [0, 1/K)$, we have $\xi_{max} = \max_{1 \leq i \leq K}\{\rho_i(1-\rho_i)\} \geq \bar{\rho}(1-\bar{\rho})$.

With (5), condition (6) and inequality (7), we have

$$\hat{D}_g = \frac{K\sigma}{1-\bar{\rho}} + \frac{2\sigma}{\bar{\rho}(1-\bar{\rho})} + \frac{\sigma}{\bar{\rho}} + \frac{1}{\rho_{min}}^{1}. \tag{8}$$

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

Let

$$g_1(\bar{\rho}) = \frac{K}{1-\bar{\rho}} + \frac{2}{\bar{\rho}(1-\bar{\rho})} + \frac{1}{\bar{\rho}}, g_2(\bar{\rho}) = \frac{K}{1-K\bar{\rho}}. \tag{9}$$

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 $g_2'(\bar{\rho}) \geq g_1'(\bar{\rho}), 0 < \bar{\rho} < \frac{1}{K}$, 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, $g_1(\bar{\rho}) \leq g_2(\bar{\rho})$ when $\bar{\rho} \in [\rho^*, 1/K)$, and $g_1(\bar{\rho}) \geq g_2(\bar{\rho})$ 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 following 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 the single regulated end host with K homogeneous flows.

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 such that $R_i \sim (\sigma_0, \rho), 1 \leq i \leq K$ and $\rho \leq 1/K$. When*

[1]By (7), with $\xi_{max} \geq \bar{\rho}(1-\bar{\rho})$ and $K\bar{\rho} = \sum_{i=1}^{K} \rho_i$, for the first part in (7), we have $\frac{(\sum_{i=1}^{K}\rho_i)\sigma}{\xi_{max}} \leq \frac{K\bar{\rho}\sigma}{\bar{\rho}(1-\bar{\rho})} = \frac{K\sigma}{1-\bar{\rho}}$; with $\xi_{max} \geq \bar{\rho}(1-\bar{\rho})$, for the second part in (7), we have $\frac{2\sigma}{\xi_{max}} \leq \frac{2\sigma}{\bar{\rho}(1-\bar{\rho})}$; with $\frac{\xi_{max}-\xi_{min}}{\xi_{max}} \leq \frac{\rho_{min}}{\bar{\rho}}$, for the third part in (7), we have $\frac{\xi_{max}-\xi_{min}}{\xi_{max}} \cdot \frac{\sigma}{\rho_{min}} \leq \frac{\sigma}{\bar{\rho}}$; with $\sigma_i \leq 1$, $\sigma \leq 1$ and $\sigma_i \geq \sigma$, for the fourth part in (7), we have $\frac{\sigma_i-\sigma}{\rho_{min}} \leq \frac{1}{\rho_{min}}$. Therefore, in the worst case, it can be inferred that $\hat{D}_g = \frac{K\sigma}{1-\bar{\rho}} + \frac{2\sigma}{\bar{\rho}(1-\bar{\rho})} + \frac{\sigma}{\bar{\rho}} + \frac{1}{\rho_{min}}$.

$K \geq 2$, there exists a rate threshold $0 < \rho^* < 1/K$ such that

(i) if $\rho^* \leq \rho < \frac{1}{K}$, $\hat{D}_g \leq D_g$, and if $0 < \rho \leq \rho^*$, $D_g \leq \hat{D}_g$, where $\hat{D}_g$ and D_g are the worst-case delay bounds of real-time flows constrained by the (σ, ρ, λ)-regulated general MUX and the (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 the proof of Theorem 3 and thus is omitted here.

C. 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. As we will see that the worst-case delay in $(\sigma_i, \rho_i, \lambda_i)$-regulated general MUX, $1 \leq i \leq K$, can be reduced effectively when the average rate $\bar{\rho}$ of K input flows is above the input rate threshold ρ^*.

Theorem 5 *Let the rate functions of input traffic be given by $R_i \sim (\sigma, \rho_i)(1 \leq i \leq K)$ with $\sum_{1 \leq i \leq K} \rho_i \leq 1$, and D_g and $\hat{D}_g$ be the worst-case delay bounds for a general MUX regulated by the (σ_i, ρ_i) and $(\sigma_i, \rho_i, \lambda_i)$ 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 and Remark 1, when the general MUX is regulated by the (σ_i, ρ_i) and $(\sigma_i, \rho_i, \lambda_i)$ regulators, the worst-case delay bounds are expressed as $D_g = \frac{\sum_{1 \leq i \leq K} \sigma_i}{1 - \sum_{1 \leq i \leq K} \rho_i}$ and $\hat{D}_g = \sum_{i=1}^{K} \frac{\sigma_i^*}{1-\rho_i} + 2 \min_{1 \leq i \leq K}\{\frac{\sigma_i}{\rho_i(1-\rho_i)}\} + \max_{1 \leq i \leq K}\{\frac{\sigma_i - \sigma_i^*}{\rho_i}\}$ respectively.

When K is large enough, by Theorem 3, we can prove 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 and $\hat{D}_g$ be the worst-case delay bounds for a general MUX regulated by the (σ, ρ) and the (σ, ρ, λ) regulators, respectively. When the number of input links $K \geq 2$, there exists a rate threshold $0 < \rho^* < 1/K$, 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.

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

Based on the above theorems, we achieve the theoretical results on the *worst-case delay bound*, the *input rate threshold* and the *worst-case delay bound improvement* for the regulated EMcast tree in this section. In our analysis, we use DSCT tree in [18] as the model of EMcast. DSCT arranges the end hosts in each group $G^I (I \in [1, K])$ to construct a DSCT multicast tree. DSCT tree is a location-aware hierarchy and cluster tree architecture. It partitions the group members into different *local domains*. Each *local domain* only contains the group members attaching to the same backbone routers. In terms of *round trip time value*, the closest s_{ina} group end hosts are assigned into the same "intra-cluster". As expressed by (1) in [14], the "intra-cluster" size s_{ina} is a random integer between k and $3k - 1$ if the number of unassigned members is greater than $3k - 1$; otherwise, s_{ina} is the number of *unassigned* group members. Each cluster has a cluster core that joins in the immediate upper layer and forms clusters in this layer with other $(s_{ina} - 1)$ closest cluster cores. Each *local domain* has a *local core* who is the end host in the upmost layer of the *local domain*. For the connections of different *local domains*, the closest *local cores* form "inter-clusters" with the size s_{ine} that is a random integer between k and $3k - 1$ as expressed by (2) in [14]. The local cores then continue constructing upper layers by the same way to layer the end hosts in each *local domain*. We first analyze the height bound H of DSCT tree in Lemma 2 when there are n members in the group.

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

$$H = \lceil \log_k^{[k+(n-j_1)(k-1)]} \rceil, \tag{10}$$

where k is a random integer that is decided by the group size and the application requirements (k is set as 3 in the computer experiments [8]), $j_1 (0 \leq j_1 \leq k-1)$ is the number of the last unassigned members in the lowest layer L_1 of DSCT tree.

Proof. According to (1) and (2) in [14], the n members will construct the highest DSCT tree when the sizes of all clusters equal 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, the following equation can be inferred $i_1 + 1 = i_2 k + j_2$, 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 the following equation

$$i_{l-1} + 1 = i_l k + j_l, \tag{11}$$

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

$$\begin{aligned} n &= j_1 + i_1 k = j_1 + (j_2 - 1 + i_2 k)k = j_1 + (j_2 - 1)k + i_2 k^2 = ... \\ &= j_1 + (j_2 - 1)k + (j_3 - 1)k^2 + ... + (j_l - 1)k^l + i_l k^{l+1}. \end{aligned} \tag{12}$$

Because there is only one member in the highest layer L_l, we have $i_l = 0$ and $j_l = 1$. And, (12) shows that the tree will have the maximum layer number when $j_2 = j_3 = ... = j_{l-1} = 2$. Thus, we can infer from (12)

$$n = j_0 + k + k^2 + ... + k^{l-1} = j_0 + \frac{k - k^l}{1 - k}. \tag{13}$$

It can be achieved from (13) that $l = \lceil \log_k^{[k+(n-j_0)(k-1)]} \rceil$. In other words, the height of DSCT tree that covers n members is upper bounded by $H = \lceil \log_k^{[k+(n-j_0)(k-1)]} \rceil$. Q.E.D.

By applying Lemma 2, we analyze the worst-case delay bounds of EMcast with K heterogeneous flows and K homogeneous flows in Theorem 7 and Theorem 8 respectively.

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 such that $R_i \sim (\sigma_i, \rho_i)$, stability condition $\sum_{i=1}^{\hat{K}} \rho_i \leq 1$* at each end host who joins in $\hat{K} (\hat{K} \in [1, K])$ groups, let $\lambda_i = \frac{1}{1-\rho_i}$, and $\sigma_i^* = \rho_i (1 - \rho_i) \min_{1 \leq j \leq \hat{K}} \{ \frac{\sigma_j}{\rho_j (1-\rho_j)} \}$, then,

(i) the maximum multicast delays experienced by any bit passing through the multi-group network with the $(\sigma_i^*, \rho_i, \lambda_i)$-regulated general MUXs are upper bounded by $\hat{D_{mg}} = \sum_{i=1}^{K} \frac{(\hat{H}-1)\sigma_i^*}{1-\rho_i} + 2 \min_{1 \leq i \leq K} \{ \frac{(\hat{H}-1)\sigma_i}{\rho_i (1-\rho_i)} \} + \max_{1 \leq i \leq K} \{ \frac{(\hat{H}-1)(\sigma_i - \sigma_i^*)}{\rho_i} \}$, where $\hat{H} = \max_{1 \leq i \leq K} \{ H_i \}$ and H_i is the height bound of DSCT tree in the group G^i that can be derived by Lemma 2;

(ii) if $K \geq 2$ and condition (6) are satisfied, then there exists a rate threshold $0 < \rho^* < \frac{1}{K}$ such that $\hat{D_{mg}} \leq D_{mg}$ if $\rho^* \leq \rho < \frac{1}{K}$, and $D_{mg} \leq \hat{D_{mg}}$ if $0 \leq \rho \leq \rho^*$, where D_{mg} is the worst-case delay bound of DSCT with the (σ_i, ρ_i)-regulated general MUX and we give its value in Remark 2;

(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 $\frac{\frac{1}{K} - \rho^*}{\frac{1}{K}} \approx \frac{5-\sqrt{21}}{2} \approx 0.21$;

(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 there are F forwarders on the path

$< s^i \rightarrow r^i >$ that are denoted as the set of $\{\gamma_1^i, ..., \gamma_m^i, ..., \gamma_F^i\}(m \in [1, F])$ 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 $\hat{D_{mg}^i}$ in G^i is calculated by $\hat{D_{mg}^i} = \hat{D_g^i}(< s^i \rightarrow \gamma_1^i >) + \hat{D_g^i}(< \gamma_F^i \rightarrow r^i >) + \sum_{m=1}^{F-1} \hat{D_g^i}(< \gamma_m^i \rightarrow \gamma_{m+1}^i >)$, where $\hat{D_g^i}(< s^i \rightarrow \gamma_1^i >)$, $\hat{D_g^i}(< \gamma_F^i \rightarrow r^i >)$ and $\sum_{m=1}^{F-1} \hat{D_g^i}(< \gamma_m^i \rightarrow \gamma_{m+1}^i >)$ 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 $\sum_{i=1}^{K} \frac{\sigma_i^*}{1-\rho_i} + 2\min_{1\le i\le K}\{\frac{\sigma_i}{\rho_i(1-\rho_i)}\} + \max_{1\le i\le K}\{\frac{(\sigma_i-\sigma_i^*)}{\rho_i}\}$. Hence, the worst-case delay $\hat{D_{mg}^i}$ of any bit passing through the DSCT tree in G^i is $\hat{D_{mg}^i} = (H_i - 1)[\sum_{i=1} K\frac{\sigma_i^*}{1-\rho_i} + 2\min_{1\le i\le K}\{\frac{\sigma_i}{\rho_i(1-\rho_i)}\} + \max_{1\le i\le K}\{\frac{(\sigma_i-\sigma_i^*)}{\rho_i}\}]$, where H_i is the height bound of DSCT tree in G^i. Actually, $H_i = m + 1$, where m is the number of forwarders in the longest path $< s^i \rightarrow r^i >$.

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

The proof of (ii), (iii) and (iv) can be similarly established as the proof of Theorems 3 and 5 and thus is omitted here. For the homogeneous flows in the multi-group network, we present Theorem 8.

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 one real-time flow exists in each group and the flow is constrained by the rate function R_i such that $R_i \sim (\sigma_0, \rho)$ with the stability condition $\rho \le \frac{1}{\hat{K}}(\hat{K} \in [1, K])$ at each end host that joins in $\hat{K}$ groups, then,*

(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 $\hat{H} = \max_{1\le i\le K}\{H_i\}$ and H_i is the height bound of DSCT tree in G^i that can be derived by Lemma 2;

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

(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 $\bar{\rho} \in [\frac{1}{K} - \frac{1}{K^{n+1}}, \frac{1}{K})$.

Remark 2 *By (13) in [15], if any bit passing through the network with K groups that are regulated by the general MUXs. The general MUX at each end host has $\hat{K}$ input links (i.e., the end host joins in $\hat{K}$ groups) and the rate functions for the flows in the $\hat{K}$ groups are given by R_i such that $R_i \sim (\sigma_i, \rho_i)$ ($R_i \sim (\sigma_0, \rho)$), $i \in [1, \hat{K}]$, with the stability condition $\sum_{i=1}^{\hat{K}} \rho_i \leq 1$ ($\rho \leq \frac{1}{\hat{K}}$) at each end host that joins in $\hat{K}(\hat{K} \in [1, K])$ groups, the maximum delay of the data bit is upper bounded by $D_{mg} = \frac{(\hat{H}-1)\sum_{i=1}^{K}\sigma_i}{1-\sum_{i=1}^{K}\rho_i}(D_{mg} = \frac{(\hat{H}-1)K\sigma_0}{1-K\rho})$ where $\hat{H}$ is the maximum value of the DSCT tree height bounds of K groups.*

VI. Simulation Evaluation

In this section, we use the simulations to evaluate the worst-case delay bounds of network communications with and without our *adaptive control algorithm* respectively. We have done two groups of simulations in ns-2 [22] and run them on a group of SUN SOLARIS workstations.

A. Simulation I

In first group simulations, we observe the WDB performances of single (σ, ρ, λ)/(σ, ρ)-regulated end host. Fig. 3 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 (σ, ρ, λ)/(σ, ρ)-regulated general MUXs respectively. Two types of real-time streams are employed: 64Kbps audio streams and 1.5Mbps MPEG-1 video streams. 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.

Fig. 4 (a) illustrates the worst-case delay performances when there are three 64Kbps audio steams pass through the network in Fig. 3. The simulation results meet our theoretical analysis. The cross point of the

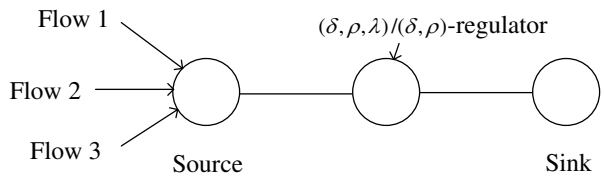
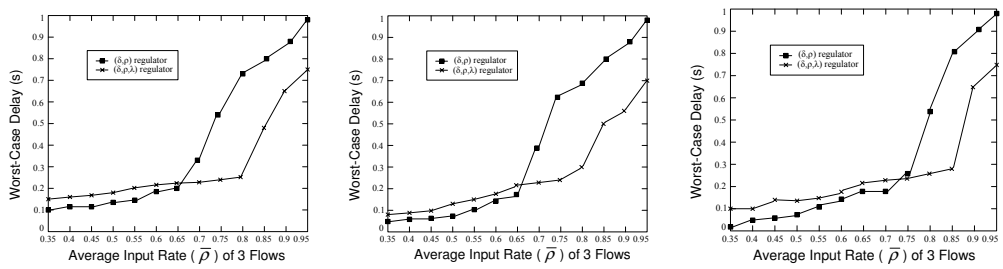


Fig. 3. The simulation topology with only one $(\sigma, \rho, \lambda)/(\sigma, \rho)$-regulated end host.



(a) (b) (c)

Fig. 4. The worst-case delay performances when there are (a) three 64Kbps audio streams, (b) three 1.5Mbps video streams and (c) one 1.5Mbps video stream and two 64Kbps audio streams in the network of Fig. 4.

two curves is 0.66, i.e., the input rate threshold in this simulation is 0.66. When $\bar{\rho} < 0.66$, the worst-case delays of packets passing through the (σ, ρ, λ) regulator are longer than the worst-case delays of packets passing through 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 theoretical analysis) is because our theoretical analysis does not take into account of the fluctuation of network throughput in the simulation. The throughput fluctuation is mainly caused by the following reasons: a) the cluster size is a random integer between k and $3k-1$ which possibly makes the same end host have different child members in different multicast schemes; b) the audio and video streams in the simulation are all variable bit rate (VBR) flows. The transmission rates of VBR flows are changing over time which causes the fluctuation of throughput. When the number of VBR flows increases, the fluctuation of network throughput becomes large. Also, it can be seen from the figure when $\bar{\rho} \geq 0.66$, the maximum worst-case delay improvement of (σ, ρ, λ) regulator over (σ, ρ) regulator is at $\bar{\rho} = 0.8$ and has the value of $\frac{0.72}{0.26} \approx 2.8$. According to Theorem 6, we can derive $n \approx 1$ from the simulation parameter $K = 3$.

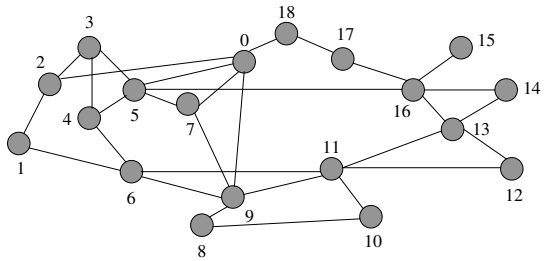


Fig. 5. The backbone network topology in the simulations.

TABLE I

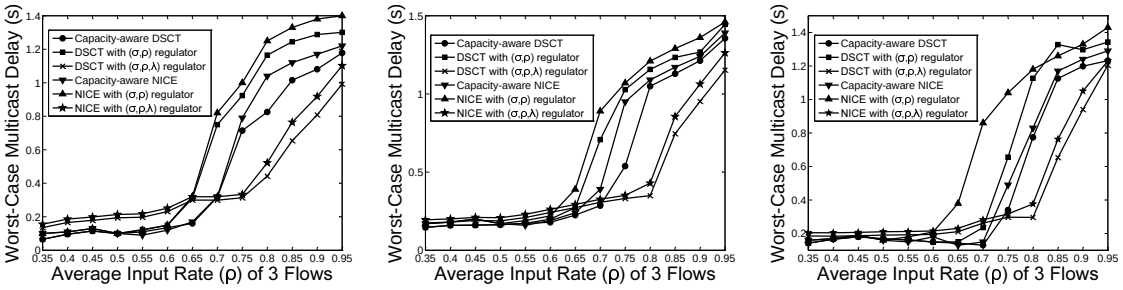
COMPARISON OF TREE LAYER NUMBERS (3 GROUPS WITH HOMOGENEOUS AUDIO STREAMS).

$\bar{\rho}$	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
Capacity-aware DSCT	5	5	5	6	6	6	7	8	7	8	8	9	9
DSCT with (σ, ρ, λ) regulator	6	6	6	6	6	6	6	6	6	6	6	6	6

Fig. 4 (b) illustrates the WDB performances of 3 homogeneous video streams. The rate threshold of 3 flows is 0.67 that is a little less than the theoretical result 0.73 for the fluctuation of network throughput. The maximum improvement in worst-case delays of the (σ, ρ, λ) regulator over the (σ, ρ) regulator is at $\bar{\rho} = 0.8$ and has the value of $\frac{0.72}{0.26} \approx 2.82$. According to Theorem 6, we can also derive $n \approx 1$ from the simulation parameter $K = 3$. Fig. 4 (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 of 3 flows is 0.74 that is a little less than the theoretical value 0.79 in Theorem 3. When $\bar{\rho} \geq 0.74$, the worst-case delays with the (σ, ρ, λ) regulator are much shorter than the ones with the (σ, ρ) regulator. The maximum improvement in the worst-case delay is at $\bar{\rho} = 0.85$ and with the value of $\frac{0.85}{0.27} \approx 3.15$ that meets the theoretical results in Theorem 5 when $n = 1$.

B. 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. Fig. 5 shows



(a) (b) (c)

Fig. 6. The worst-case delay performances when each of the three groups is fed with the same (a) 64Kbps audio stream, (b) 1.5Mbps video streams and (c) one 1.5Mbps video stream and two 64Kbps audio streams.

TABLE II

COMPARISON OF TREE LAYER NUMBERS (3 GROUPS WITH HOMOGENEOUS VIDEO STREAMS).

$\bar{\rho}$	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
Capacity-aware DSCT	6	7	7	8	8	8	9	9	9	10	9	10	10
DSCT with (σ, ρ, λ) regulator	7	7	7	7	7	7	7	7	7	7	7	7	7

the backbone network topology. The 665 group members directly or indirectly through some intermediate network components (e.g., the hubs) attach to the routers in the backbone network and are with the (σ, ρ, λ)/(σ, ρ)-regulated general MUXs. Each group has one real-time flow. Namely, each end 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. In this group of simulations, we compare the WDB performances under the following three EMcast schemes for NICE and DSCT multicast trees respectively: *the capacity-aware multicast tree*, *the multicast tree with* (σ, ρ) *regulator* and *the multicast*

TABLE III

COMPARISON OF TREE LAYER NUMBERS (3 GROUPS WITH HETEROGENEOUS STREAMS).

$\bar{\rho}$	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
Capacity-aware DSCT	5	5	6	6	6	7	7	7	8	8	8	9	9
DSCT with (σ, ρ, λ) regulator	6	6	6	6	6	6	6	6	6	6	6	6	6

tree with (σ, ρ, λ) *regulator.* The traffic pattern are 3 video streams, 3 audio streams and heterogeneous streams (1 video and 2 audio streams) respectively.

Fig. 6 (a) illustrates the worst-case delay performances of *the capacity-aware multicast tree*, *multicast tree*

with (σ, ρ) *regulator* and *multicast tree with* (σ, ρ, λ) *regulator* for NICE and DSCT respectively 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* can achieve shorter delay performances than *DSCT with* (σ, ρ) *regulator*. And, when $\bar{\rho} \geq 0.7$, *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 of 3 flows in the simulation is 0.65 that is a little less than the theoretical value 0.73 in Theorem 8. And, the maximum improvement in the worst-case delays of *DSCT with* (σ, ρ, λ) *regulator* over *DSCT with* (σ, ρ) *regulator* is at $\bar{\rho} = 0.75$ and has the value of $\frac{0.95}{0.27} \approx 3.52$ that meets the theoretical results in Theorem 8 when $n = 1$. Also, from this figure, we can see that the curves of NICE tree with (σ, ρ) regulator, NICE tree with (σ, ρ, λ) regulator, and the capacity-aware NICE tree have the similar comparison trend as the ones of DSCT tree schemes. NICE tree with (σ, ρ, λ) regulator achieves shorter worst-case delay performances than the capacity-aware NICE tree when the average transmission rate becomes high. Furthermore, the curves show that DSCT tree achieves shorter worst-case delays than NICE tree when they employ the same traffic control schemes. The results meet our analysis in [14]. It is mainly because that DSCT employs the hosts' location knowledge to build up the multicast architecture. Fig. 6 (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 $\bar{\rho} \geq 0.7$, *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 at $\bar{\rho} = 0.8$ and with the value of $\frac{1.18}{0.32} = 3.69$. Similar as the curves in Fig. 6 (a), in homogeneous video communications, NICE

tree with (σ, ρ, λ) regulator achieves shorter 26 worst-case delay performances than the capacity-aware NICE tree when the average transmission rate becomes high. And, the worst-case delays of NICE tree in each traffic control scheme are longer than the corresponding ones of DSCT tree. Fig. 6 (c) gives the worst-case delay performance comparison when one group is fed with the 1.5Mbps video stream and each of other two groups is fed with the 64Kbps audio stream. The simulation results also tell us that the rate threshold of 3 flows is 0.735 that is a little less than the theoretical 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 at $\bar{\rho} = 0.8$ and has the value of $\frac{1.15}{0.27} \approx 4.26$. Also, the comparison of NICE tree schemes and DSCT tree schemes shows the similar trend as the one in Fig. 6 (a).

Table I gives the comparison of multicast tree layer numbers when the 3 groups are with homogeneous audio streams. Data in the table show that DSCT with (σ, ρ, λ) regulator achieves shorter delay performances without increasing the tree height. But, the height of the capacity-aware DSCT increases with the increment of average input rate. Tables II and III are the comparison of multicast tree layer numbers when the 3 groups are with homogeneous video streams and heterogeneous streams respectively. Similar to the results in Table I, data in these two tables prove that the (σ, ρ, λ) regulator reduces the worst-case delay without increasing the tree height.

VII. 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 without

increasing the multicast tree height. Through using *network calculus*, we proved a set of theorems on the *input rate threshold* ρ^* above which the (σ, ρ, λ) regulator benefits the shorter delay performances, the *worst-case delay bound* for the (σ, ρ, λ) regulator achieves and the *improvement of worst-case delay bound* of (σ, ρ, λ) regulator over (σ, ρ) regulator for single end host and EMcast with homogeneous and heterogeneous flows respectively.

We then study our algorithm in the simulation environments. We ran two groups of simulations with the single regulated end host topology and the EMcast topology respectively. We observed the worst-case delay improvement of (σ, ρ, λ) regulator over (σ, ρ) regulator and the rate threshold. The simulation results meet our theoretical analysis. Therefore, the possible bottleneck in multi-group network can be avoided without increasing the lengths of multicast paths. When the flow co-exist with other traffic, the number of input traffic at the end host is changed and the flows' average input rate may be increased or decreased for the changed traffic load. Such change influences the values of each flow's working period and vacation period. But, we think that the same process of adaptive control algorithm can be implemented to control the traffic and its co-existed flows when the traffic priority is ignored. When the traffic priority is considered, we should extend our algorithm to deal with the flows with different priorities. For example, adding new parameters into (σ, ρ, λ) regulator to enable it to recognize and process flows with different priorities. In the next step, we also propose to test our algorithm in the real world network environment (e.g., PlanetLab). And, to study the algorithms on other QoS requirements (e.g., error control and packet loss) in multicast communications through theorems and simulations is our nearly future work.

ACKNOWLEDGMENTS

The authors are grateful to the editor and reviewers for their invaluable and helpful comments. The authors thanks Prof. Hanxing Wang for help with mathematics, and Mr. Mark O'Brien for help with English.

This work is supported by Embark Postdoctoral Fellowship of Ireland with the funding code: 501-et-504 4890.

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] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

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

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

[18] 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.

[19] 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.

[20] 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.

[21] 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.

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

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

[24] 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.

[25] 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.

[26] 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.

[27] M. R. Prez, M. F. Veiga, C. L. Garcła, J. C. L. Ardao, and S. H. Alonso. An Open-Loop Multicast Layered Congestion Protocol for Real-Time Multimedia Transmission. In Proc. of IEEE 2003 Global Communications Conference (Globecom 2003), pp. 2855-2859, December 1-5, 2003, San Francisco, CA, USA.

[28] L. Rizzo. Pgmcc: A TCP-friendly Single-Rate Multicast Congestion Control Scheme. In Proc. Of ACM SIGCOMM 2000.

[29] J. Widmer, and M. Handley. Extending Equation-based Congestion Control to Multicast Applications. In Proc. of ACM SIGCOMM 2001, August 2001.

[30] A. Chaintreau, F. Baccelli, and C. Diot. Impact of Network Delay Variation on Multicast Sessions Performance with TCP-Like Congestion Control. IEEE Transactions on Networking, 2002.

[31] Y. Amir, and C. Danilov. Reliable Communication in Overlay Networks. IEEE DSN 2003.

[32] G. U. Keller, and E. W. Biersack. A Congestion Control Model for Multicast Overlay Networks and Its Performance. In Proc. of NGC 2002, October 2002.

[33] Y. Amir, B. Awerbuch, C. Danilov, and J. Stanton. Global Flow Control for Wide Area Overlay Networks: A Cost-Benefit Approach.

IEEE OpenArch 2002.

[34] G. I. Kwon, and J. W. Byers. ROMA: Reliable Overlay Multicast with Loosely Coupled TCP Connections. Boston University, INFOCOM 2004.

[35] J. Lu, Q. Ruan, and R. Ni. A Scalable Overlay Multicast Congestion Control for Multimedia Streaming. In Proc. of The 30th IEEE Conference on Local Computer Networks (LCN), November 15-17, 2005, Sydney, Australia.

[36] M. Handley, S. Floyd, J. Padhye, and J. Widmer. TCP Friendly Rate Control (TFRC): Protocol Specification. Request for Comments (RFC): 3448, Network Working Group, January 2003.

Wanqing Tu is a Postdoctoral Researcher of Computer Science at University College Cork in Ireland. She received her PhD degree from the Department of Computer Science at the City University of Hong Kong in 2005. Her research interests include QoS, overlay networks, wireless mesh networks, end host multicast, and distributed computing. She received the Embark Postdoctoral Fellowship of Ireland. She received one Best Paper Awards as a co-author. She is a member of the IEEE.

Cormac J. Sreenan (M'93) is Professor of Computer Science at University College Cork (UCC) in Ireland. Prior to joining UCC in 1999 he was on the research staff at AT & T Labs - Research, Florham Park, NJ, and at Bell Labs, Murray Hill, NJ. At UCC he directs the Mobile Internet Systems Laboratory (MISL), which is a group of over 10 research staff and students with research activity in multimedia and wireless networking and systems, including wireless sensor networks and media content distribution. Prof. Sreenan is currently on the Editorial Boards of ACM/Springer Multimedia Systems Journal and Elsevier Computer Networks Journal. In the past he has served as Guest Editor for Communications of the ACM, IEEE Journal on Selected Areas in Communications and IEEE Wireless Communications Magazine. During 2007 he will host and Co-Chair the Workshop on Embedded Networked Sensors (EmNets) and the Conference on Mobile & Wireless Communication Networks (MWCN). He has a Ph.D. in Computer Science from Cambridge University and is a Fellow of the British Computer Society.

Weijia Jia received his BSc, MSc, and PhD degrees, all in computer science, in 1982, 1984, and 1993 from Center South University, China and Polytechnic Faculty of Mons, Belgium. He is an associate professor in Dept. of Computer Science, City University of Hong Kong. His research interests include computer network, distributed systems, multicast and anycast QoS routing protocols for Internet and wireless communications. He is a member of the IEEE.